

FUNDAÇÃO OSWALDO CRUZ
INSTITUTO AGGEU MAGALHÃES
DOUTORADO EM BIOCÊNCIAS E BIOTECNOLOGIA EM SAÚDE

GUSTAVO BARBOSA DE LIMA

Contribuições ao estudo de múltiplos complexos atuantes na iniciação da
tradução em *Leishmania sp.*:
identificação de interações proteicas e mecanismos de regulação

Recife
2021

Gustavo Barbosa de Lima

Contribuições ao estudo de múltiplos complexos atuantes na iniciação da tradução em *Leishmania sp.*: identificação de interações proteicas e mecanismos de regulação

Tese apresentada ao Doutorado em Biociências e Biotecnologia em Saúde do Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, para obtenção do grau de Doutor em Ciências.

Área de concentração: Biologia celular e molecular de patógenos, vetores e hospedeiros.

Orientador: Dr. Osvaldo Pompílio de Melo Neto

Coorientador: Dr. Christian Robson de Souza Reis

Recife

2021

**Catálogo na fonte: Biblioteca Luciana Borges Abrantes dos Santos
Instituto Aggeu Magalhães / Fundação Oswaldo Cruz**

L732c Lima, Gustavo Barbosa de.
Contribuições ao estudo de múltiplos complexos atuantes na iniciação da tradução em *Leishmania* sp.: identificação de interações proteicas e mecanismos de regulação / Gustavo Barbosa de Lima. — Recife: [s.n.], 2022.
140 p. : il., graf., tab., 30 cm.

Tese (doutorado) – Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2022.

Orientador: Osvaldo Pompílio de Melo Neto;
coorientador: Christian Robson de Souza Reis

Programa de Pós-graduação em Doutorado em Biotecnologia e Biotecnologia em Saúde

1. 1. *Leishmania infantum* - crescimento & desenvolvimento. 2. Fator de Iniciação 2 em Eucariotos. 3. Fator de Iniciação 3 em Eucariotos. 4. Fator de Iniciação 4F em Eucariotos. 5. Leishmaniose - parasitologia. 6. Análise de Sequência de Proteína. I. Melo Neto, Osvaldo Pompílio de. II. Reis, Christian Robson de Souza. III. Título.

CDU 616.92

Gustavo Barbosa de Lima


Contribuições ao estudo de múltiplos complexos atuantes na iniciação da tradução em *Leishmania sp.*: identificação de interações proteicas e mecanismos de regulação

Tese apresentada ao Doutorado em Biociências e Biotecnologia em Saúde do Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, para obtenção do grau de Doutor em Ciências.


Área de concentração: Biologia celular e molecular de patógenos, vetores e hospedeiros.

Aprovado em: 30 de outubro de 2021.


BANCA EXAMINADORA



Dr. Osvaldo Pompílio de Melo Neto (Orientador)
Instituto Aggeu Magalhães/Fundação Oswaldo Cruz



Dr. Rafael Dhalia (Examinador interno)
Instituto Aggeu Magalhães/Fundação Oswaldo Cruz



Dr. Roberto Dias Lins Neto (Examinador interno)
Instituto Aggeu Magalhães/ Fundação Oswaldo Cruz

Dra. Fabíola Barbieri Holetz (Examinadora externa)
Instituto Carlos Chagas/ Fundação Oswaldo Cruz

Dra Janaina de Freitas Nascimento (Examinadora externa)
Universidade de São Paulo

*A minha mãe, Rosa (in memoriam), que com muito
carinho e apoio, não mediu esforços para que eu
chegasse até esta etapa da minha vida.
Saudades mãe.*

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me dado a oportunidade e a capacidade intelectual para realizar esse projeto.

Agradeço ao meu Orientador Osvaldo Pompílio de Melo Neto, por ser um grande exemplo de profissional e pai a se seguir, sou muito grato pela paciência e empatia em momentos difíceis como foram estes em que houve pandemia e outras dificuldades, muito obrigado.

Sou grato de forma única a Adalúcia, minha parceira em todas as ideias loucas que tenho, nunca largou a minha mão, é a pessoa mais importante na minha vida e que vou lhe ser grato enquanto eu viver por me aturar.

Gostaria de agradecer ao meu coorientador, psicólogo, conselheiro, coach, no qual eu pude contar em momentos de dificuldades durante esta jornada, Christian Reis, assim como a Éden, grande amigo e meu segundo coorientador na vida científica.

Agradeço a todos que passaram pela minha vida durante esses anos de doutorado, principalmente os colegas de bancada, orientadora não oficial Dani, Adriana, Ludmila, Moézio, Maria, Wagner, Camila, Matheus, Artur, Renata, Rafa, Allana, as minhas orientadoras de iniciação científica, Amaranta e Larissa vocês tornaram um pouco mais leve a caminhada, amo vocês.

Agradeço ao meu Pai e aos meus irmãos que sempre se mantiveram presente em minha vida e que torcem pelo meu sucesso.

Agradeço a mãe que tive, muito do que sou hoje eu devo a ela, e mesmo não me acompanhando até o fim, sei que torce sempre por mim.

Ao órgão de fomento FACEPE, pelo apoio financeiro concedido, ao IAM-FIOCRUZ pela infraestrutura e ao PPGBBS – IAM pelo suporte acadêmico

Existem Coisas que você não consegue enxergar se não mudar seu ponto de vista.

Trafalgar D. Water Law, One piece.

RESUMO

DE LIMA, Gustavo Barbosa. **Contribuições ao estudo de múltiplos complexos atuantes na iniciação da tradução em *Leishmania sp.*: identificação de interações proteicas e mecanismos de regulação.** 2021. Tese (Doutorado em Biociências e Biotecnologia em Saúde) – Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2021.

A família Trypanosomatidae de protozoários parasitas apresenta características moleculares diferenciadas, como a regulação da sua expressão gênica feita em nível pós-transcricional. Como em outros eucariotos, acredita-se que a iniciação da tradução, mediada por diferentes fatores de iniciação da tradução (eIFs), seja uma etapa crítica dessa regulação. Etapas importantes desse processo são o reconhecimento do mRNA e tRNAs, o recrutamento dos ribossomos e a identificação do códon AUG de iniciação da tradução, atividades realizadas por diversos eIFs como os complexos eIF2, eIF3 e eIF4F. O eIF2, formado pelas subunidades eIF2 α , eIF2 β e eIF2 γ , leva o tRNA ligado a metionina inicial da proteína nascente à subunidade menor ribossomal (40S). Já o eIF3, formado por 13 subunidades em mamíferos (eIF3a até eIF3m), com múltiplas funções na tradução, atua ligado a subunidade 40S e facilitando sua associação ao mRNA. O eIF4F, por sua vez, é responsável por reconhecer o mRNA e facilitar o recrutamento dessa subunidade ribossomal através de sua interação com o eIF3. Esse complexo é formado pelas subunidades eIF4E, eIF4A e eIF4G. Os tripanosomatídeos se destacam por apresentarem particularidades específicas em cada um destes complexos e que os distinguem dos demais eucariotos, como dois homólogos de eIF2 γ , cinco eIF4Es formando diferentes tipos de complexos do tipo eIF4F e um complexo eIF3 com 12 subunidades. Este trabalho buscou avançar no conhecimento sobre cada um destes complexos a partir de estudos voltados a caracterização de subunidades selecionadas de cada complexo (EIF2 γ -1 e EIF2 γ -1; seis subunidades do eIF3; EIF4E3 e EIF4E5), avaliando interações proteicas a partir de duas espécies de *Leishmania*. Estas proteínas foram estudadas a partir de sua expressão em linhagens transgênicas fusionadas ao epítipo HA. Cada proteína alvo foi utilizada em experimentos de imunoprecipitação, com parceiros co-precipitados identificados por espectrometria de massa. Foram definidas redes de interações proteicas para cada um destes complexos. Assim, diferenças importantes foram identificadas entre proteínas parceiras dos dois homólogos de eIF2 γ , sugerindo mecanismos de atuação distintos. As subunidades do complexo eIF3 apresentam interações entre si semelhantes a outras espécies, mas interações distintas com outras proteínas parceiras. Para o EIF4E3 foi observado a relevância da sua fosforilação sobre todo o seu perfil de interação com proteínas parceiras. Já para o EIF4E5, a resolução da sua estrutura, nos trouxe novas informações sobre como seu modo de ação e interação com parceiros proteicos. Os resultados gerados contribuem no conhecimento sobre a tradução nos tripanosomatídeos e sobre como essa pode ser regulada, subsidiando ainda o desenvolvimento de moléculas inibidoras específicas.

Palavras chaves: eIFs; expressão gênica; tripanosomatídeos.

ABSTRACT

DE LIMA, Gustavo Barbosa. **Contribuições ao estudo de múltiplos complexos atuantes na iniciação da tradução em *Leishmania* sp.: identificação de interações proteicas e mecanismos de regulação.** 2021. Tese (Doutorado em Biociências e Biotecnologia em Saúde) – Instituto Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2021.

The family Trypanosomatidae of parasitic protozoa is characterized by unique molecular features, such as their gene expression regulation being carried out at the post-transcriptional level. As in other eukaryotes, translation initiation, mediated by different translation initiation factors (eIFs), is believed to be a critical step in this regulation. Important steps in this process are the recognition of mRNA and tRNAs, the recruitment of ribosomes and the identification of the translation initiation codon AUG, activities performed by several eIFs, such as the eIF2, eIF3 and eIF4F complexes. eIF2, formed by the eIF2 α , eIF2 β and eIF2 γ subunits, delivers the tRNA bound to the first methionine of the nascent protein to the smaller ribosomal subunit (40S). eIF3, formed by 13 subunits in mammals (eIF3a to eIF3m), with multiple functions in translation, acts bound to the 40S subunit, facilitating its association with the mRNA. eIF4F, is responsible for recognizing the mRNA and facilitating the recruitment of this ribosomal subunit through its interaction with eIF3. This complex is formed by the eIF4E, eIF4A and eIF4G subunits. Trypanosomatids are noteworthy for specific features found in each of these complexes and that distinguish them from other eukaryotes, such as two eIF2 γ homologues, five eIF4Es forming distinct eIF4F-like complexes and a 12 subunit eIF3. This work sought to advance the knowledge on each of these complexes from studies aimed at the characterization of selected subunits from each complex (EIF2 γ -1 and EIF2 γ -1; six eIF3 subunits; EIF4E3 and EIF4E5), evaluating protein interactions from two *Leishmania* species. These proteins were studied from their expression in transgenic lines fused to the HA epitope. Each target protein was used in immunoprecipitation experiments, with co-precipitated partners identified by mass spectrometry. Protein interaction networks were defined for each of these complexes. Thus, important differences were identified between partner proteins of the two eIF2 γ homologues, suggesting distinct mechanisms of action. The subunits of the eIF3 complex have interactions with each other similar to other species, but distinct interactions with other protein partners. For EIF4E3 the relevance of its phosphorylation on its entire interaction profile with partner proteins was observed. As for the EIF4E5, the resolution of its structure produced new information regarding its mode of action and interaction with protein partners. The results generated contribute to the knowledge about translation in trypanosomatids and how it can be regulated, further supporting the development of specific inhibitor molecules.

Key Words: eIFs; regulation of gene expression; Trypanosomatids.

LISTA DE FIGURAS

Figura 1 - Árvore filogenética esquemática representando os super-filos principais do reino protozoa, nomeados Alveolata e Excavata	20
Figura 2 - Estrutura do kinetoplasto.	22
Figura 3 - Filogenia global Kinetoplastida.	22
Figura 4 - Arquitetura do mapa do cromossomo 26 de <i>Leishmania infantum</i>	25
Figura 5 - Representação esquemática da via de iniciação da tradução eucariótica.	27
Figura 6 - Esquema que descreve o papel de eIF2 na iniciação da tradução.	29
Figura 7 - Interações do eIF2 com o tRNA Iniciador e o eIF2B.	30
Figura 8 - O núcleo PCI/MPN do complexo eIF3 de mamíferos.	32
Figura 9 - Estrutura tridimensional do fator eIF4E3 em humano.	34
Figura 10 - Análise da estrutura do eIF3 de mamíferos e <i>Trypanosoma cruzi</i>	38
Figura 11 - Estrutura química do <i>cap4</i>	41
Figura 12 - Representação esquemática dos homólogos de EIF4E3 e EIF4E4 de <i>Trypanosoma brucei</i> e <i>Leishmania major</i>	44
Figura 13 - Esquema do complexo EIF4E5-EIF4G1-G1-IP-GI-IP2.	47
Figura 14 - Sítio de interação ao <i>cap4</i> e o fator EIF4E5 de <i>Trypanosoma cruzi</i>	48
Figura 15 - Mapa do vetor plasmidial de expressão pSP-BT1-Y-Neo-alfa-HA.	55
Figura 16 - Representação do mecanismo do ensaio de termoforese em microescala.	61
Figura 17 - Alinhamento esquemático dos domínios encontrados em fatores da família EIF2 γ	64
Figura 18 - Alinhamento dos homólogos EIF2 γ -1 e EIF2 γ -2 com espécies de Kinetoplastida e espécies modelo utilizando o Pfam.	65
Figura 19 - Arvore filogenética dos homólogos EIF2 γ 1 e EIF2 γ 2 com outras espécies de kinetoplastidae e espécies modelo.	66
Figura 20 - Western blot de imunoprecipitação de fatores EIF2 γ 1 e EIF2 γ 2.	67
Figura 21 - Diagrama de Venn de interação com os fatores EIF2 γ 1 e EIF2 γ 2.	68
Figura 22 - <i>Western blot</i> de imunoprecipitação de fatores EIF3.	69
Figura 23 - Mapa de co-precipitação de subunidades do complexo eIF3 com outras subunidades desse mesmo complexo.	71

Figura 24 - Mapa de co-precipitação de subunidades do complexo eIF3 com homólogos selecionados fatores de iniciação da tradução.	72
Figura 25 - Mapa de co-precipitação de subunidades do complexo eIF3 com proteínas de ligação ao RNA.....	73
Figura 26 - Mapa de co-precipitação de subunidades do complexo eIF3 com proteínas kinases, RNA helicases e proteínas ribossomais.	75
Figura 27 - Ensaio de <i>western-blot</i> avaliando a expressão em <i>Leishmania amazonensis</i> da proteína EIF4E3 fusionada ao epítipo HA e dois mutantes.....	76
Figura 28 - Ensaio de <i>western-blot</i> para avaliação de imunoprecipitação do fator EIF4E3 de <i>Leishmania amazonensis</i> expressando o EIF4E3 e seu mutante S75A, ambos fusionados ao epítipo HA.	77
Figura 29 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com fatores de iniciação da tradução, helicases, proteínas de ligação ao RNA, quinases e NTPases.....	79
Figura 30 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com proteínas ribossomais, ubiquitinas, proteínas do proteasoma e nucleares.	80
Figura 31 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com proteínas transportadoras, mitocondriais, de atividade metabólica e hipotéticas.	81
Figura 32 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com proteínas diversas.	82
Figura 33 - Distribuição das proteínas que co-precipitaram com o fator EIF4E3.	83
Figura 34 - Expressão do fator EIF4E5 em <i>Escherichia coli</i> após indução.....	85
Figura 35 - Purificação do fator EIF4E5.	86
Figura 36 - Cromatograma de purificação da proteína EIF4E5 por cauda de histidina.	87
Figura 37 - Eletroforese das frações do cromatograma de purificação por afinidade a cauda de histidina do fator EIF4E5.....	87
Figura 38 - Cromatograma de purificação da proteína EIF4E5 por troca iônica.....	88
Figura 39 - Eletroforese das frações do cromatograma de purificação por troca iônica do fator EIF4E5.	89
Figura 40 - Cromatograma de purificação da proteína EIF4E5 por exclusão de tamanho.	90

Figura 41 - Eletroforese das frações do cromatograma de purificação por exclusão de tamanho do fator EIF4E5.	91
Figura 42 - Formação dos cristais através de soluções precipitantes.	92
Figura 43 - Refinamento de cristais da proteína recombinante EIF4E5 através de soluções precipitantes específicas.	92
Figura 44 - Análise estrutural do EIF4E5 de <i>Leishmania major</i> na forma apo.	95
Figura 45 - Análise da interface eIF4E – eIF4G.	96
Figura 46 - Análise das superfícies de interação putativas EIF4E5 dos tripanosomatídeos.....	97

LISTA DE TABELAS

Tabela 1 - Homólogos do fator eIF4E em diferentes espécies.....	35
Tabela 2 - Sequências de peptídeos da proteína alvo por espectrometria de massas do fator EIF4E3.	84
Tabela 3 - Dados cristalográficos e estatísticas de refinamento.	93

LISTA DE ABREVIATURAS E SIGLAS

BLAST	<i>Basic Local Alignment Search Tool</i> - Ferramenta Básica De Pesquisa De Alinhamento Local
CID	Collision Induced Dissociation - dissociação induzida por colisão
DP	Lisado Depletado
ECL	<i>Enhanced Chemiluminescence</i> - reação de quimioluminescência
eIFs	<i>Eukaryotic Initiation Factor</i> - fatores de iniciação eucarióticos
EXT	Extrato Total
FDR	False discovery rate – razão de falsa descoberta
FPLC	<i>Fast Protein Liquid Chromatography</i> - rápida cromatografia líquida de proteínas
GDP	<i>Guanosine Diphosphate</i> - difosfato de guanosina
GEF	<i>Guanine Nucleotide Exchange Factor</i> - fator de troca de nucleotídeos de guanina
GTP	<i>Guanosine Triphosphate</i> - trifosfato de guanosina
HA	<i>Human Influenza Hemagglutinin</i> - hemaglutinina da influenza humana
IP	Imunoprecipitação
IPTG	Isopropil β -d-1-Tiogalactopiranósido
IR	<i>Infra Red</i> - Laser Infravermelho
kDa	KiloDaltons
LC	Leishmaniose Cutânea
LC-MSMS	<i>Liquid Chromatography with tandem Mass Spectrometry</i> - cromatografia líquida acoplada a espectrometria de massas
LFQ	Label-free quantitation – marcação de quantificação livre
Lis	Lisado Celular
LMC	Leishmaniose Mucocutânea
LV	Leishmaniose Visceral
MAFFT	<i>Multiple Alignment using Fast Fourier Transform</i> - alinhamento múltiplo usando transformada rápida de Fourier
MPN	<i>Mpr1-Pad1 N-Terminal</i>
MST	<i>Microscale Thermophoresis</i> - termoforese em microescala

NCBI	<i>National Center For Biotechnology Information</i> – Centro Nacional de Informações Sobre Biotecnologia
OMS	Organização Mundial da Saúde
ORF	<i>Open Reading Frame</i> - quadro de leitura aberta
PABP	<i>Poly(A)-Binding Protein</i> - Proteína de Ligação a Cauda Poli-A
PEG	Polietilenoglicol
PCR	<i>Polymerase Chain Reaction</i> - reação em cadeia da polimerase
PIC	<i>Preinitiation Complex</i> - complexo de pré-iniciação
RMSD	<i>Root Mean Square Deviation</i> - desvio da raiz média quadrada
SDSPAGE	<i>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</i> - eletroforese em gel de poliacrilamida com dodecil sulfato de sódio
SFB	Soro Fetal Bovino
SL	<i>Spliced Leader</i> – sequência líder
SLBP	<i>Stem-Loop Binding Protein</i> - proteína de ligação grampo do RNA da Histona
TBS	<i>Tris-Buffered Saline</i> – solução salina tamponada com Tris
TC	<i>Ternary Complex</i> - complexo ternário
UTR	<i>Untranslated Region</i> - região não traduzida (5')
WT	<i>Wild Type</i> - tipo selvagem

SUMÁRIO

1	INTRODUÇÃO	18
2	REFERENCIAL TEÓRICO.....	20
2.1	Características gerais da família Trypanosomatidae.....	20
2.1.1	<i>Ordem Kinetoplastida</i>	<i>20</i>
2.1.2	<i>Família Trypanosomatidae</i>	<i>22</i>
2.2	Expressão gênica em tripanosomatídeos.....	24
2.3	Iniciação da tradução eucariótica	25
2.4	Visão geral do papel de eIFs selecionados na iniciação da tradução	28
2.4.1	<i>Complexo eIF2</i>	<i>28</i>
2.4.2	<i>Complexo eIF3.....</i>	<i>31</i>
2.4.3	<i>Complexo eIF4F.....</i>	<i>33</i>
2.4.3.1	Superfamília eIF4E.....	33
2.5	eIFs de tripanosomatídeos	36
2.5.1	<i>Complexo eIF2</i>	<i>36</i>
2.5.2	<i>Complexo eIF3.....</i>	<i>37</i>
2.5.3	<i>Complexos do tipo eIF4F em tripanosomatídeos</i>	<i>39</i>
2.5.3.1	Homólogos de eIF4E.....	41
2.5.3.1.1	Grupo 1: EIF4E1 e EIF4E2.....	42
2.5.3.1.2	Grupo 2: EIF4E3 e EIF4E4.....	43
2.5.3.1.3	Grupo 3: EIF4E5 e EIF4E6.....	46
3	JUSTIFICATIVA.....	50
4	OBJETIVO GERAL.....	51
4.1	OBJETIVOS ESPECÍFICOS	51
5	METODOLOGIA	52
5.1	Análises de sequência e filogenia	52
5.2	Cultivo e avaliação de crescimento celular.....	53
5.3	Transfecção dos parasitas	53
5.4	Construções plasmidiais dos fatores de iniciação da tradução	54
5.5	Sequenciamento de DNA	55
5.6	Ensaio de Western-Blot.....	56
5.7	Preparação de lisados celulares e ensaios de imunoprecipitação	56

5.8	Análise por cromatografia líquida acoplada a espectrometria de massas (LC-MS/MS)	57
5.9	Cristalografia do fator EIF4E5	58
5.9.1	<i>Purificação de proteínas com cauda histidina</i>	58
5.9.2	<i>Gel filtração</i>	59
5.9.3	<i>Cristalização</i>	59
5.10	Ensaio de interação entre o fator EIF4E5 e os <i>cap0</i> e <i>cap4</i> através de termoforese em microescala (MST)	60
6	RESULTADOS	62
6.1	Análise comparativa dos dois homólogos de EIF2 γ de <i>L. infantum</i>	63
6.1.1	Alinhamento dos EIF2 γ 1 e EIF2 γ 2 de espécies de kinetoplastidas e espécies representativas	63
6.1.2	Identificação de parceiros funcionais dos EIF2 γ 1 e EIF2 γ 2	66
6.2	Análises de interação com parceiros funcionais do complexo eIF3	68
6.2.1	Expressão heteróloga de subunidades do eIF3 de <i>Leishmania infantum</i>	68
6.2.2	Análise de espectrometria de massas de proteínas do complexo eIF3	69
6.3	Análises de interação com parceiros funcionais do fator EIF4E3	75
6.3.1	Expressão heteróloga do EIF4E3 e seus mutantes	75
6.3.2	Lisados celulares e imunoprecipitação do EIF4E3 nativo e mutante S75A	77
6.3.3	Identificação de parceiros do EIF4E3 e avaliação do impacto da mutação S75A sobre sua capacidade de interação a parceiros selecionados	78
6.4	Caracterização estrutural do fator EIF4E5	84
6.4.1	Expressão e purificação do fator EIF4E5 com cauda de Histidina	85
6.4.1.1	Expressão e purificação do EIF4E5 por FPLC por afinidade, troca iônica e filtração em gel	86
6.4.1.2	Purificação por afinidade a histidina	86
6.4.1.3	Purificação por troca iônica	88
6.4.1.4	Purificação por gel filtração	89
6.4.2	Cristalização do EIF4E5 recombinante	91
6.4.3	Refinamento estatístico e dados cristalográficos	93
6.4.4	Análise do fator EIF4E5 com interface ao EIF4G	94
6.4.5	Análise comparativa da superfície de interação do fator EIF4E5	97

6.5	Ensaio de interação entre o fator EIF4E5 e o <i>cap4</i> através de termoforese em microescala (MST)	98
7	DISCUSSÃO	100
8	CONCLUSÕES	109
9	REFERENCIAS	110
	APÊNDICE A – ARTIGO PUBLICADO	120

1 INTRODUÇÃO

A família Trypanosomatidae é formada por protozoários flagelados que incluem os gêneros *Leishmania* e *Trypanosoma*, agentes causadores de enfermidades de importância para saúde pública, em especial em países em desenvolvimento. Juntas, as doenças provocadas por esses patógenos (Leishmaniose, a Doença de Chagas e a Doença do Sono) chegam a afetar cerca de 21 milhões de pessoas, mas métodos de prevenção e tratamento para essas enfermidades se mostram pouco efetivos, custosos ou possuem diversos efeitos colaterais. Portanto ainda existe uma grande necessidade de uma maior compreensão sobre mecanismos básicos da biologia molecular desses parasitas para se buscar métodos mais efetivos de combate.

Os tripanosomatídeos são organismos que divergiram precocemente durante a evolução e possuem características biológicas distintas dos demais eucariotos, como a complexa regulação da sua expressão gênica. Esta ocorre em etapas pós-transcricionais, como no processamento, transporte e estabilidade dos RNAs mensageiros (mRNAs) e ainda durante sua tradução em proteínas. Existem indícios de que o controle da tradução de mRNAs tenha um papel importante na regulação da expressão gênica, porém os mecanismos deste controle não são adequadamente conhecidos. Em eucariotos mais bem caracterizados, a maioria dos eventos de controle da expressão gênica ao nível da tradução ocorre na sua etapa de iniciação, quando os mRNAs são selecionados por complexos proteicos associados aos fatores de iniciação da tradução (eIFs de *eukaryotic Initiation Factors*) e levados aos ribossomos.

Dentre os fatores de iniciação da tradução podemos destacar vários complexos formados por múltiplas subunidades, como os eIF2, eIF3 e eIF4F. O complexo heterotrimérico eIF2, formado pelas subunidades eIF2 α , eIF2 β e eIF2 γ , tem atividade de ligação ao GTP e sua principal função é entregar o tRNA ligado a primeira metionina da proteína a ser sintetizada (Met-tRNA_i) à subunidade menor ribossomal (40S). Sua ação se dá em todas as células eucarióticas e garante que a síntese proteica comece pelo códon AUG correto. Já o eIF3, envolvido em quase todas as etapas da tradução, atua na iniciação associado a subunidade ribossomal

40S, junto com os eIF2 (com o Met-tRNAⁱ), eIF1 e eIF1A. O eIF3 é o maior dos eIFs e em mamíferos é composto por 13 subunidades não idênticas (eIF3a até eIF3m). Entre suas funções está a de mediar a interação entre a subunidade 40S e o mRNA e a de atuar na varredura do mRNA para o reconhecimento do códon AUG de iniciação da tradução. O complexo eIF4F típico, por sua vez, formado pelos fatores eIF4E, eIF4G e eIF4A, é responsável pelo reconhecimento dos mRNAs e, junto com o eIF3, media a interação entre estes e a subunidade ribossomal 40S.

Nos tripanosomatídeos, características próprias foram observadas em diferentes fatores de iniciação da tradução e que podem estar associadas ao controle da sua função. No caso do complexo eIF2, sua subunidade eIF2 γ é representada por dois homólogos distintos, chamados de EIF2 γ -1 e EIF2 γ -2, embora pouco se sabe sobre como se diferenciam funcionalmente. No caso do eIF3, poucos estudos tem avaliado seu papel e o subunidades individuais na tradução, embora apenas 12 subunidades (EIF3A até EIF3L) tenham sido identificadas, todas representadas por um único homólogo. Já no que diz respeito ao complexo eIF4F, foram descobertos nos tripanosomatídeos vários homólogos para suas subunidades eIF4E (EIF4E1 a 6) e eIF4G (EIF4G1 a 5), formando múltiplos complexos do tipo eIF4F atuantes na tradução.

Este trabalho se propôs a investigar diferenças e semelhanças nos complexos eIF2, eIF3 e eIF4F de tripanosomatídeos a partir do estudo de subunidades selecionadas e previamente identificadas. Primeiramente, foram estudados os dois homólogos de EIF2 γ , avaliando suas interações e informações sobre como atuam na tradução. A análise realizada para evidenciar interações do eIF3, a partir do estudo de seis de suas subunidades, nos trouxe uma maior clareza das interações entre si e com novas ou conhecidas proteínas parceiras. Estudos semelhantes com os EIF4E3 e EIF4E5, além da análise estrutural focada apenas no EIF4E5, revelaram importantes aspectos sobre o modo de atuação destas proteínas e complexos nas quais atuam, inclusive no que concerne a regulação de suas atividades por modificações pós-traducionais do tipo fosforilação. Importantes contribuições foram então realizadas no estudo de cada uma dessas proteínas e complexos associados com impacto importante no conhecimento existente a respeito da iniciação da tradução em tripanosomatídeos.

2 REFERENCIAL TEÓRICO

2.1 Características gerais da família Trypanosomatidae

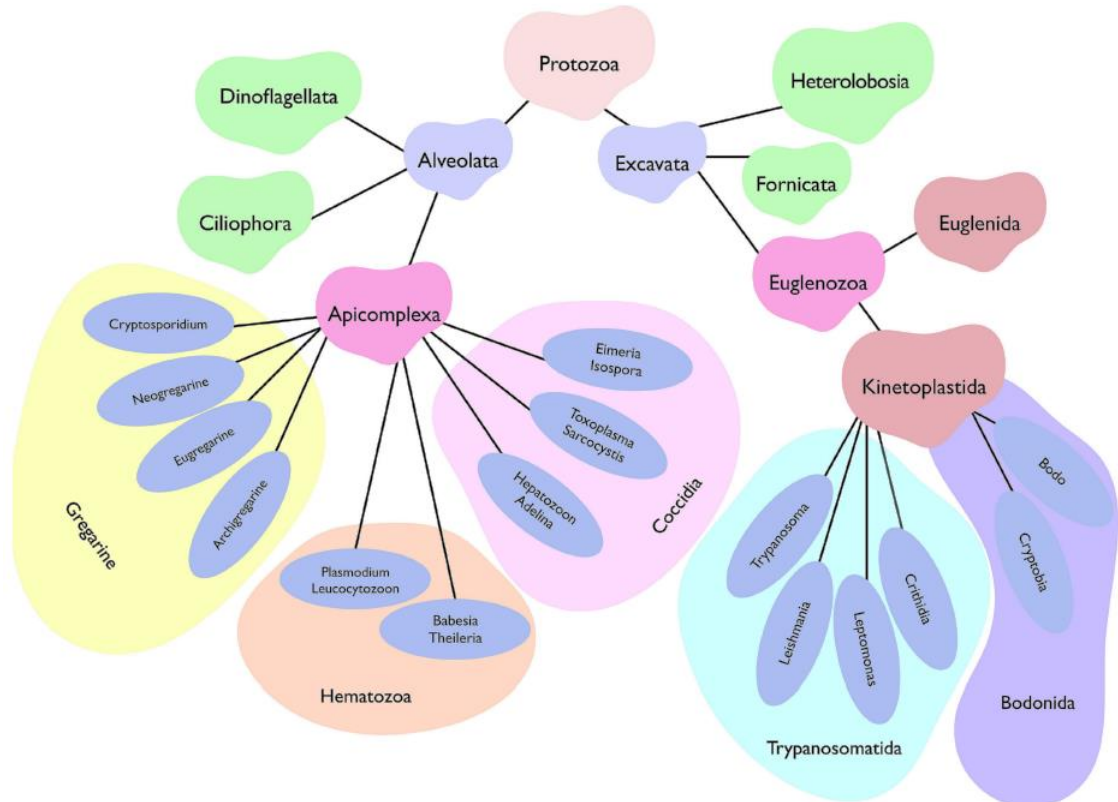
A família Trypanosomatidae possuem características celulares e moleculares que os tornam organismos com funções e metabolismo distinto de outros eucariotos.

2.1.1 *Ordem Kinetoplastida*

O Reino *Protozoa* possui mais de 40.000 espécies existentes de seres unicelulares, das quais, cerca de 25.000 são espécies de vida livre, enquanto as demais possuem um estilo de vida parasítico (REN; GUPTA, 2017). Dentro deste Reino existe um supergrupo denominado Excavata e neste se encontra o filo Euglenozoa onde se encontra a ordem Kinetoplastida (STUART; BRUN; CROFT; FAIRLAMB *et al.*, 2008) (

Figura 1).

Figura 1 - Árvore filogenética esquemática representando os super-filos principais do reino protozoa, nomeados Alveolata e Excavata

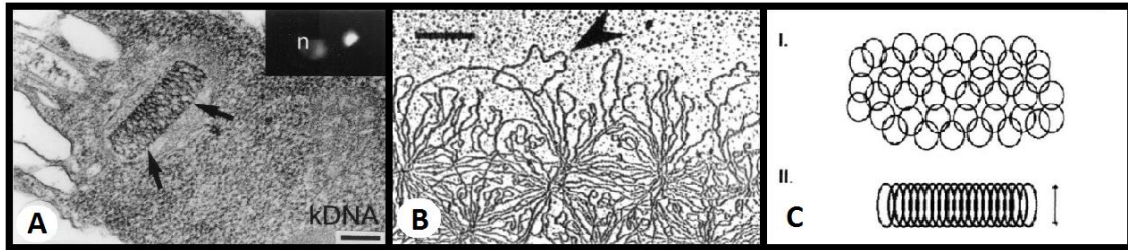


Fonte: Ren e Gupta (2017).

Nota: Dois dos filios mostrados, Apicomplexa e Kinetoplastida, compreendem a grande maioria de patógenos humanos e de animais. Enquanto a maioria dos apicomplexos (exceto gregarinas) favorecem um estilo de vida intracelular, os kinetoplastídeos preferem uma vida extracelular (exceto certos estágios de *T. cruzi* e *Leishmania*).

A ordem Kinetoplastida é constituída por um grande grupo de protozoários de vida livre ou parasitas de invertebrados, vertebrados e plantas. Todos são caracterizados por possuírem uma estrutura intracelular exclusiva, dentro de sua única e grande mitocôndria (LUKES; GUILBRIDE; VOTÝPKA; ZÍKOVÁ *et al.*, 2002). Essa estrutura possui forma de disco e é denominada de kinetoplasto ou também chamado de kinetoplasto, de onde vem a denominação da ordem (Figura 2A). Essa estrutura é formada por um grande número de moléculas circulares de DNA mitocondrial entrelaçadas (~20.000 minicírculos, 20-50 maxicírculos), o kDNA (Figura 2B e 2C). Todos os protozoários que fazem parte desta ordem possuem um flagelo, que pode ser longo e livre, incorporado na superfície celular (formando uma membrana ondulatória), ou ainda ser pequeno e fechado na bolsa flagelar (BASTIN; PULLEN; MOREIRA-LEITE; GULL, 2000).

Figura 2 - Estrutura do kinetoplasto.



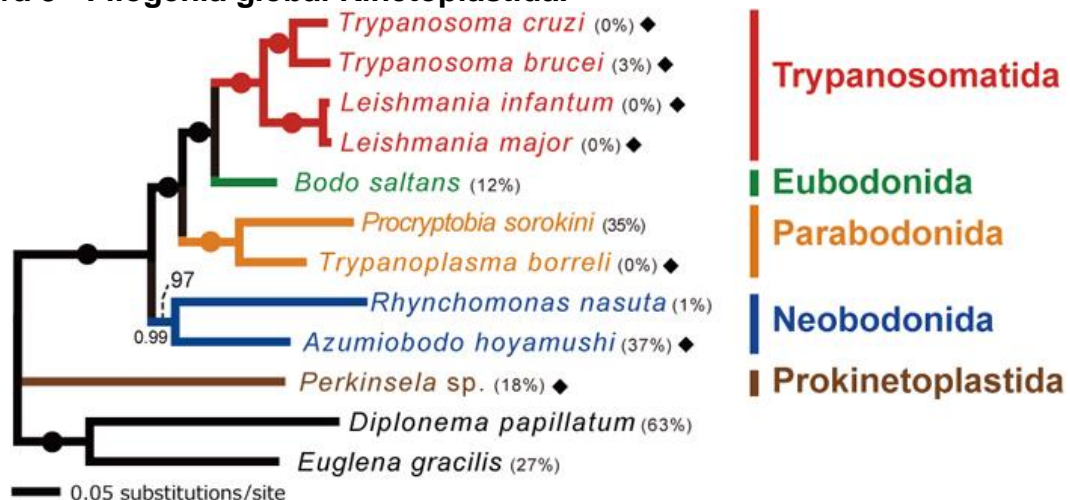
Fonte: adaptado de Lukes *et al.*(2002)

Nota: Em **A)** Eletromicrografia da região do kinetoplasto evidenciando o kDNA de *Trypanosoma avium*. **B)** Eletromicrografia evidenciando a rede de DNA e sua organização no kinetoplasto, na seta um minicírculo claramente visível. **C)** Diagrama esquemático da organização do kDNA, evidenciando as suas duas formas de organização, identificadas através do estudo de cortes transversais no kinetoplasto.

2.1.1.1 Família Trypanosomatidae

Dentro da ordem Kinetoplastida encontra-se a família Trypanosomatidae. Essa família possui atualmente 13 gêneros confirmados, todos compreendendo espécies exclusivamente parasitas: *Leptomonas*, *Chritidia*, *Herpetomonas*, *Blastocrithidia*, *Strigomonas*, *Angomonas*, *Wallaceina*, *Sergeia*, *Blechomonas*, *Phytomonas*, *Endotrypanum*, *Trypanosoma* e *Leishmania* (Figura 3) (LUKEŠ; SKALICKÝ; TÝČ; VOTÝPKA *et al.*, 2014; YAZAKI; ISHIKAWA; KUME; KUMAGAI *et al.*, 2017).

Figura 3 - Filogenia global Kinetoplastida.



Fonte: YAZAKI *et al.*(2017)

Nota: A árvore foi inferida a partir de um alinhamento compreendendo 43 genes que codificam proteínas altamente conservadas. A topologia da árvore foi inferida a partir do método de máxima verossimilhança. Os nós marcados por pontos foram suportados por valores de bootstrap de máxima

verossimilhança de 100% e probabilidades bayesianas posteriores de 1,0. Para cada táxon, a porcentagem de dados ausentes é apresentada entre parênteses. Os parasitas são marcados por diamantes.

A família Trypanosomatidae inclui vários parasitas responsáveis por causarem doenças em humanos e animais. Os mais relevantes por causarem doenças em humanos e pelo seu impacto são *Trypanosoma brucei*, *Trypanosoma cruzi* e *Leishmania*, causadores da Doença do Sono, da Doença de Chagas e as várias manifestações clínicas da Leishmaniose, respectivamente (GUARNER, 2019; KENNEDY, 2019; MCCALL; MCKERROW, 2014). No caso da Leishmaniose, esta é caracterizada por um amplo espectro de manifestações clínicas, entre as quais se destacam a leishmaniose visceral (LV), a leishmaniose Mucocutânea (LMC) e a leishmaniose cutânea (LC), sendo uma doença tropical negligenciada, afetando populações em áreas tropicais e subtropicais em todo o mundo (TASLIMI; ZAHEDIFARD; RAFATI, 2018). A Organização Mundial da Saúde (OMS) estima que mais de 1 bilhão de pessoas vivem em áreas endêmicas de leishmaniose e estão sob risco de infecção. Também é estimado que 30.000 novos casos de LV e mais de 1 milhão de novos casos de LC ocorram anualmente (WHO, 2020). O gênero *Leishmania* é transmitido pela picada de fêmeas de flebotomíneos infectados (PACE, 2014; WHO, 2018) e compreende aproximadamente 53 espécies, dentre as quais 20 são causadoras de leishmaniose em humanos (KAUFER; ELLIS; STARK; BARRATT, 2017).

Atualmente não há vacina contra qualquer forma da leishmaniose aprovada em humanos, e a quimioterapia é a única abordagem para o tratamento. Os tratamentos atuais para leishmaniose incluem quimioterapia com antimoniais para as LC e LMC e Anfotericina B para LV, sendo as opções de tratamento muito limitadas, em parte devido ao alto custo e efeitos adversos significativos, além da citotoxicidade, a resistência aos medicamentos é o principal obstáculo para a terapia antileishmanicida atual, sendo necessário o desenvolvimento de novas formas terapêuticas de rápida ação e com baixa toxicidade (TASLIMI; ZAHEDIFARD; RAFATI, 2018). Um cenário semelhante é observado para as doenças provocadas por outras espécies de tripanosomatídeos e o desenvolvimento de novas ferramentas de controle terapêutico ainda depende de um conhecimento mais fino sobre a biologia destes parasitas voltados a identificação de novos alvos para quimioterápicos. Entretanto as características únicas já observadas em processos

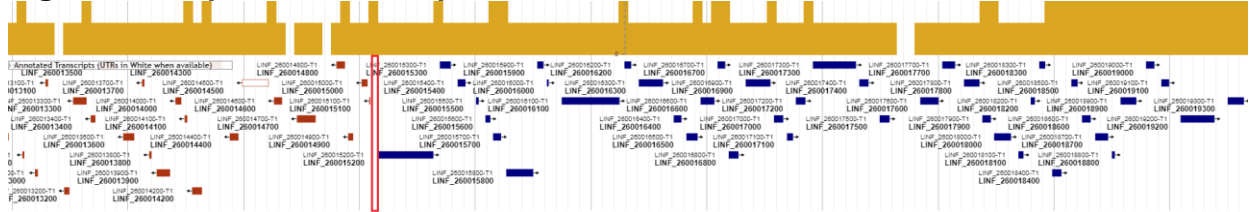
associados a expressão gênica destes parasitas sugerem possíveis processos passíveis de investigação e voltados a identificação destes alvos.

2.2 Expressão gênica em tripanosomatídeos

De uma forma geral, os tripanosomatídeos são caracterizados por ciclos de vida digenéticos, ou seja, se alternam entre dois hospedeiros distintos, sendo um deles o vetor, geralmente um invertebrado, e o segundo o hospedeiro definitivo (planta ou animal). Ao longo do seu ciclo de vida estes parasitas sofrem alterações morfológicas e genéticas que os permitem se adaptar a diferentes microambientes em seus respectivos hospedeiros (LUKEŠ; SKALICKÝ; TÝČ; VOTÝPKA *et al.*, 2014; SIMPSON; STEVENS; LUKES, 2006).

Do ponto de vista de sua biologia molecular, e mesmo tendo divergido entre si há cerca de 200 a 500 milhões de anos, os tripanosomatídeos de uma forma geral apresentam uma notável conservação tanto no número como na organização de seus genes (BRINGAUD; MÜLLER; CERQUEIRA; SMITH *et al.*, 2007; KISSINGER, 2006). Todos têm seu genoma organizado em grandes blocos policistrônicos (Figura 4), ou seja, dezenas a centenas de genes codificadores de proteínas arranjadas na mesma fita de DNA e separados por uma ou mais regiões de troca de fita em cada cromossomo (BUBIS; MARTÍNEZ; CALABOKIS; FERREIRA *et al.*, 2018; SIMPSON; STEVENS; LUKES, 2006). Como consequência do arranjo gênico em blocos, os tripanosomatídeos tem seus mRNAs gerados a partir de precursores policistrônicos gigantes que são transcritos simultaneamente por conta da ausência de promotores para transcrição individual de cada gene. Isso requer um processamento diferenciado destes precursores em mRNAs maduros, conhecido como *trans-splicing*. Este processo é direcionado a todos os RNAs codificantes de proteínas que, ao serem processados, recebem uma sequência de nucleotídeos chamada de sequência *Spliced Leader* (SL) (CAMPBELL; THOMAS; STURM, 2003). Por possuírem sua organização genômica peculiar e também por não possuírem promotores típicos para RNA polimerase II, a regulação da expressão gênica nos tripanosomatídeos é dependente de eventos pós-transcricionais, onde um papel crucial parece ser o do controle da tradução dos mRNAs em proteínas (KRAMER, 2012).

Figura 4 - Arquitetura do mapa do cromossomo 26 de *Leishmania infantum*.



Fonte: Adaptado Tritrypdb. (2022)

Nota: Os genes estão representados por faixas azuis (esquerda para direita) e faixas vermelhas (direita para esquerda) alternando assim a fita positiva e negativa de leitura transcricional dos genes em cluster. Uma região de inversão da transcrição (troca de fita) está presente em retângulo vermelho.

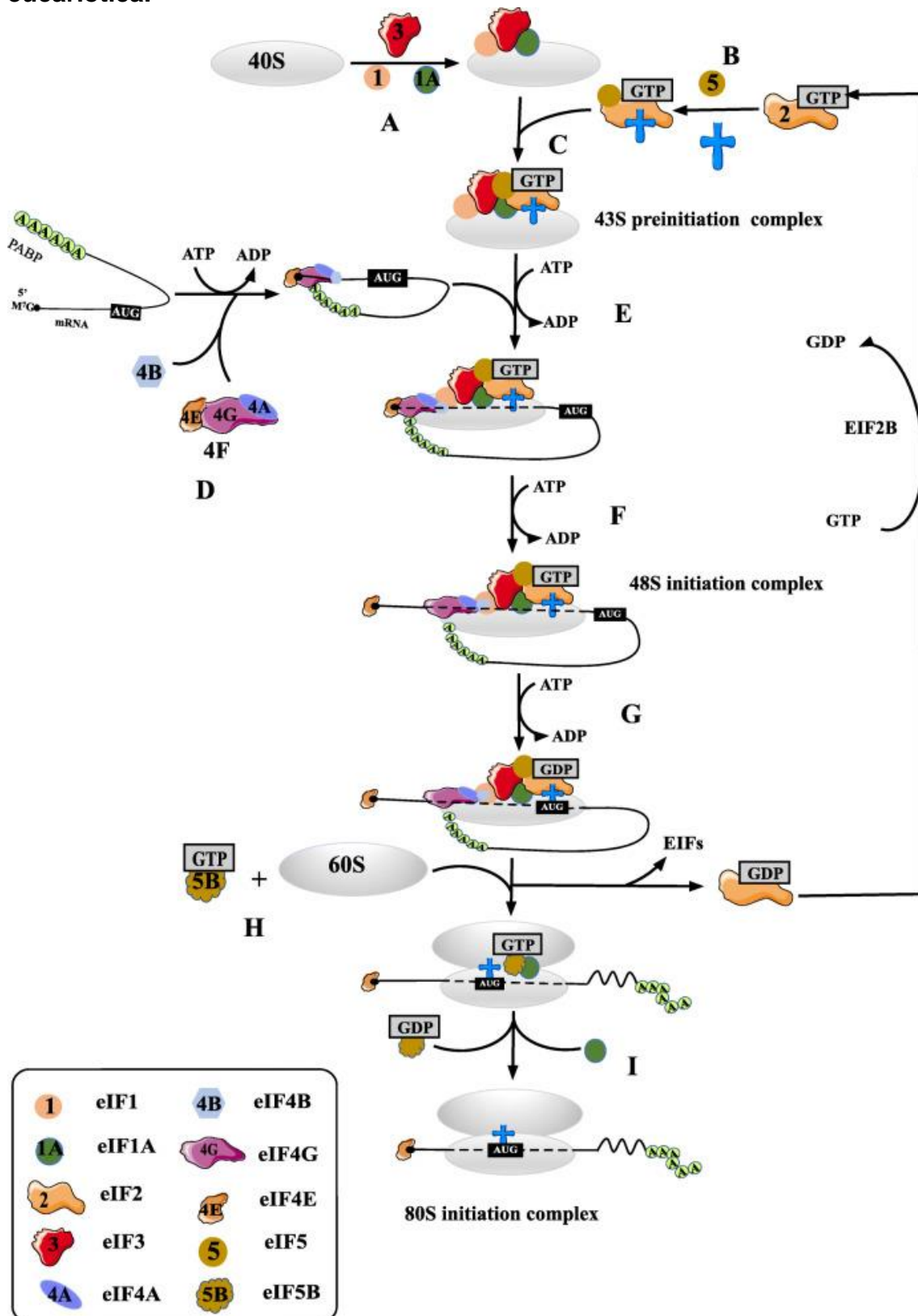
2.3 Iniciação da tradução eucariótica

A síntese proteica, que consiste na tradução dos mRNAs em proteínas, representa o objetivo final da expressão gênica e é essencial para a sobrevivência de todos os organismos vivos. Este é um dos processos celulares mais conservados, sendo controlado por vários fatores proteicos que atuam em quatro etapas principais: iniciação, alongamento, terminação e reciclagem (SONENBERG; HINNEBUSCH, 2009). A regulação da tradução pode ser mediada por mecanismos que atuam em diferentes etapas, mas o alvo principal para essa regulação é a etapa de iniciação, onde um grande número de proteínas denominadas fatores de iniciação eucarióticos (eIFs de *eukaryotic Initiation Factors*), são necessários. Os eIFs participam de muitos pontos críticos requeridos para a tradução prosseguir e estes incluem a identificação de mRNA maduros, proteção contra nucleases, e recrutamento dos ribossomos para o início da tradução corretamente (SONENBERG; HINNEBUSCH, 2009). Embora os mecanismos e a regulação da iniciação da tradução, em diferentes grupos de seres vivos, venham sendo estudados há décadas, muitos aspectos permanecem obscuros (FRASER, 2015).

A iniciação da tradução em eucariotos é a fase mais altamente regulada na tradução da maioria dos mRNAs. Esta etapa compreende todas as etapas que levam à montagem de um ribossomo completo (80S) junto ao mRNA, através da união das subunidades ribossômicas maior (60S) e menor (40S) junto ao mRNA, posicionadas em torno do códon AUG de início da tradução (Figura 5) (SHIROKIKH; PREISS, 2018). O processo começa com a montagem do “complexo de pré-iniciação” 43S (PIC), composto por: o tRNA transportando com anticódon para o primeiro AUG da sequência codificadora e transportando o aminoácido metionina (Met-tRNAi); vários fatores de iniciação, como os eIF1, eIF1A, eIF2 e eIF3; e a

subunidade ribossomal menor 40S. O mRNA maduro é recrutado de forma independente através da ligação do complexo eIF4F a sua extremidade 5'. Em seguida se dá a associação entre o PIC e o mRNA, através da sua extremidade 5' não traduzida, com a subsequente busca pelo códon AUG e recrutamento da subunidade 60S junto com o fator eIF5B. A partir daí se inicia a síntese da cadeia polipeptídica, propriamente.

Figura 5 - Representação esquemática da via de iniciação da tradução eucariótica.



Fonte: HAO *et al.*(2020)

Nota: **A)** Reciclagem de subunidades ribossômicas separadas e eIFs que são gerados a partir de eventos prévios de tradução. **B)** Formação do complexo ternário eIF2-GTP-Met-tRNAⁱ Met. **C)** Formação de 43S PIC que é composto de complexo ternário eIF2-GTP-Met-tRNAⁱ Met, subunidades ribossômicas 40S, eIF1, eIF1A, eIF3 e eIF5. **D)** Ativação do mRNA pelo complexo eIF4F com a ajuda

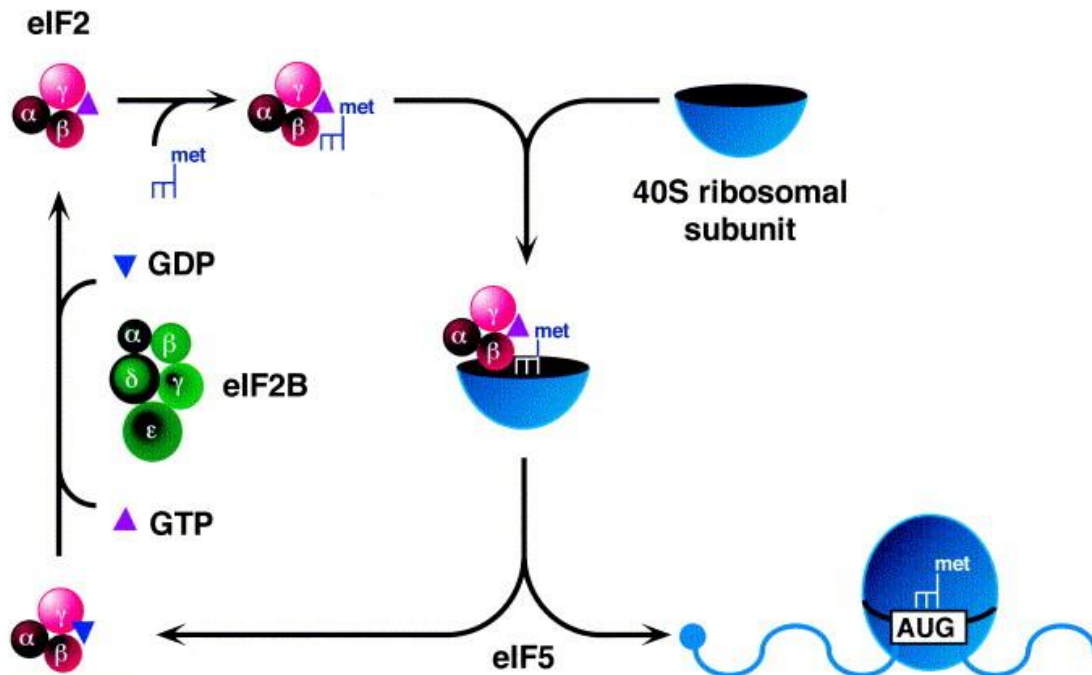
de eIF4B, eIF3 e PABP. **E)** Ligação de 43S PIC ao mRNA. **F)** Varredura de mRNA 5'UTR em uma direção 5'-3' por 43S PIC. **G)** Reconhecimento do códon de iniciação e a formação do complexo de iniciação 48S. **H)** União da subunidade 60S ao complexo 48S com a assistência de eIF5B-GTP e eIF1A, e deslocamento concomitante de eIF2-PIB e outros fatores, incluindo eIF1, eIF3, eIF4B, eIF4F e eIF5. **I)** Hidrólise de GTP ligado a eIF5B e liberação de eIF1A e eIF5B-GDP do ribossomo 80S, com a tradução de mRNA entrando no estágio de alongamento.

2.4 Visão geral do papel de eIFs selecionados na iniciação da tradução

2.4.1 Complexo eIF2

O fator de iniciação eIF2 é uma proteína de ligação ao GTP (proteína G) heterotrimérica formada pelas subunidades eIF2 α , eIF2 β e eIF2 γ . Sua principal função é entregar o Met-tRNA_i à subunidade ribossomal menor, processo que ocorre em todas as células eucarióticas e garante que a síntese proteica comece pelo códon AUG correto (JACKSON; HELLEN; PESTOVA, 2010; WORTHAM; PROUD, 2015). Semelhante a outras proteínas G, o eIF2 alterna entre os estados inativo (ligado ao GDP) e ativo (ligado ao GTP) e este ciclo da proteína G conduz rodadas sucessivas de iniciação da tradução. Juntamente com outros fatores da tradução, o eIF2 ajuda na seleção de mRNA e na varredura para localizar o códon AUG iniciação da tradução. Durante a seleção de AUG pelo ribossomo, o GTP ligado a eIF2 é hidrolisado pela atividade da proteína aceleradora de GTPase (GAP) de eIF5 e uma molécula de fosfato inorgânica (Pi) é liberada. A liberação de Pi reduz a afinidade de eIF2 para Met-tRNA_i e desencadeia a liberação de ambos os fatores do ribossomo, permitindo a formação do ribossomo 80S completo e o alongamento da tradução. O eIF2 é, portanto, liberado dos ribossomos em seu estado inativo em complexo com o eIF5. Para a próxima rodada de iniciação da tradução, o GDP do eIF2 γ deve ser substituído por GTP, essa troca é realizada pelo fator eIF2B, um complexo heterodecamérico de dois conjuntos das subunidades α , β , γ , δ e ϵ . Este é um fator de troca de nucleotídeo guanina (GEF) específico para eIF2 (Figura 6) (ALGIRE; MAAG; LORSCH, 2005; KASHIWAGI; ITO; YOKOYAMA, 2017; KOMAR; MERRICK, 2020).

Figura 6 - Esquema que descreve o papel de eIF2 na iniciação da tradução.

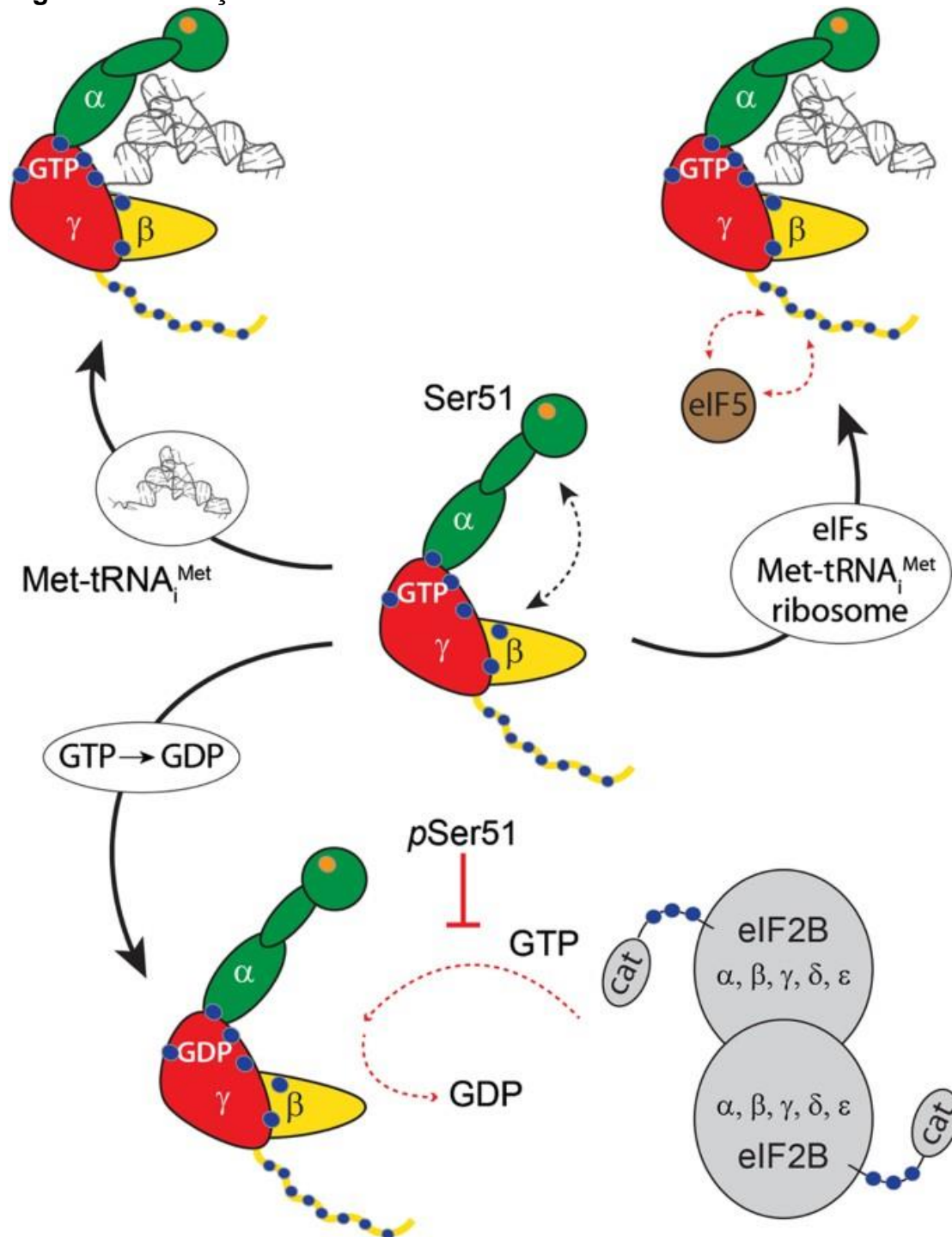


Fonte: KIMBALL.(1999)

Nota: O papel principal do eIF2 na iniciação da tradução, é transferir Met-tRNA_i para a subunidade ribossomal 40S. Seguindo a associação de mRNA com a subunidade 40S e localização da subunidade no códon de início AUG, o eIF5 se liga a eIF2 e estimula a hidrólise de GTP ligado a eIF2. O eIF5 estimula a hidrólise de GTP apenas quando o eIF2 está ligado à subunidade 40S com complexo ternário com GTP e Met-tRNA_i. Após a hidrólise do GTP, o complexo eIF2 -GDP é liberado do ribossomo. No próximo ciclo o GDP do eIF2 é substituído por GTP pelo eIF2B.

Diversos estudos bioquímicos, genéticos e estruturais do eIF2 demonstraram que as três subunidades que compõem o complexo têm funções distintas. A subunidade eIF2 α contém o principal resíduo-alvo para a regulação de eIF2 (por fosforilação), uma serina na posição 51, sendo a sua fosforilação uma das vias regulatórias de inibição da síntese de proteínas. A subunidade eIF2 β contém motivos de ligação para eIF2B (o fator de troca de nucleotídeo da guanina) e eIF5 (a proteína ativadora de GTPase). A subunidade eIF2 γ é a maior subunidade do complexo (52 kDa) e é responsável pela ligação aos nucleotídeos de guanina, GTP e GDP e ao Met-tRNA_i (Figura 7) (BORCK; SHIN; STILLER; MIMOUNI-BLOCH *et al.*, 2012; DONNELLY; GORMAN; GUPTA; SAMALI, 2013; KOMAR; MERRICK, 2020).

Figura 7 - Interações do eIF2 com o tRNA Iniciador e o eIF2B.



Fonte: BEILSTEN-EDMANDS *et al.* (2015)

Nota: Representação do eIF2 (α, verde; β, amarelo; e γ, vermelho), com a serina 51 do eIF2α destacada (laranja). A localização de motivos de fosforilação e acetilação são indicadas (azul). O eIF2 ligado a GTP é altamente flexível em solução. Após a ligação de Met-tRNA_i, o complexo ternário adota uma conformação mais rígida. O N-terminal não estruturado do eIF2β, quando o complexo está ligado a GTP, interage com eIF5. A liberação de Met-tRNA_i do eIF2 e a transferência para o ribossomo é desencadeada pela hidrólise do GTP. O Pi é trocado pelo fator de troca de nucleotídeos eIF2B, preparando o eIF2 para uma nova rodada de iniciação da tradução. A troca de nucleotídeos é inibida pela fosforilação da serina 51.

2.4.2 Complexo eIF3

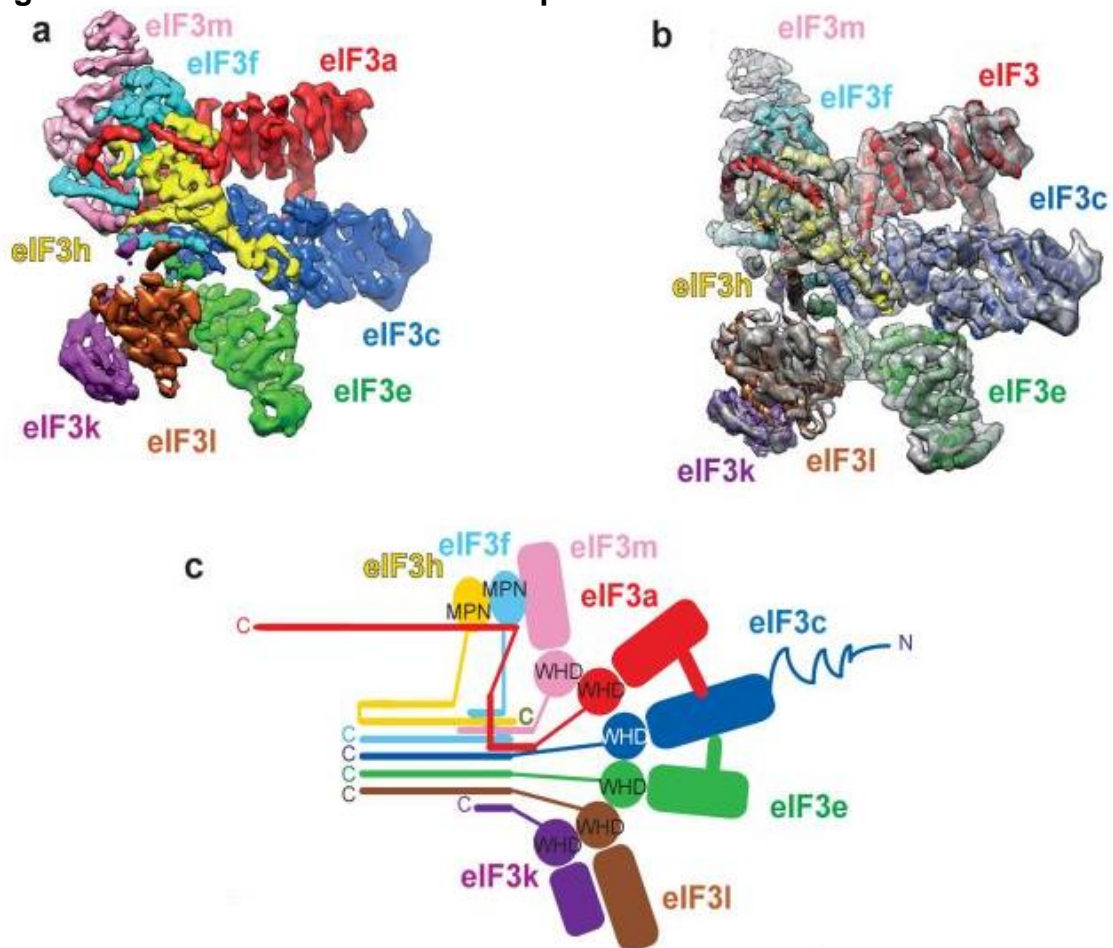
Entre as diversas proteínas envolvidas na iniciação da tradução vale ressaltar o fator eIF3, o maior dos eIFs. Este em mamíferos é composto por 13 subunidades (a – m) não idênticas, com aproximadamente 800 kDa (CATE, 2017; DES GEORGES; DHOTE; KUHN; HELLEN *et al.*, 2015). A composição do eIF3 é conservada na maioria dos eucariotos como em plantas e fungos filamentosos (BROWNING; BAILEY-SERRES, 2015; CATE, 2017), mas difere em *S. cerevisiae*, onde o complexo compreende apenas seis subunidades (a, b, c, g, i e j) (GENUTH; BARNA, 2018; SENOHRABKOVA; MALCOVA; HASEK, 2019).

O eIF3 está envolvido em quase todas as etapas da tradução. Na iniciação, ele atua interagindo com o complexo eIF4F associado ao mRNA e atraindo o PIC 43S e auxiliando na varredura e reconhecimento do códon AUG (DES GEORGES; DHOTE; KUHN; HELLEN *et al.*, 2015; JACKSON; HELLEN; PESTOVA, 2010; WOLF; LIN; DUAN; CHENG, 2020). Por atuar ainda no recrutamento do complexo ternário (TC) para formar o PIC 43S, o eIF3 também atua como um ativador ou repressor da tradução (AITKEN; BEZNOŠKOVÁ; VLČKOVA; CHIU *et al.*, 2016; GOMES-DUARTE; LACERDA; MENEZES; ROMÃO, 2018). Além disso, tem envolvimento em mecanismos não canônicos de iniciação da tradução do mRNA refletindo a sua complexidade (GENUTH; BARNA, 2018; THAKOR; SMITH; ROBERTS; FAYE *et al.*, 2017)

Como dito anteriormente, o complexo eIF3 de mamíferos é formado por 13 subunidades, sendo que oito delas se organizam no chamado octâmero e são relacionadas a subunidades encontradas em dois outros grandes complexos celulares, o Proteassoma e um complexo distinto chamado COP9. Seis das subunidades do eIF3 contém um domínio denominado de PCI comum aos três complexos (eIF3a, eIF3c, eIF3e, eIF3k, eIF3l e eIF3m) e duas subunidades, eIF3f e eIF3h, possuem o MPN, um segundo domínio comum aos três complexos e envolvido em mediar interações entre as subunidades (Figura 8) (DES GEORGES; DHOTE; KUHN; HELLEN *et al.*, 2015; SIMONETTI; BRITO QUERIDO; MYASNIKOV; MANCERA-MARTINEZ *et al.*, 2016). As subunidades eIF3b, eIF3g e eIF3i se ligam de maneira flexível ao octâmero de subunidades PCI/MPN por meio do eIF3a, organizadas como um submódulo. Já o eIF3d está localizado em uma

posição mais periférica anexada à subunidade eIF3e do octâmero, próximo a saída do mRNA e interage diretamente com eIF4G e com o *cap* e está implicado na tradução de mRNAs. A subunidade eIF3j, por sua vez, é fracamente anexada ao resto do eIF3. As subunidades eIF3b e eIF3a são ainda consideradas como um núcleo em torno do qual outras subunidades se agrupam de forma ordenada em dois módulos interconectados (DES GEORGES; DHOTE; KUHN; HELLEN *et al.*, 2015; ELLISDON; STEWART, 2012; GOMES-DUARTE; LACERDA; MENEZES; ROMÃO, 2018; KRAMER, 2012; LEE; KRANZUSCH; DOUDNA; CATE, 2016; VILLA; DO; HERSHEY; FRASER, 2013).

Figura 8 - O núcleo PCI/MPN do complexo eIF3 de mamíferos.



Fonte: Adaptado de DES GEORGES *et al.*(2015)

Nota: A) Núcleo segmentado do eIF3, colorido de forma variável por subunidade, visto de diferentes orientações. B) Montagem do modelo de núcleo eIF3, com base em crio-EM C) Representação bidimensional da estrutura tridimensional do núcleo do octâmero eIF3. O feixe helicoidal é representado por barras coloridas. A linha em ziguezague em eIF3c indica cauda N-terminal possivelmente não estruturada. A região C-terminal do eIF3a (não presente na estrutura relatada aqui), não é mostrada nesta representação esquemática.

2.4.3 Complexo eIF4F

O complexo heterotrimérico eIF4F é formado pela associação das subunidades eIF4E, eIF4G e eIF4A, onde eIF4E é responsável pelo reconhecimento e ligação específica ao *cap* de m⁷GTP encontrado na extremidade 5' do mRNA. O eIF4G é o núcleo central do complexo e se liga tanto ao eIF4E quanto ao eIF4A, organizando o complexo eIF4F. Já o eIF4A é uma helicase de RNA que possui o papel de desenrolar quaisquer estruturas secundárias encontradas na região 5' não traduzida (5' UTR) do mRNA. Para o reconhecimento propriamente do mRNA pela maquinaria de tradução, o PIC 43S é conduzido para a região 5' do mRNA por meio de uma interação entre múltiplas subunidades do complexo eIF3 e a subunidade eIF4G do eIF4F. Simultaneamente, o eIF4F, através da subunidade eIF4G, se liga a proteína de ligação a cauda poli-A (PABP), por sua vez associada a extremidade 3' dos mRNAs, pela sua cauda poli-A (HAO; YU; WARD; LIU *et al.*, 2020; PELLETIER; GRAFF; RUGGERO; SONENBERG, 2015; SHIROKIKH; PREISS, 2018; SONENBERG; HINNEBUSCH, 2009).

A ligação da PABP a cauda poli-A do mRNA, na sua extremidade 3', e a sua interação com o eIF4G, preso a extremidade 5', promove a circularização do mRNA durante a tradução (JACKSON; HELLEN; PESTOVA, 2010)(JACKSON *et al.*, 2010). Acredita-se que esta formação em circuito fechado aumenta a eficiência do processo. A atividade helicase do complexo eIF4F facilita ainda a mobilidade de todo o PIC ao longo do mRNA, um requisito para a varredura da região 5'UTR e a identificação do primeiro AUG. Após a identificação do AUG, as subunidades do complexo eIF4F são dispersas e em seguida ocorre o recrutamento da subunidade ribossomal maior e subsequente formação da primeira ligação peptídica da proteína nascente (BROWNING; BAILEY-SERRES, 2015; SONENBERG; HINNEBUSCH, 2009).

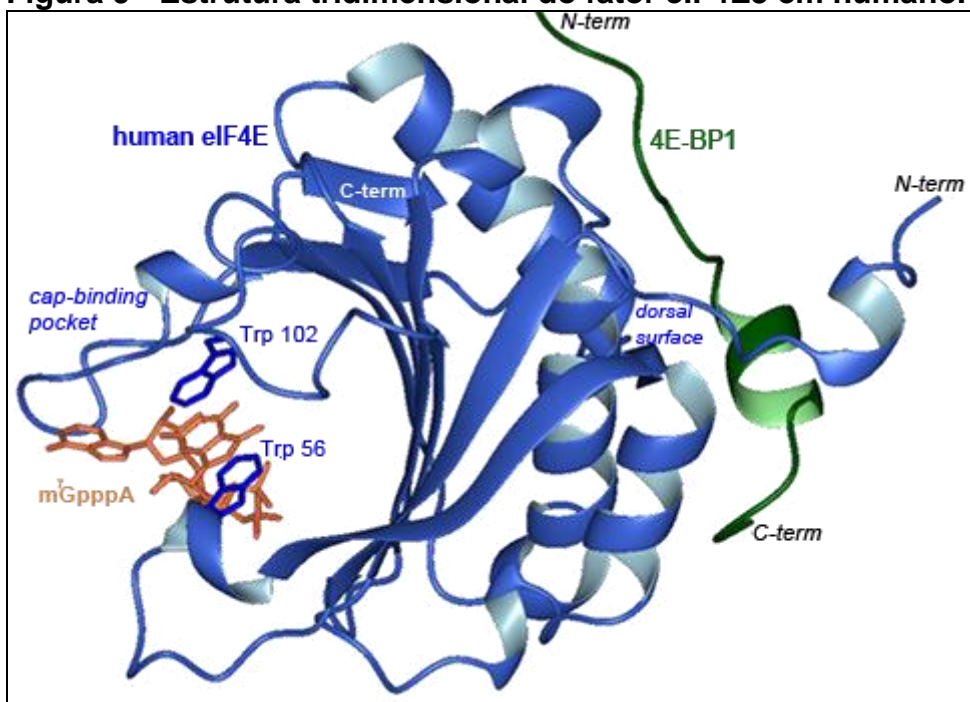
2.4.3.1 Superfamília eIF4E

O fator eIF4E, a proteína de ligação ao *cap*, é reconhecido como representante de um grande grupo de proteínas eucarióticas classificadas na superfamília eIF4E, geralmente definida pela presença de até oito resíduos

aromáticos, em geral triptofanos, conservados. Estes foram originalmente descritos com base nos homólogos do eIF4E de mamífero e de leveduras, com a numeração de sequência dos aminoácidos em mamíferos utilizada como referência aqui: W43, W46, W56, W73, W102, W112, W130 e W166. Quatro desses resíduos são essenciais para as funções clássicas do eIF4E, com os W56, W102 e W166 sendo necessários para o reconhecimento do *cap*, enquanto que o W73 está envolvido na ligação ao eIF4G (JAGUS; BACHVAROFF; JOSHI; PLACE, 2012). Ao se comparar os eIF4Es de vários organismos, foi possível determinar que esses resíduos de triptofanos são espaçados dentro de uma região central conservada. Este região constitui o núcleo típico (*core*) dos eIF4Es e é encontrado em todos os membros dessa família de proteínas (JOSHI; LEE; MAEDER; JAGUS, 2005).

Do ponto de vista estrutural, os eIF4Es se doblam em uma estrutura terciária em formato de “luva de beisebol”, tendo oito folhas antiparalelas e três alfa hélices (Figura 9), sendo essa estrutura crítica para sua habilidade de ligar ao *cap* de m7GTP (GINGRAS; RAUGHT; SONENBERG, 1999). A ligação ao *cap* é mediada pelos dois triptofanos conservados localizados no lado côncavo do eIF4E, W56 e W102 (MATSUO; LI; MCGUIRE; FLETCHER *et al.*, 1997). Uma rede de pontes de hidrogênio também fortalece a interação eIF4E-*cap* envolvendo o W56 e outros resíduos (LAMA; PRADHAN; BROWN; EAPEN *et al.*, 2017).

Figura 9 - Estrutura tridimensional do fator eIF4E3 em humano.



Fonte: BORDEN; CULJKOVIC-KRALJACIC e VOLPON.(2015)

Nota: Complexo ternário do peptídeo m7GpppA-eIF4E-4EBP1. eIF4E (azul) e o peptídeo 4E-BP1 (verde) são mostrados com o modelo de fita. O modo de reconhecimento do cap do eIF4E é caracterizado por um sanduíche pi-pi da base m7G do cap (aqui m7GpppA, laranja) pelos anéis aromáticos de Trp56 e Trp102.

Através do seu lado convexo, o eIF4E pode se ligar a várias proteínas com papéis diferentes, mas seu principal parceiro na tradução é o eIF4G (KAMENSKA; SIMPSON; STANDART, 2014; RHOADS, 2009). A ligação eIF4E-eIF4G é mediada pelos resíduos VEDFW do eIF4E e o motivo canônico Y(X)₄LΦ (onde X representa qualquer aminoácido e Φ representa um resíduo hidrofóbico) localizado dentro do terço N-terminal do eIF4G (MARCOTRIGIANO; GINGRAS; SONENBERG; BURLEY, 1999; MOERKE; AKTAS; CHEN; CANTEL *et al.*, 2007). Este motivo de ligação ao eIF4E também é encontrado em várias outras proteínas, como é o caso de proteínas de ligação eIF4E (4E-BPs) de vários organismos que se ligam competitivamente ao eIF4E, impedindo a sua ligação ao eIF4G, seja para promover ou reprimir a tradução de mRNAs específicos (KAMENSKA; SIMPSON; STANDART, 2014).

O eIF4E participa também em outros processos celulares além da iniciação da tradução, como o transporte do mRNA do núcleo para o citoplasma, sequestro de mRNAs instáveis bem como a estabilização do mRNA, protegendo-o da degradação no citosol (STRUDWICK; BORDEN, 2002). Esta é uma proteína encontrada apenas em eucariotos, mas que pode ser representado por vários homólogos em uma mesma espécie (Tabela 1).

Tabela 1 - Homólogos do fator eIF4E em diferentes espécies.

Organismo modelo	Nome científico	Homólogos do eIF4E	Referências
Camundongo	<i>Mus musculus</i>	eIF4E-1, 4EHP e eIF4E-3	(JOSHI; CAMERON; JAGUS, 2004)
Humano	<i>Homo sapiens</i>	eIF4E1, eIF4E2 e 4EHP	(GAO; RYCHLIK; RHOADS, 1998; ROM; KIM; GINGRAS; MARCOTRIGIANO <i>et al.</i> , 1998)
Vegetal	<i>Arabidopsis thaliana</i>	eIF4E, eIFiso4E e nCBP	(RUUD; KUHLLOW; GOSS; BROWNING, 1998).
Levedura	<i>Schizosaccharomyces pombe</i>	eIF4E-1 e eIF4E-2	(PTUSHKINA; BERTHELOT; VON DER HAAR; GEFFERS <i>et al.</i> , 2001)
Nematóide	<i>Caenorhabditis elegans</i>	IFE-1, IFE-2, IFE-3, IFE-4 e IFE-5	(KEIPER; LAMPHEAR; DESHPANDE; JANKOWSKA-ANYSZKA <i>et al.</i> , 2000)

O eIF4E possui um papel chave na regulação da síntese proteica e, em organismos multicelulares, foi observada a relação direta deste fator com o câncer, sendo frequentemente detectado em altos níveis em tumores e correlacionado com prognósticos negativos. Estes dados reforçam a importância deste fator como um alvo a ser estudado na produção de novos fármacos (PETTERSSON; DEL RINCON; MILLER, 2014).

2.5 eIFs de tripanosomatídeos

2.5.1 Complexo eIF2

O complexo eIF2 dos tripanosomatídeos também é formado por três subunidades com funções conservadas. A subunidade eIF2 α de *T. brucei* e *T. cruzi* é fosforilada no resíduo treonina T169 e em *Leishmania* no resíduo Thr 166, o que corresponde a serina S51 em outros eucariotos. Em *Leishmania*, essa fosforilação de eIF2 α em resposta ao estresse se demonstrou ser relevante para a diferenciação celular da forma promastigota, encontrada no vetor inseto, nas amastigotas encontrados nos macrófagos de mamíferos. Da mesma forma, tanto em *T. brucei* quanto em *T. cruzi*, e em outros protozoários como *Toxoplasma gondii*, e *Plasmodium falciparum*, uma resposta equivalente associada a diferenciação celular também já havia sido observada (CASTRO MACHADO; BITTENCOURT-CUNHA; MALVEZZI; ARICO *et al.*, 2020; CHOW; CLOUTIER; DUMAS; CHOU *et al.*, 2011; NARASIMHAN; JOYCE; NAGULESWARAN; SMITH *et al.*, 2008; TONELLI; AUGUSTO; CASTILHO; SCHENKMAN, 2011; ZHANG; FENNELL; RANFORD-CARTWRIGHT; SAKTHIVEL *et al.*, 2010).

Em estudo recente sobre a estrutura de eIF2 β de *T. cruzi*, foram caracterizados resíduos envolvidos no ancoramento do eIF2 β e do eIF1 dentro do PIC 43S. Essa ligação de eIF2 β em eIF1, ocorre pelos resíduos T325, Y326 e S327 na porção C terminal, e essa porção C terminal também interage com o rRNA 18S. Também se avaliou a interação entre o eIF5 e o eIF2 γ (domínio G), onde R229, R230 e R273 do eIF5 entram em contato com D219, S224 e S220 do domínio G de

eIF2 γ . Com base nessa estrutura, o arranjo do sítio de ligação do eIF5 em eIF2 γ no contexto do 43S PIC é semelhante ao do domínio eIF2B ϵ no contexto do complexo eIF2-eIF2B, essa interação eIF5-CTD indica que a interação entre eIF5-CTD e eIF2 γ poderia, induzir uma mudança conformacional sutil em seu domínio G, permitindo que o eIF5-NTD (um domínio de ativação de GTPase de eIF5) ganhasse acesso ao bolso de ligação de GTP para promover a hidrólise reversível de GTP em eIF2 durante a varredura do mRNA (BOCHLER; QUERIDO; PRILEPSKAJA; SOUFARI *et al.*, 2020).

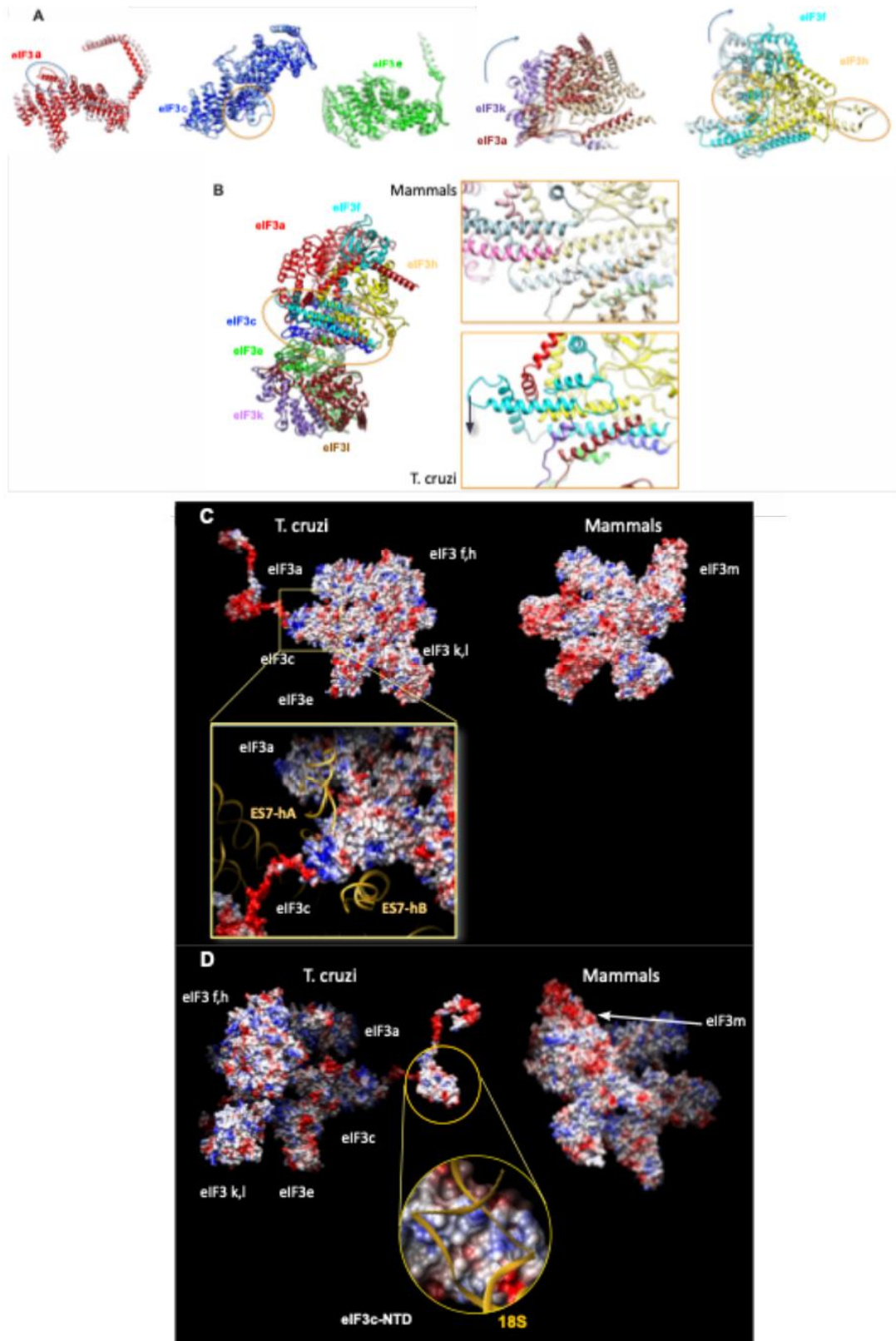
2.5.2 Complexo eIF3

Em tripanosomatídeos poucos estudos têm como foco o papel do eIF3 na tradução, além do que a presença de diversas subunidades do complexo eIF4F e os mecanismos pelo qual esses organismos recrutam o ribossomo 40S ainda não são completamente elucidados. A composição do complexo eIF3 de tripanosomatídeos foi determinada por meio de análises comparativas de sequências *in silico* e confirmadas por caracterização bioquímica e espectrometria de massas. Demonstrou-se que esses organismos possuem apenas 12 subunidades, não codificando a subunidade eIF3m. As subunidades identificadas apresentam um alto grau de conservação como os domínios PCI/MPN, porém algumas subunidades como o eIF3f e eIF3a apresentam grandes variações em sua estrutura o que implica em características distintas do complexo em tripanosomatídeos (LI; ZHOU; GUO; CHEN *et al.*, 2017; MELEPPATTU; KAMUS-ELIMELEH; ZINOVIEV; COHEN-MOR *et al.*, 2015; REZENDE; ASSIS; NUNES; DA COSTA LIMA *et al.*, 2014).

De fato, quando comparado ao de mamíferos, a conformação do núcleo estrutural eIF3 de *T. cruzi* difere devido à ausência da subunidade eIF3m, ocasionando em rearranjos de outras subunidades do eIF3, principalmente as subunidades eIF3f e eIF3h, que alteram sua conformação para suprir essa ausência. Este rearranjo é provavelmente necessário para a manutenção do núcleo central do eIF3 (Figura 10A e 10B, as setas indicam a direção da mudança). *T. cruzi* eIF3. Além disso, uma análise da superfície de carga revela padrões de distribuição de carga muito diferentes entre *T. cruzi* eIF3 e sua contraparte de mamífero (Figura 10C e 10D), em parte como consequência da superfície de ligação de 40S diferente que

é representada principalmente por rRNA, que está em contraste com outros eucariotos conhecidos (BOCHLER; QUERIDO; PRILEPSKAJA; SOUFARI et al., 2020).

Figura 10 - Análise da estrutura do eIF3 de mamíferos e *T. cruzi*.



Fonte: BOCHLER *et al.*(2020)

Nota: A) Sobreposição de estruturas de mamíferos e kinetoplastidas de subunidades individuais do eIF3 com diferenças estruturais marcadas. As estruturas do *T. cruzi* são representadas em tons escuros e os de mamíferos em tons claros. Setas curvas indicam a direção do rearranjo estrutural das subunidades de eIF3 de *T. cruzi* em comparação com suas contrapartes de mamíferos. Círculos coloridos destacam diferenças estruturais marcadas entre subunidades do eIF3 de *T. cruzi* e de mamíferos. B) Desenho representativo do modelo atômico do eIF3 mostrando o feixe helicoidal em mamíferos (painel superior) e em *T. cruzi* (painel inferior). A seta escura indica o deslocamento de uma hélice de eIF3f em *T. cruzi* para compensar a ausência de eIF3m. C) Representação de superfície da estrutura eIF3 de *T. cruzi* (esquerda) e mamífero (direita) em relação ao 40S. Painel inferior: visão aproximada do eIF3c de *T. cruzi* e sua interação com a hélice A e a hélice B. O modelo é codificado por cores de acordo com o potencial eletrostático - negativo em vermelho e positivo em azul. D) Representação de superfície da estrutura do eIF3 de *T. cruzi* (esquerda) e de mamífero (direita) vista do lado do solvente 40S. Painel inferior: visão aproximada do eIF3c-NTD de *T. cruzi* e sua interação com o RNA 18S.

2.5.3 Complexos do tipo eIF4F em tripanosomatídeos

A partir do que já foi descrito nos eucariotos, foram feitos avanços para se melhor compreender o complexo eIF4F e suas subunidades em tripanosomatídeos. Com a publicação dos genomas das três primeiras espécies de tripanosomatídeos, os TriTryps, um total de seis homólogos eIF4E foram identificados, sendo quatro deles descritos inicialmente, EIF4E1 a EIF4E4 (DHALLIA; REIS; FREIRE; ROCHA *et al.*, 2005), e outros dois bem depois, EIF4E5 e EIF4E6 (FREIRE; MALVEZZI; VASHISHT; ZUBEREK *et al.*, 2014; FREIRE; VASHISHT; MALVEZZI; ZUBEREK *et al.*, 2014). No que concerne aos homólogos de eIF4G foram descritos cinco homólogos no total (EIF4G1 a EIF4G5) (DHALLIA; REIS; FREIRE; ROCHA *et al.*, 2005), além de dois homólogos de eIF4A (EIF4AI e EIF4AIII) (DHALLIA; MARINSEK; REIS; KATZ *et al.*, 2006). Até o momento foi confirmada a formação de dois complexos principais do tipo eIF4F, um entre os fatores EIF4E4/EIF4G3/EIF4AI e o outro baseado nos fatores EIF4E3/EIF4G4 com a provável participação do EIF4AI (DHALLIA; REIS; FREIRE; ROCHA *et al.*, 2005). Foi observado ainda a formação de mais três complexos do tipo eIF4F, mas que ainda não se sabe seu papel na síntese proteica. Estes são baseados nos EIF4E5/EIF4G1, EIF4E5/EIF4G2 (FREIRE; VASHISHT; MALVEZZI; ZUBEREK *et al.*, 2014) e no EIF4E6/EIF4G5 (FREIRE; MALVEZZI; VASHISHT; ZUBEREK *et al.*, 2014).

Em trabalhos realizados em *T. brucei* foi visto uma diminuição brusca e acelerada do crescimento celular quando o fator EIF4G3 foi depletado através de ensaio de RNAi, com um resultado diferente observado para o EIF4G4 com o mesmo ensaio, onde ocorre uma diminuição mais lenta do seu crescimento e com

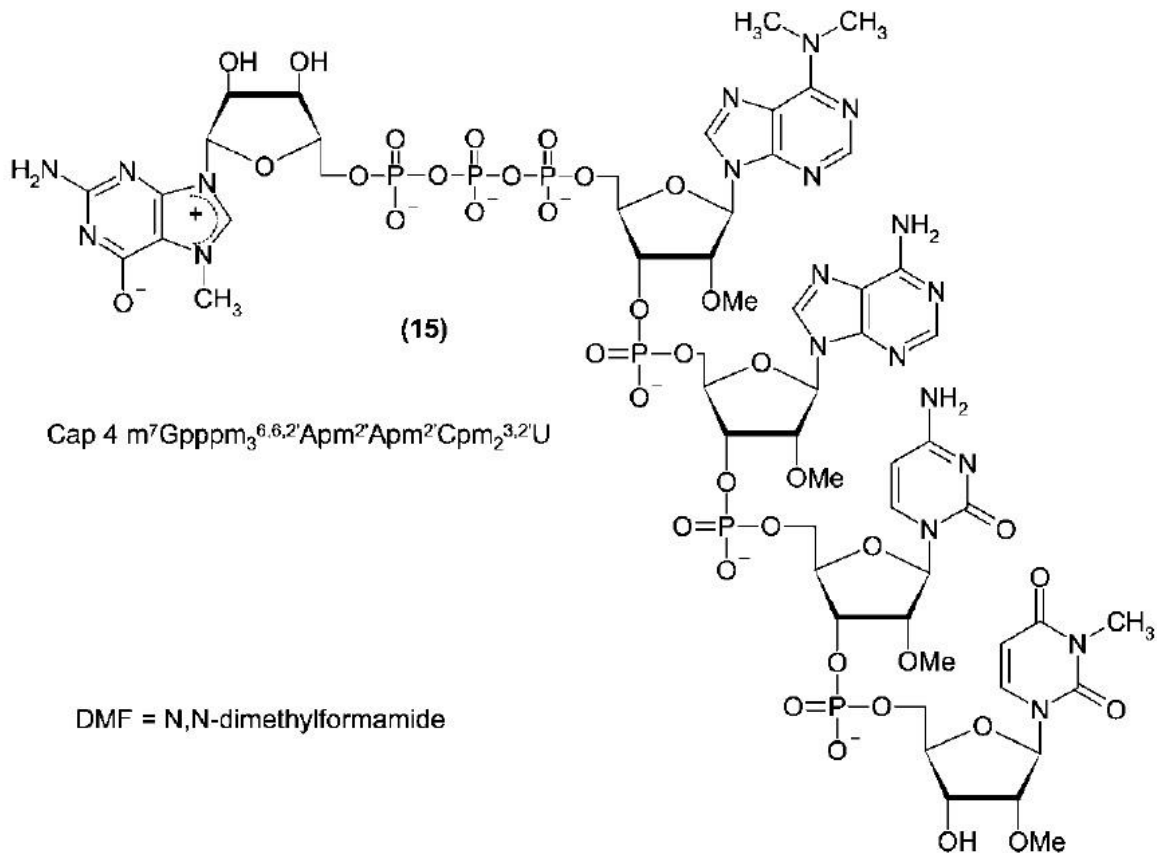
alterações morfológicas. Como estes fatores possuem sequências e estruturas similares, acredita-se que o complexo baseado no EIF4E4/EIF4G3 tem a função principal na síntese proteica, enquanto o complexo EIF4E3/EIF4G4 também atue na tradução, mas de forma mais restrita (MOURA; REIS; XAVIER; DA COSTA LIMA *et al.*, 2015). Além dos fatores que compõem o eIF4F, no genoma de *L. major* foram encontrados três sequências homólogas da PABP (PABP1, PABP2 e PABP3), porém ortólogos de apenas as duas primeiras foram observadas em *T. cruzi* e *T. brucei* (DA COSTA LIMA; MOURA; REIS; VASCONCELOS *et al.*, 2010). Estudos destas proteínas em *Leishmania* indicam que apenas a PABP1 participa do complexo formado pelos EIF4G3/EIF4E4/EIF4A1. As PABPs 2 e 3 se ligam aos mesmos mRNAs, porém a mRNAs distintos daqueles nos quais a PABP1 se liga, reforçando funções diferenciadas para estas proteínas (DA COSTA LIMA; MOURA; REIS; VASCONCELOS *et al.*, 2010).

2.5.3.1 Cap4

Enquanto ocorre a maturação do RNA mensageiro é adicionado o *cap*, uma estrutura composta por uma guanossina metilada, também chamada de m⁷GTP, na região 5' de todos os mRNAs, sendo importante para estabilizar o mRNA e atuar na direção da síntese proteica. (COWLING, 2009).

Em tripanosomatídeos é encontrado uma *cap* modificado, o *cap4*, este recebe essa nomenclatura pois a numeração é dada pelo número de modificações por metilação na guanossina invertida e nos nucleotídeos próximos a este (Figura 11). Este *cap* é característico da família Trypanosomatidae, e de forma geral foi observado uma grande importância do *cap4* para esta família atuando principalmente na síntese proteica.

Figura 11 - Estrutura química do *cap4*



Fonte: MAGDALENA *et al.* (2004)

Nota: Estrutura química representativa do cap modificado cap4.

2.5.3.2 Homólogos de eIF4E

A análise dos homólogos de eIF4E de *Leishmania* mostrou que estes diferem na sua afinidade pelo nucleotídeo *cap*, na sua associação a polissomos e, como já ressaltado, na sua capacidade de se ligar a homólogos de eIF4G (DHALIA; REIS; FREIRE; ROCHA *et al.*, 2005; YOFFE; LÉGER; ZINOVIEV; ZUBEREK *et al.*, 2009; YOFFE; ZUBEREK; LERER; LEWDOROWICZ *et al.*, 2006). Estudos de homologia por modelagem molecular sugeriram que a estrutura básica de ligação ao *cap* observada em leveduras e mamíferos está conservada nos diferentes homólogos de eIF4E (ZINOVIEV; LÉGER; WAGNER; SHAPIRA, 2011).

A caracterização dos quatro primeiros homólogos de eIF4E em *T. brucei* os classificou em dois grupos, com base em propriedades estruturais e moleculares (FREIRE; DHALIA; MOURA; DA COSTA LIMA *et al.*, 2011). No entanto, com a identificação de mais dois homólogos eIF4E, um terceiro grupo foi proposto

(FREIRE; VASHISHT; MALVEZZI; ZUBEREK *et al.*, 2014). O grupo 1 então é formado por EIF4E1 e EIF4E2, proteínas pouco abundantes, de alta similaridade de tamanho e maior homologia ao eIF4E de humano. Estes homólogos não formam complexos do tipo EIF4F e parecem não ser necessários para tradução geral. O grupo 2 dos eIF4Es é formado pelos EIF4E3 e EIF4E4, proteínas mais abundantes que possuem longas extensões N-terminais. Estas proteínas possuem uma menor homologia entre os eIF4Es, a exceção da região do *core*, mas, no entanto, formam complexos eIF4F típicos implicados na tradução geral e essenciais para a sobrevivência celular. O grupo 3 é composto por EIF4E5 e EIF4E6, proteínas pequenas e com baixa homologia do *core* do eIF4E, mas que são capazes de se ligar a homólogos eIF4G como parte de novos complexos eIF4F, que provavelmente não estão envolvidos na tradução geral (NETO; REIS; MOURA; CARRINGTON, 2016). Análises comparando apenas sequências dos eIF4Es de vários grupos protistas, confirmam que o EIF4E1-EIF4E2 e EIF4E3-EIF4E4 pertencem a clados de eIF4E filogeneticamente independentes, confirmando assim sua separação, baseada em suas propriedades biológicas, em dois grupos independentes (FREIRE; DHALIA; MOURA; DA COSTA LIMA *et al.*, 2011).

2.5.3.2.1 Grupo 1: EIF4E1 e EIF4E2

O primeiro homólogo de eIF4E relatado de tripanosomatídeos foi o EIF4E1, caracterizado de *L. major* e descrito como uma proteína com 42% de similaridade com o eIF4E de mamíferos. Sua estrutura 3D predita indicou que ele deveria se dobrar em uma estrutura geral compatível com o descrito para o eIF4E de mamíferos, apesar da baixa conservação em sequência entre as duas proteínas (DHALIA; REIS; FREIRE; ROCHA *et al.*, 2005; YOFFE; ZUBEREK; LEWDOROWICZ; ZEIRA *et al.*, 2004). O EIF4E2 de *L. major* apresenta 41% de semelhança com o homólogo humano. Esta proteína conservou todos os resíduos de triptofano típicos mais relevantes dos homólogos de eIF4E, bem como os resíduos carregados positivamente necessários para se ligarem aos grupos fosfato do *cap*. O EIF4E2 é também caracterizado pela presença de inserções curtas localizadas dentro da região central do eIF4E e uma inserção muito maior perto do

seu terminal C, mas a sua estrutura 3D predita também é compatível com a do eIF4E de mamíferos (DHALLIA; REIS; FREIRE; ROCHA *et al.*, 2005; YOFFE; ZUBEREK; LERER; LEWDOROWICZ *et al.*, 2006).

Recentemente foi visto que o EIF4E2 de *T. brucei* forma um complexo com um homólogo da proteína de ligação a mRNAs de histonas de mamíferos, a SLBP, possivelmente desempenhando um papel na seleção diferencial de mRNAs contendo estruturas do tipo *stem-loop*. É possível que este fator esteja então desempenhando um papel na seleção diferencial de mRNAs estruturalmente específicos (FREIRE; MOURA; BEZERRA; XAVIER *et al.*, 2018).

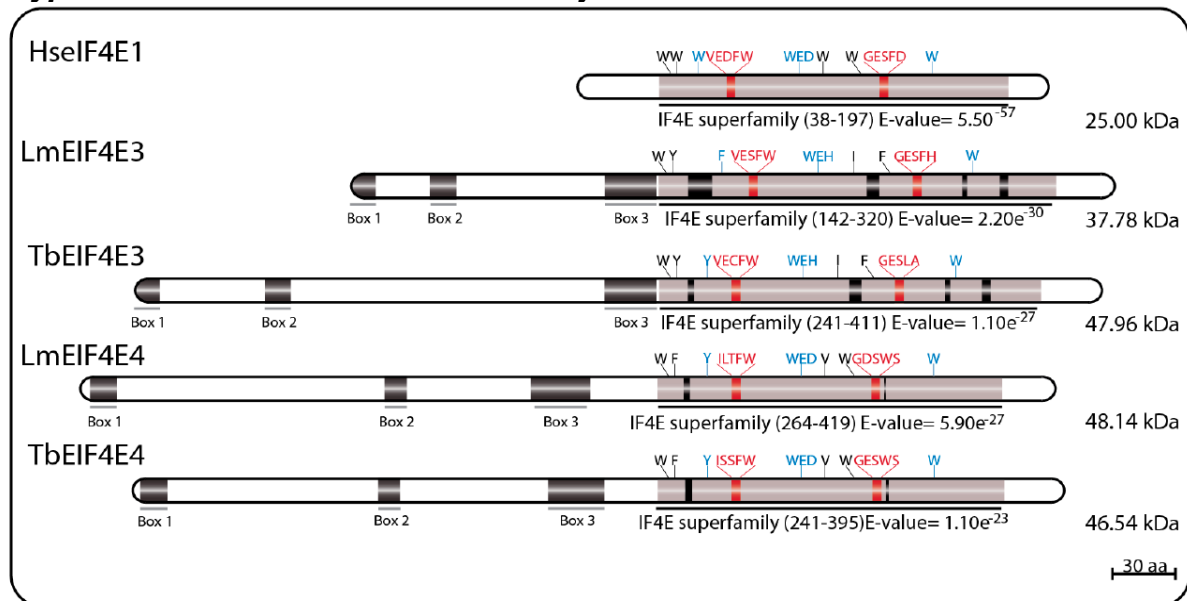
2.5.3.2.2 Grupo 2: EIF4E3 e EIF4E4

Os fatores EIF4E3 e EIF4E4 são os homólogos de eIF4E que participam dos complexos mais bem caracterizados em tripanosomatídeos (Figura 12). Tanto o EIF4E3 como o EIF4E4 são proteínas abundantes, localizadas no citoplasma e que interagem com homólogos de eIF4G (FREIRE; DHALLIA; MOURA; DA COSTA LIMA *et al.*, 2011). O complexo EIF4E4/EIF4G3 se associa ao EIF4AI e PABP1 e parece ter como principal papel sua ação na síntese proteica. No que diz respeito ao complexo EIF4E3/EIF4G4, ainda não está claro qual a sua real importância para o processo de síntese proteica como um todo, embora acredita-se que este complexo esteja responsável num papel seletivo a grupos específicos de mRNAs para síntese proteica (FREIRE; DHALLIA; MOURA; DA COSTA LIMA *et al.*, 2011; MOURA; REIS; XAVIER; DA COSTA LIMA *et al.*, 2015). Resultados em *Leishmania* para o EIF4E3 indicariam uma função dissociada da síntese proteica (ZINOVIEV; MANOR; SHAPIRA, 2012).

A proteína EIF4E3 de *L. major* apresenta 43% de similaridade com homólogos de mamíferos, mas com apenas quatro resíduos aromáticos conservados das sequências típicas de EIF4E: os resíduos de ligação ao cap (F172, W216 e W286), equivalente aos W56, W102 e W166 de mamíferos; e o resíduo de ligação ao EIF4G, W187, equivalentes ao W73 de mamíferos (FREIRE; DHALLIA; MOURA; DA COSTA LIMA *et al.*, 2011) Em ensaios de fracionamento celular e

localização subcelular foi visto que o EIF4E3 é exclusivamente citoplasmático, o mesmo local onde foi encontrado o EIF4G4 (MOURA; REIS; XAVIER; DA COSTA LIMA *et al.*, 2015). O EIF4E3 se mostrou mais abundante que todos os outros homólogos de eIF4E de tripanosomatídeos, com cerca de 2×10^4 a 10^5 moléculas por célula na forma promastigota de *Leishmania* (DHALLIA; REIS; FREIRE; ROCHA *et al.*, 2005; FREIRE; DHALLIA; MOURA; DA COSTA LIMA *et al.*, 2011).

Figura 12 - Representação esquemática dos homólogos de EIF4E3 e EIF4E4 de *Trypanosoma brucei* e *Leishmania major*.



Fonte: Adaptado de FREIRE *et al.*(2017)

Nota: As regiões centrais do eIF4E são representadas em cinza; inserções dentro dessas regiões são representadas por caixas pretas; os resíduos de triptofano envolvidos na ligação ao cap são destacados em azul; e resíduos envolvidos na ligação eIF4G são destacados em vermelho. Os Box 1, 2 e 3 representam três regiões conservadas envolvidas na ligação a homólogos de PABP e localizadas dentro da região N terminal do EIF4E3 e EIF4E4.

A ligação do EIF4E3 e EIF4G4 foi observada através de ensaios de imunoprecipitação e ambos co-precipitavam com o EIF4A1 em *T. brucei* (FREIRE; DHALLIA; MOURA; DA COSTA LIMA *et al.*, 2011), mas a interação direta entre estes fatores foi também confirmada por ensaios de duplo híbrido e de *pull-down* (MOURA; REIS; XAVIER; DA COSTA LIMA *et al.*, 2015). Em *L. amazonensis*, análises de espectrometria de massas confirmaram a co-precipitação de ambas as proteínas e o papel da interação do W187 do EIF4E3 com uma leucina (L26) no EIF4G4 (ZINOVIEV; MANOR; SHAPIRA, 2012). O papel da proteína EIF4E3 na

tradução é reforçado por ensaios de RNAi em *T. brucei*, onde sua depleção levou a morte célula tanto na forma procíclica como na sanguínea (FREIRE; DHALIA; MOURA; DA COSTA LIMA *et al.*, 2011). Um efeito letal também é observado quando é depletado o EIF4G4, mas ainda são necessários estudos pra entender melhor o papel destas proteínas na síntese proteica (MOURA; REIS; XAVIER; DA COSTA LIMA *et al.*, 2015).

Outra característica dos EIF4E3 e EIF4E4 é que ambos possuem uma extensão em suas porções N terminais não observada para outros eIF4Es (DE MELO NETO; DA COSTA LIMA; XAVIER; NASCIMENTO *et al.*, 2015). A ocorrência de múltiplos eventos de fosforilação nos resíduos N terminal de ambos EIF4E3 e EIF4E4 mostrou associação com o crescimento celular (PEREIRA; MALVEZZI; NASCIMENTO; LIMA *et al.*, 2013). Além de se ligar ao EIFG3, o EIF4E4 de *Leishmania infantum* também é capaz de se ligar diretamente a PABP1 por meio de motivos conservados em sua porção N terminal. Tanto o EIF4E4 como a PABP1 são alvos de múltiplas fosforilações em resíduos de serina-prolina ou treonina-prolina (DE MELO NETO; DA COSTA LIMA; XAVIER; NASCIMENTO *et al.*, 2015). Resultados de alinhamento de sequência sugerem que as interações entre o 5' – 3' do mRNA mediada por EIF4E4 e PABP1 desempenham um papel distinto e possivelmente dominante na iniciação da tradução em *Leishmania*. Há possibilidade de que a interação EIF4E4-PABP1 pode ser aplicada em outras espécies de tripanosomatídeos (DOS SANTOS RODRIGUES; FIRSZUK; BREEZE; CAMERON *et al.*, 2019).

Em pesquisas nos últimos anos vem sendo estudado a região N-terminal do EIF4E3, especificamente de *L. amazonensis*, onde um sítio de fosforilação em uma serina na posição 75 foi identificado que não é encontrado em outras espécies do gênero *Leishmania*. Observou-se a importância deste resíduo em estudos que avaliaram a formação de grânulos em situação de estresse nutricional, onde a proteína fosforilada nesta região, não conseguiu migrar para os grânulos de estresse, e sua interação com o seu provável parceiro funcional EIF4G4 não foi evidenciada (SHRIVASTAVA; DRORY-RETWITZER; SHAPIRA, 2019).

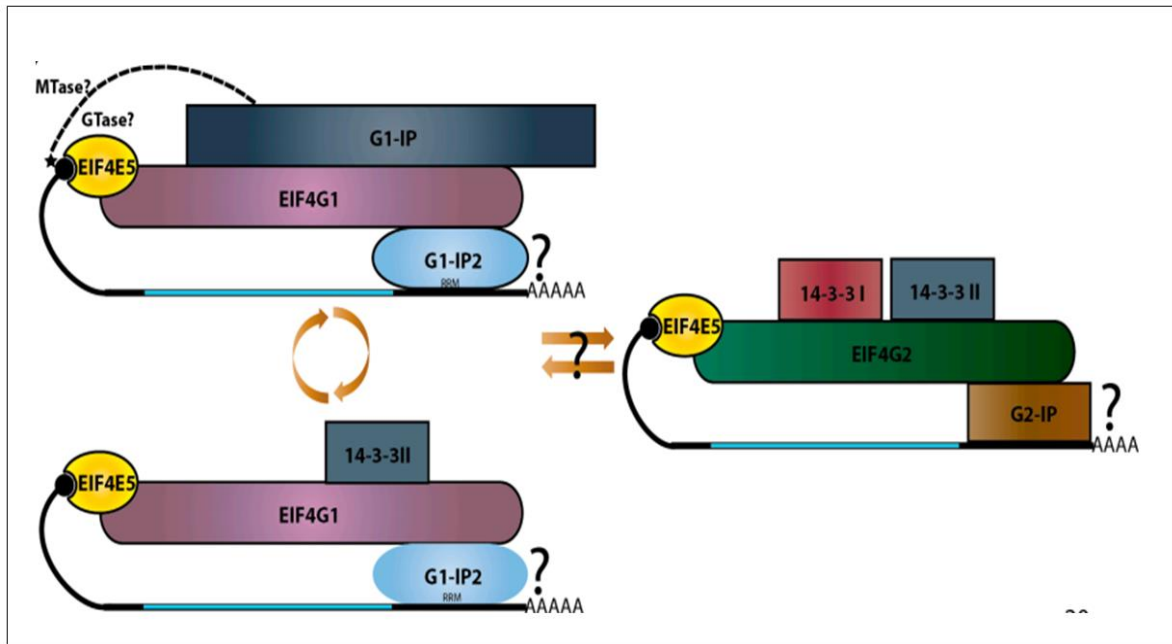
2.5.3.2.3 Grupo 3: EIF4E5 e EIF4E6

Os fatores EIF4E5 e EIF4E6 são proteínas citoplasmáticas que são capazes de se ligar tanto ao *cap* comum como ao *cap4*, típico de tripanosomatídeos. Ambos parecem apresentar estruturas típicas de eIF4E para poder fazer interação ao *cap*. Através de ensaios de ligação ao *cap* foi visto que o EIF4E5 recombinante interage tanto com o m⁷GTP quanto com o *cap4*. Estudos de imunofluorescência realizados em *T. brucei* indicam que tanto o EIF4E5 quanto o EIF4E6 estão presentes em todo o citoplasma, e ensaios de RNAi em *T. brucei* indicaram que o EIF4E5 parece não ser um fator necessário para a síntese proteica em células procíclicas. Por outro lado, a sua depleção levou a perda da motilidade celular, avaliada por meio de ensaios de sedimentação. Outro estudo que também utilizou RNAi em *T. brucei* mostrou que ocorreu um crescimento anormal das células procíclicas após a depleção do EIF4E5 (ALSFORD; TURNER; OBADO; SANCHEZ-FLORES *et al.*, 2011). A deleção dupla do gene do fator EIF4E5 em *T. brucei* não foi viável, podendo indicar a essencialidade do mesmo (FREIRE; MALVEZZI; VASHISHT; ZUBEREK *et al.*, 2014; FREIRE; VASHISHT; MALVEZZI; ZUBEREK *et al.*, 2014).

Em ensaios de interação do EIF4E5 de *T. brucei* com os cinco homólogos do eIF4G observou-se uma especificidade de ligação do EIF4E5 com os EIF4G1 e EIF4G2 (Figura 13), com a interação do EIF4E5 com o EIF4G2 sendo mais intensa quando comparada com a com o EIF4G1. Análises de proteômica e espectrometria de massas também mostraram a interação do EIF4E5 ao EIF4G1 e ao EIF4G2 *in vivo*, formando assim dois complexos do tipo EIF4F em *T. brucei*, os EIF4E5-EIF4G1 e EIF4E5-EIF4G2, além de interações com outras proteínas hipotéticas associadas (FREIRE; VASHISHT; MALVEZZI; ZUBEREK *et al.*, 2014). Uma das proteínas com a qual o EIF4E5 interagiu foi a G1-IP, uma proteína grande que contém dois domínios, de guanilil-transferase e metil-transferase. Esses domínios são encontrados em proteínas que estão implicadas na adição do *cap* e maturação do mRNA (GHOSH; LIMA, 2010). Ensaios em duplo híbrido demonstraram que a G1-IP se liga diretamente ao EIF4G1 e parece estar relacionada com a ligação do EIF4E5 ao EIF4G1 (FREIRE; VASHISHT; MALVEZZI; ZUBEREK *et al.*, 2014). Em relação ao EIF4E6, ensaios de interação em *T. brucei* mostraram uma associação específica com o EIF4G5, tanto em ensaios de duplo-híbrido quanto em ensaios *in vivo*. Foi

também observada a interação do complexo EIF4E6/EIF4G5 com proteínas com domínios relacionados a formação do *cap*, hidrolase trifosfatase e guanililtransferase. Sua depleção levou a alterações de morfologia e também a perda de movimentação celular em cultura (FREIRE; MALVEZZI; VASHISHT; ZUBEREK *et al.*, 2014).

Figura 13 - Esquema do complexo EIF4E5-EIF4G1-G1-IP-G1-IP2.

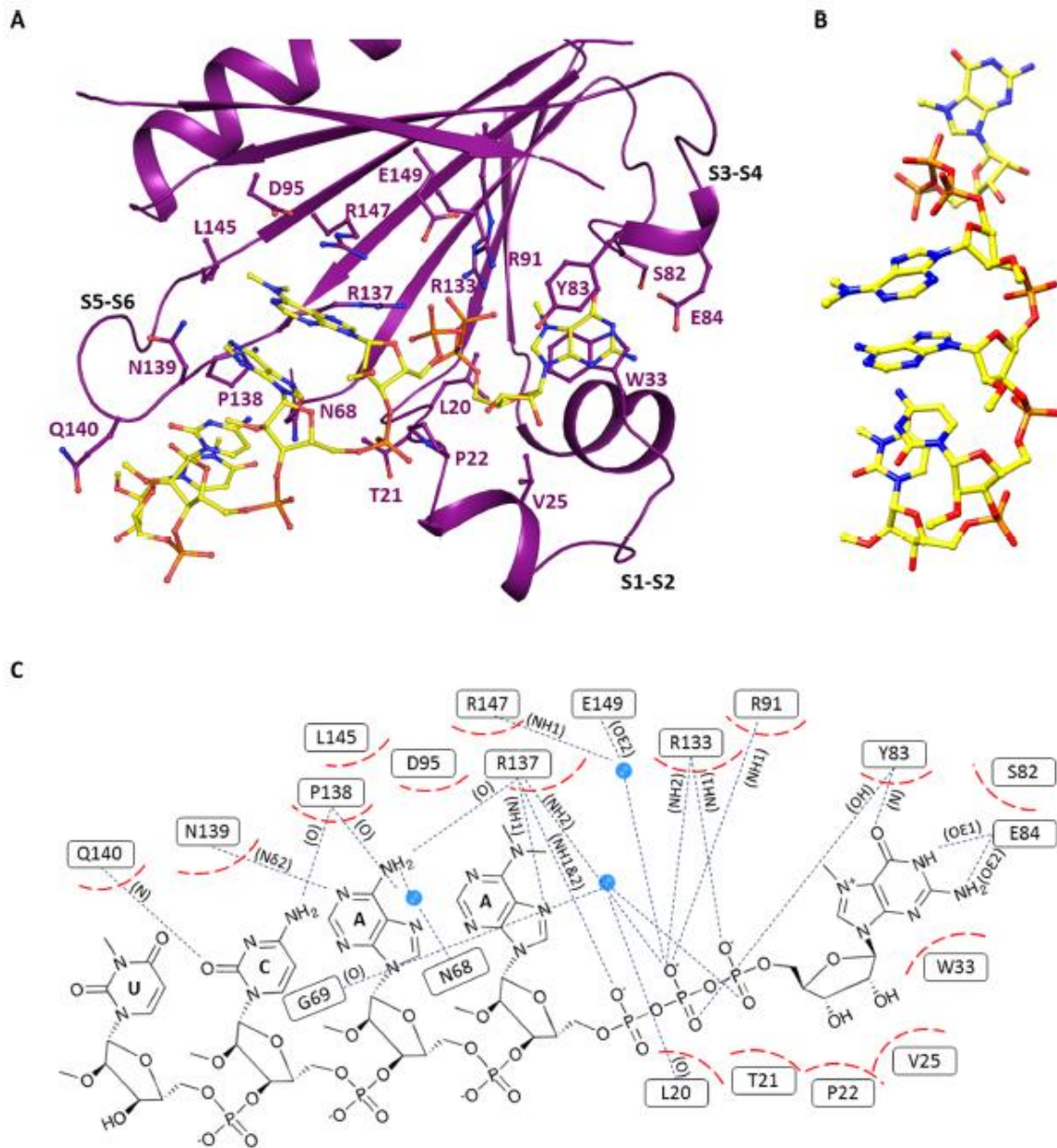


Fonte: FREIRE *et al.*(2017)

Nota: O esquema apresentado mostra o complexo EIF4E5-EIF4G1-G1-IP-G1-IP2. O complexo do EIF4E5 é composto por sua interação com o EIF4G1, que por sua vez se liga a proteína G1-IP que contém o domínio da guanil transferase/ metil transferase e também está ligada à proteína G1-IP2, que é uma proteína de ligação ao RNA. O outro complexo é composto pela interação do EIF4E5 ao EIF4G2, que também se liga a uma proteína hipotética denominada de G2-IP.

Recentemente foi obtida a estrutura do fator EIF4E5 de *Trypanosoma cruzi* (Figura 14) e através desta pode se observar de forma estrutural a interação do fator EIF4E5 com o *cap4*, este encontrado nos mRNAs de tripanosomatídeos. A interação encontrada se mostrou um pouco diferente da ligação que ocorre com o EIF4E1 de *Leishmania major*, evidenciando e corroborando com a diferença de capacidade de ligação entres os homólogos do EIF4E1 e EIF4E5. Uma vez que este estudo compara duas espécies muito diferentes, é necessário ainda, entretanto, observar a interação destes em espécies mais próximas (REOLON; VICHIER-GUERRE; DE MATOS; DUGUÉ *et al.*, 2019).

Figura 14 - Sítio de interação ao cap-4 e o fator EIF4E5 de *Trypanosoma cruzi*.



Fonte: REOLON *et al.* (2019)

Nota: **A)** Representação estrutural dos resíduos envolvidos na interação ao cap4. As regiões que conectam as regiões betas estão indicadas por S1-S2, S3-S4 e S5-S6. **B)** Estrutura tridimensional evidenciando como as bases AACU do cap4 estão posicionadas durante a ligação ao fator EIF4E5. **C)** Representação esquemática em detalhes da interação entre o fator EIF4E5 e o cap4, as regiões hidrofóbicas estão representadas tracejadas vermelhos e as ligações de hidrogênios em linhas e pontilhadas em azul. Os átomos que envolvem as ligações de hidrogênio estão indicados em parênteses, moléculas de água estão representadas por esferas azuis.

A biologia molecular do tripanosomatídeos e as peculiaridades que são encontradas para cada proteína nova estudada, em especial no caso dos fatores envolvidos na iniciação da tradução, nos leva a novas perguntas a partir de cada nova descoberta. O estudo dos fatores que atuam na síntese proteica nos leva a cada vez mais estar a um passo próximo de entender importantes aspectos particulares de uma etapa crítica para a biologia molecular e expressão gênica destes microrganismos. Nesse sentido, neste trabalho, foi proposto avançar de forma sistemática no estudo de proteínas selecionadas que parecem ter ações críticas no processo porém com particularidades nos tripanosomatídeos não observadas em outros eucariotos.

3 JUSTIFICATIVA

Os gêneros *Trypanosoma* e *Leishmania* são responsáveis por diferentes enfermidades de impacto mundial, como a Doença de Chagas, a Doença do Sono e as várias formas de leishmaniose. No Brasil, as políticas de saúde empregadas vêm sendo consideradas fracassadas, ocasionando o aumento da incidência e a expansão das áreas de transmissão, sendo que a leishmaniose visceral está distribuída em 17 dos 27 estados, com maior incidência no Nordeste, com 92% do total de casos. Considerando a ausência de ferramentas eficientes de tratamento e prevenção da leishmaniose, o desenvolvimento de novos quimioterápicos menos tóxicos e mais eficazes depende de um conhecimento maior sobre os processos biológicos básicos desses agentes etiológicos. A síntese proteica é um alvo amplamente utilizado por quimioterápicos que agem sobre diferentes grupos de organismos (CENCIC; DESFORGES; HALL; KOZAKOV *et al.*, 2011; MALINA; CENCIC; PELLETIER, 2011; MOERKE; AKTAS; CHEN; CANTEL *et al.*, 2007). Devido às diferenças biológicas no processo de tradução entre eucariotos superiores e os tripanosomatídeos, o melhor entendimento e estudo de como ocorre os processos biológicos nestes organismos ajudará no desenvolvimento de ferramentas de controle como moléculas inibidoras da síntese proteica.

4 OBJETIVO GERAL

Avaliar a função na iniciação da tradução nos tripanosomatídeos de subunidades dos complexos eIF2, eIF3 e eIF4F, pela investigação de suas interações com proteínas parceiras, modificações pós-traducionais e/ou ultraestrutura.

4.1 OBJETIVOS ESPECÍFICOS

- Analisar as diferenças gênicas e caracterizar quais as principais proteínas que interagem com as proteínas homólogas do complexo eIF2 (EIF2 γ 1 e EIF2 γ 2) e avaliar diferenças de interação entre estas.
- Identificar as interações proteicas de subunidades do complexo EIF3 (EIF3a, EIF3d, EIF3e, EIF3g, EIF3i e EIF3j) avaliando interações internas dentro do complexo e com outros parceiros funcionais.
- Avaliar o papel da proteína de ligação ao *cap* (EIF4E3) frente aos seus parceiros funcionais e da importância da serina 75 em *L. amazonenses*.
- Caracterizar a estrutura tridimensional da proteína de ligação ao *cap* (EIF4E5) e realizar estudo comparativo de sítios de ligação a prováveis parceiros funcionais em *Leishmania sp.*

5 METODOLOGIA

5.1 Análises de sequência e filogenia

Para a pesquisa de sequências ligadas aos dos EIF2Y1 e EIF2Y2 no banco de dados do TritypDB (<https://tritrypdb.org/tritrypdb/app>), e em todos os bancos de dados disponíveis pelo *National Center for Biotechnology Information* (NCBI) (<https://www.ncbi.nlm.nih.gov/>), foram avaliados um total de 40 espécies entre kinetoplastidas e outros organismos considerados ancestrais a esta família. Também foram incluídas 14 espécies representando grupos eucarióticos distintos como os Opisthokonta, Archaeplastida, Chromalveolata, SAR (grupo composto por Stramenophila, Alveolata e Rhizaria), Amebozoa e Excavata.

No banco de dados de famílias de proteínas Pfam (<https://pfam.xfam.org/>) se buscou regiões funcionais, comumente denominada de domínios, que por sua vez estão presentes em determinados grupos funcionais onde é possível identificá-los e trazer informações relevantes ao papel funcional das proteínas estudadas. Desta forma foram definidas, através de sete sequências de caráter relevante para análise obtidas por alinhamento, regiões de domínio característicos destas proteínas.

Com todas as sequências obtidas anteriormente em bancos de dados, foi feito o alinhamento pela ferramenta online MAFFT (Multiple Alignment program for amino acid or nucleotide sequence) versão 7 (<https://mafft.cbrc.jp/alignment/software/>) que como nome já nos informa, realiza alinhamento de sequências múltiplas, utilizando diferentes métodos de alinhamento. Após o alinhamento foi utilizado o trimal (<http://trimal.cgenomics.org/>) ferramenta que irá gerar um novo alinhamento após avaliar as inserções e deleções envolvidas no alinhamento, melhorando o alinhamento removendo regiões mal alinhadas, gerando um novo alinhamento só com as áreas que alinharam de forma melhorada.

Após o alinhamento finalizado, o alinhamento foi submetido ao Phylemon.2 (<http://phylemon2.bioinfo.cipf.es/>) ferramentas online de evolução molecular, filogenética, filogenômica e teste de hipóteses. Foi utilizado neste trabalho o modelo evolutivo Phyml para geração de árvores filogenéticas, obtido como melhor modelo o

Blossum 62+I+G. Com a árvore determinada, foi utilizado a ferramenta FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) para geração de árvores visuais a partir dos modelos gerados pelo Phylemon.2 e avaliado pelo *Bootstrap* para verificar a confiança da árvore gerada.

5.2 Cultivo e avaliação de crescimento celular

Células de *Leishmania infantum* e *L. amazonensis* foram mantidas em fase exponencial de crescimento através de repiques sucessivos em meio Schneider pH 7,2 suplementado com 10% de soro fetal bovino (SFB) e 0,1% v/v de Penicilina e Estreptomicina (Sigma-Aldrich Co.) em estufa BOD a 26°C. Culturas de *Leishmania* foram utilizadas durante todos os experimentos de modificação genética através da técnica de transfecção dos parasitas com as construções obtidas para expressão heteróloga e conseguinte análise do crescimento e diferenciação das diversas linhagens obtidas.

Para a avaliação e crescimento das linhagens selvagens e mutantes, células em fase estacionária de cultivo foram repassadas numa concentração inicial de 10^6 células/mL em meio Schneider, como descrito anteriormente, durante curvas de crescimento (cinco a sete dias) e observadas quanto ao número de células, morfologia e motilidade para observar a sua viabilidade celular.

5.3 Transfecção dos parasitas

Apenas formas promastigotas de *Leishmania sp.* foram utilizadas nos experimentos de transfecção. Alíquotas de $2,5 \times 10^8$ célula/mL de *Leishmania sp.* foram centrifugadas a 3.000 rpm por 5 minutos, lavadas com 5 ml de tampão HEPES-NaCl pH 7,05 (HEPES 20 mM, NaCl 85 mM, KCl 5 mM, Na_2PO_4 0,7 mM e Glicose 20 mM) e ressuspendidas em 400 μl do mesmo tampão. As células foram então repassadas para cuvetas de 2 mm (Biolabs), já contendo de 5 a 10 μg do DNA plasmidial ou 2 a 5 μg do DNA linear a ser transfectado. Após um período de 15

minutos a 4°C as células foram eletroporadas com 1 pulso de 450 Volts e capacitância de 450µF, utilizando o eletroporador *Gene Pulser Xcell™ system* (BIO-RAD). Imediatamente depois da eletroporação, foi adicionado 1 ml de meio Schneider pH 7,2 na cuveta e o conteúdo repassado para um frasco de cultura, com 5 ml do mesmo meio.

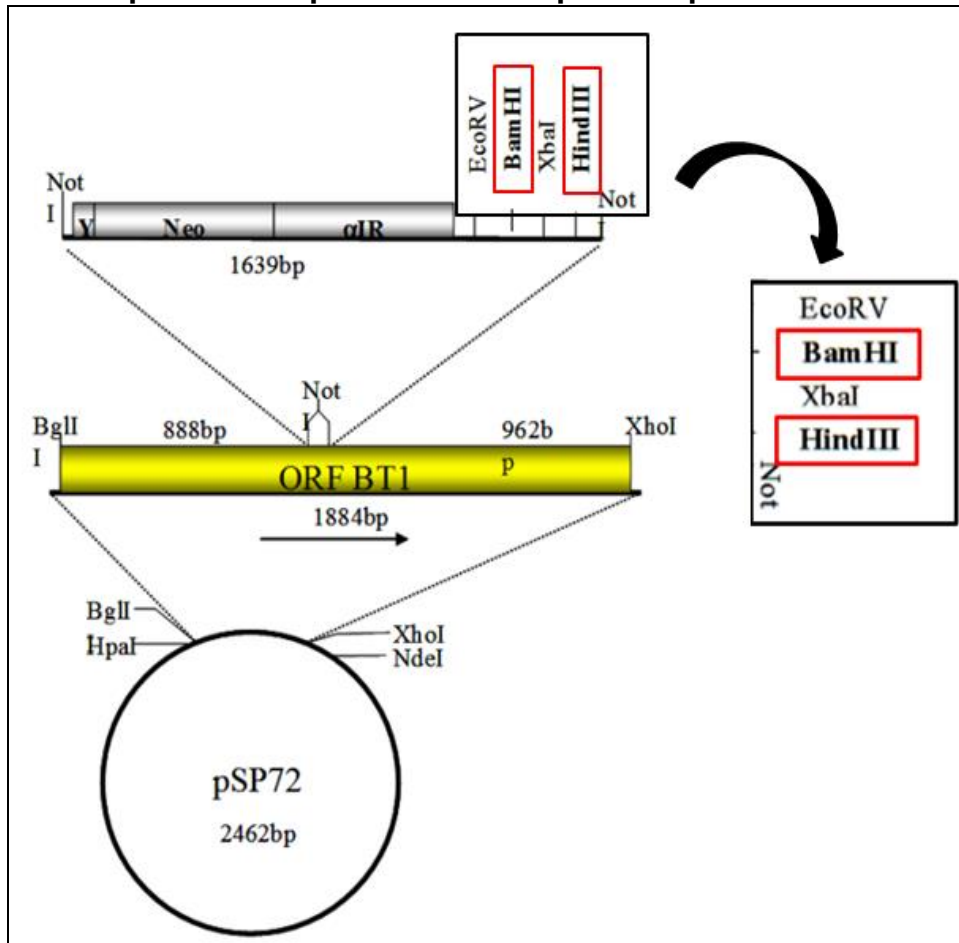
A pré-seleção foi realizada no dia seguinte, adicionando mais 5 ml de meio Schneider pH 7,2 e a adição de 50% da quantidade adequada do antibiótico específico. Após 24 horas da pré-seleção foi realizada a seleção, onde 1 ml da célula pré-selecionada foi diluída em 9 ml de meio novo adicionando a concentração adequada do antibiótico específico. Os transfectantes foram recuperados entre 7 e 14 dias após o início da etapa de seleção e foi utilizado como controle experimental células eletroporadas com água purificada no lugar do DNA.

5.4 Construções plasmidiais dos fatores de iniciação da tradução

As construções utilizadas neste trabalho foram obtidas através de diferentes trabalhos prévios. Foram utilizadas neste trabalho para expressão em células eucarióticas do gênero *Leishmania* o vetor pSP-BT1-Y-Neo-alfa-HA (Figura 15) e para expressão procariótica o pET21d (Novagen). Todas as construções utilizadas neste trabalho tiveram a integridade confirmada por digestão enzimática.

A construções para imunoprecipitação que foram construídas produzidas junto ao vetor pSP-BT1-Y-Neo-alfa-HA continham os genes *EIF2γ1* e *EIF2γ2*, do complexo EIF2 (NASCIMENTO, 2016); os genes *EIF3a*, *EIF3d*, *EIF3e*, *EIF3g*, *EIF3i* e *EIF3j*, codificantes de subunidades pertencentes ao complexo EIF3 (ASSIS, 2015); o gene *EIF4E3* selvagem e com a mutação S75A (RÔMULO; MORAES, 2015); e o gene *EIF4E5* de *L. infantum* (DE LIMA; DE LIMA CAVALCANTI; DE BRITO; DE ASSIS *et al.*, 2021). Também foi usada a construção derivada do plasmídeo pET-21d com o gene *EIF4E5* de *L. major*.

Figura 15 - Mapa do vetor plasmidial de expressão pSP-BT1-Y-Neo-alfa-HA.



Fonte: Adaptado de ADDGENE (2018)

Nota: Fusionado ao epítipo HA (Hemaglutinina). Os sítios das enzimas de restrição utilizados foram: BamHI e HindIII (em vermelho).

5.5 Sequenciamento de DNA

Após confirmação dos tamanhos dos insertos por digestão, as construções em vetor pSP-BT1-YneoαIR-HA e pET21d foram submetidas ao sequenciamento de DNA, que foi realizado utilizando o *kit BigDye Terminator v3.1 Cycle Sequencing* (Applied Biosystems®) e os produtos purificados com Etanol/EDTA (*Applied Biosystems*®), ambos os métodos seguindo as instruções do fabricante. Todos os clones foram analisados através do seqüenciador automático de DNA ABI Prism 3100 (*Applied Biosystems*®) disponível para utilização no Núcleo de Plataformas Tecnológicas do IAM / FIOCRUZ. Após o sequenciamento, as sequencias obtidas foram avaliadas para verificar a sua integridade.

5.6 Ensaios de Western-Blot

Os extratos proteicos obtidos a partir de alíquotas retiradas de cultivo de *Leishmania sp.* foram fracionados em gel SDS-PAGE 20% em concentrações de 5×10^6 a 1×10^7 células por poço e transferidos para membrana de PVDF (100 mA por 1h). A membrana foi então bloqueada com solução de TBS (20 mM Tris-HCl, 500 mM NaCl, pH 7,5), leite desnatado 5% e Tween-20 a 0,05% em temperatura ambiente por uma hora. Após o bloqueio, a membrana foi incubada com anticorpo primário com a solução de bloqueio com diluições dos anticorpos utilizados. Após três lavagens de 10 minutos em solução de TBS/Tween-20 a 0,05%, houve uma segunda incubação, desta vez com o segundo anticorpo secundário conjugado a peroxidase em diluições semelhantes às do primeiro anticorpo na solução de bloqueio, por uma hora. Esta incubação foi seguida de três lavagens de 10 minutos cada.

A visualização final foi obtida através de uma reação de quimioluminescência (ECL, do inglês *Enhanced chemiluminescence*) e exposição a filme *Biomax Light* (Kodak). Tal reação usa uma solução contendo 25 mL de luminol a 1,2 mM diluído em solução de Tris-HCl 0,1 M de pH 8,5, adicionado de iodofenol para uma concentração final de 0,4 mM e de peróxido de hidrogênio para 0,03%. A membrana foi exposta ao filme em tempos que variaram de 30 segundos a 5 minutos de acordo com anticorpo utilizado.

5.7 Preparação de lisados celulares e ensaios de imunoprecipitação

Para a obtenção da fração citoplasmática de células de *Leishmania sp.*, cerca de 2×10^8 células promastigotas foram coletadas por centrifugação e lavadas com PBS suplementado com 10% de inibidor de protease (Roche®). Em seguida, estas células foram ressuspensas de 400 a 800 μ l de tampão de lise (20 mM de HEPES-KOH pH 7,4/ 75 mM acetato de potássio/ 4mM de acetato de magnésio/ 2 mM de DTT, suplementado com inibidores de protease). Para a lise, as células foram

submetidas à diferença de pressão, utilizando um cavitador (para realizar a ruptura celular), e esta lise monitorada por microscopia óptica. As amostras foram centrifugadas a 12000 rpm durante 10 minutos para remover restos celulares e o sobrenadante foi armazenado a -80°C.

Para a imunoprecipitação, como todas as proteínas de interesse possuíam o epítipo HA, utilizou-se para cada 250 µl de lisado, 0,2 a 0,5 mg de microesferas magnéticas contendo anticorpos anti-HA (Pierce™) de acordo com o protocolo dos fabricantes. As microesferas foram lavadas três vezes com tampão PBS e em seguida foram incubadas com os lisados celulares produzidos como descrito anteriormente, de 30 a 60 minutos. Após a incubação, o sobrenadante (depletado) foi guardado e a resina, lavada mais três vezes com tampão PBS. Alíquotas do lisado celular original, do lisado depletado após a imunoprecipitação e as microesferas (imunoprecipitado) foram ressuspensas em tampão de amostra para SDS-PAGE, fervidos por 5 minutos e guardados a -80°C para realizar o *western blot* de confirmação de imunoprecipitação e as demais análises.

5.8 Análise por cromatografia líquida acoplada a espectrometria de massas (LC-MS/MS)

A espectrometria de massas foi realizada com 30 µl das IPs (volume total resultante após testes de confirmação por *western blot*), que foram fracionadas em gel SDS-PAGE 12,5% até atingir a linha limite entre as duas fases do gel. O gel foi corado com azul de Comassive Blue 250-R de 10 a 30 minutos e depois descorado com solução solvente, sendo necessárias várias trocas para total remoção do excesso de corante, o conjunto de proteínas foi excisado do gel e enviados para análise por espectrometria de massa na plataforma tecnológica de Espectrometria de Massa do Instituto Carlos Chagas Fiocruz, Paraná. Ao chegar no destino as amostras foram digeridas com tripsina e os peptídeos submetidos a análise por LC-MS/MS (MS2) no sistema Thermo Scientific Easy-nLC 1000 Ultra Performance Liquide Chromatography (UPLC) acoplado à um espectrômetro de massas LTQ-Orbitrap XL ETD (Thermo Fisher Scientific).

Os dados brutos provenientes da análise por LC-MS/MS das IPs foram processados pelo programa MaxQuant v1.3.0.3 e os espectros MS/MS foram submetidos à busca contra o banco de dados de *Leishmania infantum* (strain: JPCM5) e *Leishmania mexicana* (strain: MHOM/GT/2001/U1103) disponível em TriTrypDB, para a identificação de proteínas. A análise se deu inicialmente pela normalização das intensidades de cada amostra geradas em amostras em triplicatas. Para isso, após a soma das intensidades de cada amostra, foi feito a razão do total de intensidade da amostra com maior intensidade pela intensidade da amostra testada, gerando um fator de normalização do total de intensidade da amostra, onde cada amostra foi multiplicada pelo respectivo fator de normalização. Após a normalização, e a partir dos valores de peptídeos únicos e das intensidades, foram calculadas tanto as médias entre as replicatas do controle quanto das amostras analisadas. Com as médias obtidas, foi calculada a razão entre a média da intensidade da amostra a ser estudada sobre a média da intensidade do controle. Foi assim obtido o valor em número de vezes que a proteína está mais enriquecida do que no controle.

Para maior robustez da análise foi realizado corte das proteínas que obteve o número de peptídeos únicos menor que 2 em pelo menos duas das triplicatas e valores de intensidade próximos de zero em pelo menos duas das réplicas. Foi removido da análise também todas as proteínas que no valor de intensidade média estivesse 1000 vezes menor que o valor da proteína alvo. Após os pontos de corte as proteínas foram classificadas como acima de 100 vezes, proteínas que não apresentavam intensidade no controle se comparado a amostra, proteínas acima de 20 vezes mais enriquecida que no controle, dez vezes, quatro vezes e em alguns casos duas vezes mais enriquecida que no controle foram mantidas para fins de análise e comparação.

5.9 Cristalografia do fator EIF4E5

5.9.1 Purificação de proteínas com cauda histidina

A construção pET21a-EIF4E5 de *L. major* foi utilizada para transformação de *Escherichia coli* linhagem BL21 star (Invitrogen). Após crescimento, em uma densidade óptica de 0,5, foi feita a indução da expressão da proteínas recombinante

com IPTG (β -D-1-tiogalactopiranosídeo de isopropílico) em uma concentração de 0,5 mM por 3 horas a 37°C em meio LB. As células foram coletadas por centrifugação e lisadas por sonicação, com seis ciclos de 30 segundos de sonicação por 30 segundos de repouso, sob amplitude de 40% em gelo e água, em aparelho sonicador MSE Soniprep 150. As proteínas superexpressas foram purificadas utilizando-se de colunas de afinidade e colunas catiônicas em sistema automatizado de purificação FPLC AKTA UV-900 em tampão A (50mM Tris-HCl, 300mM NaCl e Imidazol a 20mM) e para a eluição linear, em coletor automático P-950, usando tampão B (tampão A com KCl 1M). Após eluição das frações contendo as proteínas, as frações contendo a proteína de interesse foram analisadas em SDS-PAGE e então reunidas e dialisadas em tampão A com 20% de glicerol.

5.9.2 Gel filtração

Após a purificação das proteínas por afinidade elas foram separadas por gel filtração, para exclusão de tamanho de proteínas contaminantes que vieram na purificação, onde foram utilizadas colunas HiTrep 16/60 Sephacryl S-300 (GE-Healthcare), Superose 12 10/300 (GE-Healthcare) ou coluna Superose 6 10/30 (GE-Healthcare) conforme indicações do fabricante em FPLC AKTA purifier UV-900, coletor automático P-900 (GE-Healthcare).

5.9.3 Cristalização

Para triagem de cristais do fator EIF4E5 foram utilizadas preparações de proteínas recém purificadas (sem congelamento) para incubação pelo método de difusão de vapor em gota sentada, com soluções providas por testes comerciais (Hampton Research) PEG/Ion, INDEX; JCSG I e II, Wizard I, II e III e Crystal Screen Cryo. A condição de cristalização encontrada e reprodutível foi de 10mg/mL de proteína em 0.2 M de Citrato de potássio tribásico monoidratado, 20% (p/v) de polietilenoglicol (PEG) 3350. identificada com auxílio de robô nanopipetador (Mosquito®) (0,05:0,05 μ L) (40 μ L de solução de precipitação no reservatório) e incubação a 20°C.

Reproduções dos cristais foram realizadas sob método manual de pipetagem (1:1 μ L), com 40 μ L de solução de precipitação no reservatório. O experimento de otimização de cristais foi realizado modificando os parâmetros de concentração de PEG 3533 (10, 15, 20, 25% (p/v)) e de proteína (2, 8, 10, 15 mg/mL). Análises de difração dos cristais obtidos foram realizadas em difratômetro Rigaku. A coleta de dados foi realizada utilizando sincrotron APS, Todas as etapas de cristalização foram feitas em colaboração com laboratório de proteômica e engenharia de proteínas do Instituto Carlos Chagas, Fiocruz-PR.

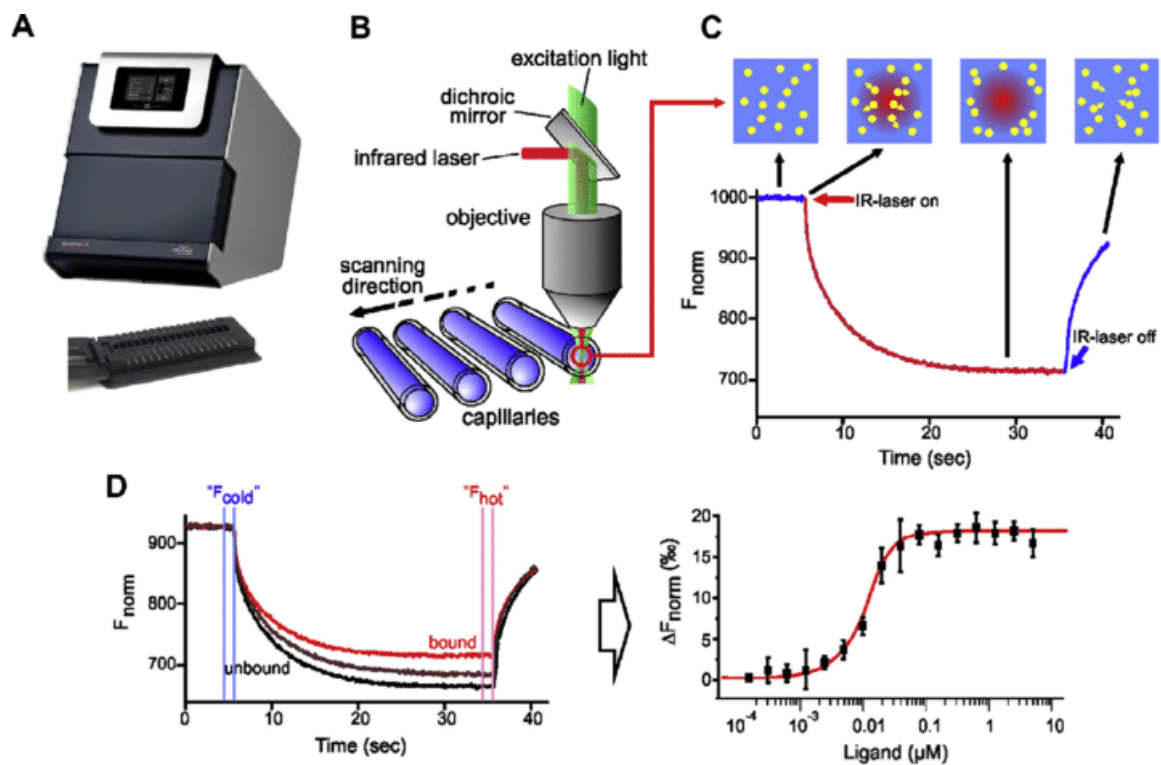
5.10 Ensaio de interação entre o fator EIF4E5 e os *cap0* e *cap4* através de termoforese em microescala (MST)

A análise da interação entre o fator EIF4E5 e o *cap4* foi realizada através da técnica de termoforese em microescala (*microscale thermophoresis*, MST). A técnica tem como princípio analisar o movimento das moléculas quando submetidas a um gradiente de temperatura. O gradiente é gerado por um laser infravermelho (IR), cuja radiação as moléculas de água presentes na solução da amostra conseguem absorver, o que desencadeia o fluxo de moléculas. Esse fluxo é decorrente de mudanças de tamanho, carga ou conformação que podem ocorrer quando as moléculas interagem. A quantificação da interação é feita através da quantificação de fluorescência emitida pelas moléculas, tornando-se necessário que uma das moléculas esteja ligada a um marcador fluorescente. O equipamento Monolith NT.115 (NanoTemper Technologies) utilizado para esta técnica é constituído de uma luz visível usada para excitação do fluoróforo, e de um laser infravermelho acoplado no mesmo caminho dessa luz. Com isso, é possível medir a intensidade de fluorescência no mesmo ponto em que a radiação é focada na amostra, gerando o fluxo das moléculas. Os gráficos referentes à interação das moléculas são gerados de forma a calcular a fluorescência relativa sobre o tempo em segundos, como pode ser observado na Figura 16. É possível também observar os intervalos de tempo referentes aos estados de salto de temperatura (*temperature jump*, *T jump*), e de termoforese. O salto de temperatura se refere ao estado do primeiro segundo em que o laser é ativado. Já o fluxo das moléculas através do gradiente de temperatura se refere ao estado da termoforese, sendo esta geralmente avaliada no intervalo de

tempo que apresenta maior distância calculada entre as duas curvas geradas referentes à proteína com ligante e à proteína sem ligante.

Foi analisada a interação entre a proteína recombinante EIF4E5 e moléculas de GTP (não ligante), como controle negativo, m7GTP (*cap0*) e o *cap4*. Os experimentos foram realizados em triplicata. As curvas geradas foram analisadas para se avaliar a constante de dissociação (k_d).

Figura 16 - Representação do mecanismo do ensaio de termoforese em microescala.



Fonte: NanoTemper Technologies

Nota: **A)** equipamento utilizado para o MST, **B)** esquema representativo de como ocorre a detecção nos capilares, **C)** esquema de como é formado o gráfico de excitação, **D)** gráficos com dados comparativos da molécula com ligante e sem ligante.

6 RESULTADOS

Com base em conhecimentos prévios obtidos em tripanosomatídeos para os três principais complexos envolvidos na iniciação da tradução, os eIF2, eIF3 e eIF4F, esse trabalho buscou preencher lacunas no conhecimento a respeito de cada um destes complexos a partir do estudo de subunidades individuais. Nesse sentido, a abordagem padrão utilizadas para boa parte dos trabalhos realizados e descritos aqui fez uso de ferramentas geradas em trabalhos prévios. Essas ferramentas são construções plasmidiais que permitem a expressão, de forma ectópica, das subunidades selecionadas em células transgênicas de *Leishmania* sp. Todas estas construções levam a produção das respectivas proteínas fusionadas com os nove aminoácidos do epítipo HA (YPYDVPDYA) na sua extremidade C-terminal. A presença deste epítipo permitiu a avaliação da expressão de cada proteína utilizando um anticorpo monoclonal anti-HA comercial. Após a confirmação da expressão de cada proteína, foram produzidos lisados celulares em condições de ausência de detergente, com os lisados utilizados para ensaios de imunoprecipitação utilizando microesferas magnéticas ligadas ao anticorpo anti-HA. Para cada proteína imunoprecipitada, seus parceiros proteicos co-precipitados foram identificados por espectrometria de massa. Estas abordagens foram utilizadas para se estudar os dois homólogos da subunidade eIF2 γ do complexo eIF2, homólogos de seis subunidades do complexo eIF3 e o EIF4E3, um dos seis homólogos de eIF4E (subunidade do complexo eIF4F) de tripanosomatídeos. Um outro aspecto investigado de proteínas selecionados, e também com o auxílio de espectrometria de massa e das ferramentas já citadas, foi a ocorrência e o efeito de mutação pós-traducionais do tipo fosforilação sobre o homólogo EIF4E3 em *Leishmania amazonenses*.

Apenas no caso específico do EIF4E5, este trabalho também investigou a estrutura dessa proteína e sua interação com análogos sintéticos de seu substrato, o *cap* 5' dos mRNAs. Esta investigação fez uso de todo um conjunto novo de ferramentas para a expressão da respectiva proteína recombinante em *Escherichia coli*, seguida de sua purificação, utilização em ensaios de interação proteína-proteína e cristalização, para definição de sua estrutura. Cada conjunto de

resultados, associados aos diferentes complexos e/ou proteínas são descritos de forma compartimentalizada nos tópicos a seguir.

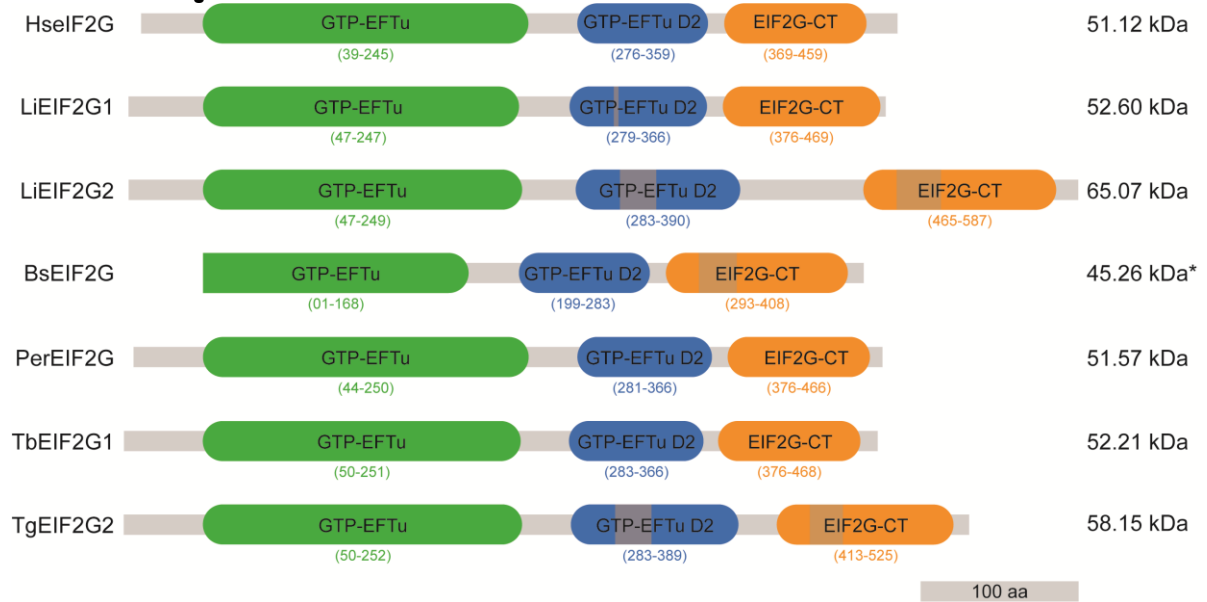
6.1 Análise comparativa dos dois homólogos de EIF2 γ de *L. infantum*

6.1.1 Alinhamento dos EIF2 γ 1 e EIF2 γ 2 de espécies de kinetoplastidas e espécies representativas.

Para se iniciar os estudos com o complexo eIF2, e buscando entender o significado funcional da presença de dois homólogos para sua subunidade central, o eIF2 γ , foram realizadas análises *in silico* a partir da busca por homólogos desta proteína em diferentes espécies de tripanosomatídeos e organismos relacionados. Primeiro foi possível observar a presença de homólogos desta proteína em todos os outros eucariotos investigados com estes em geral possuindo similaridades maiores com o EIF2 γ -1, como por exemplo o homólogo de *Bodo saltans*, espécie precursora ancestral, embora não seja da família Trypanosomatidae. Já o EIF2 γ -2 parecer ser exclusivo dos tripanosomatídeos, como melhor descrito abaixo.

Através do alinhamento é possível observar a presença de duas regiões com os domínios GTP-EFTu e GTP-EFTu D2 (Figura 17 e Figura 18). Estes são domínios presentes em fatores de alongamento responsáveis por facilitar a seleção e catalisar a ligação de um aminoacil-tRNA ao sítio A ribossomo. Também foi possível observar a região C terminal típica de fatores de tradução do complexo EIF2 γ (EIF2G-CT). Deste alinhamento foram observados ainda regiões de inserção em alguns genes, inferindo uma aproximação perante a árvore filogenética. De acordo, com as divergências encontradas nas sequências lineares, esses homólogos podem diferir quanto à capacidade de interação com o tRNA_i.

Figura 17 - Alinhamento esquemático dos domínios encontrados em fatores da família EIF2 γ .



Fonte: Autoria própria (2021)

Nota: Alinhamento esquemático de 7 sequências de proteínas homólogas do eIF2 γ , sendo: HseIF2 γ = *Homo sapiens*; LiEIF2 γ 1 e LiEIF2 γ 2 = *Leishmania infantum*; TbEIF2 γ = *Trypanosoma brucei brucei*; BsEIF2 γ = *Bodo saltans*; PerEIF2 γ = *Perkinsela sp.*; e TgEIF2 γ = *Trypanosoma grayi*. Em tracejado está delimitada os domínios identificados.

Figura 18 - Alinhamento dos homólogos EIF2 γ -1 e EIF2 γ -2 com espécies de Kinetoplastida e espécies modelo utilizando o Pfam.

HseEIF2 γ	1	MAGGEAGV	---	TLGQPH	---	-----	LSR	QDLTTL	LDVTK	LTPLS	HEVLS	RQATINI	IGTI	GHVAHGKSTV	VKALSGVHTV	67
LiEIF2 γ 1	1	M--ABEDVAD	FTLDDEKVID	T---	DQGLAK	QDFSKIDLDE	LNVDTFEVMS	RQATINV	GTI	GHVAHGKSTV	VKALSGVKTQ	75				
TbEIF2 γ	1	M--SNDVAD	FTIDETTLSS	DPTRDC	GLAA	QNL	SAMLDK	MSVET	FEVMS	RQATINI	GTI	GHVAHGKSTV	VKALSGVKTQ	78		
BsEIF2 γ	1	---	---	---	---	---	---	---	---	---	---	---	---	---	1	
PerEIF2 γ	1	M--EVDVAC	E--DPA-VD	G-RVNK	GLKV	QDLKNC	DPK	LHPLS	EELIS	RQATINI	GTI	GHVAHGKSTV	VKALSGVKTQ	72		
LiEIF2 γ 2	1	M--ABEDVAD	FTLDDEKVID	T---	DQGLAK	QDFSKIDLDE	LNVDTFEVMS	RQATINV	GTI	GHVAHGKSTV	VKALSGVKTQ	75				
TgEIF2 γ 2	1	M--GNDVAD	FTLDESTFGK	NEQNDR	GLAV	QNI	IGTL	LDL	LDK	ITVET	FEVMS	RQATINI	GTI	GHVAHGKSTV	VKALSGVKTQ	78
GTP-EFTu																
HseEIF2 γ	68	RFKNELERNI	TIKLGYANAK	IYKLLDDPSCF	RFECYRSCGS	STPDEFPTDI	PCTKGNFKLV	RHVSFVDCPG	HDILMATMLN	147						
LiEIF2 γ 1	76	KFHRDAVMNI	TIHLGYANAK	VYQCE--TCP	RPTCYQTYPS	SQLDSTPC--	PNCGETMLLK	RHFSFVDCPG	HDVLMATMLN	151						
TbEIF2 γ 1	79	KYHRDAVMNI	TIHLGYANAK	VFRCD--SCE	VFAAYHAFPS	SQPKTGC--	ETCCGAPLTLK	RHFSFVDCPG	HDVLMATMLN	154						
BsEIF2 γ	1	---	---	---	---	---	---	---	---	---						
PerEIF2 γ	73	RFKSELVQNI	TIKLGYANAK	IFKCE--TCE	AFCYQALGS	RD--DEAKC--	QFCGNPLKLV	RHVSFVDCPG	HDILMATMLN	158						
LiEIF2 γ 2	76	KFHRDAVMNI	TIHLGYANAK	VYQCE--TCP	RPTCYQTYPS	SQLDSTPC--	PNCGETMLLK	RHFSFVDCPG	HDVLMATMLN	151						
TgEIF2 γ 2	79	KYHRDAVMNI	TIHLGYANAK	VFKCD--KCE	LEAAYHAFPS	SQLDKTDC--	ETCCGSPLTLK	RHFSFVDCPG	HDVLMATMLN	154						
GTP-EFTu D2																
HseEIF2 γ	148	GAAVMDAALL	LIAANESFPQ	PQTSSEHLAAL	EIMKLRKHLI	LQNKIDLKVE	SOAKBOYEQT	LAFVQGTVA-	--EGAP	IPI	224					
LiEIF2 γ 1	152	GAAIMDAALL	LIAANESFPQ	PQTEHLKAV	EIMKLRHLV	LQNKIDLKVE	VKAHQYRHHV	RAYLDN-IA-	--LNVP	IVPI	227					
TbEIF2 γ	155	GAAIMDAALL	LIAANESFPQ	PQTEHLKAV	EIMKLRHLV	LQNKIDLKVE	VHAQDYHKT	RDYLDSTIG-	--LSAP	IPI	231					
BsEIF2 γ	69	GAAIMDAALL	LIAANESFPQ	PQTEHLKAV	EILQLKRLI	LQNKIDLKVP	VQAIPOYQT	VQFLKSTIF	SNATS	VVPI	148					
PerEIF2 γ	151	GVSIMDAALL	LIAANESFPQ	AQTRHLCAI	ETRLRNLI	LQNKIDLKVE	NEAEANRDAI	SKFVQSTPA-	--DSAP	VPI	227					
LiEIF2 γ 2	152	GAAIMDAALL	LIAANESFPQ	PQTEHLKAV	EMIGVSSLIV	LQNKIDLKVE	ERAAACHSII	HHYLATKTAY	--AKAP	VPI	229					
TgEIF2 γ 2	155	GAAIMDAALL	LIAANESFPQ	PQTEHLKAV	EVLVCSLLI	LQNKIDLKVE	EAHEVQQSVV	RYLRSRTNY	--GGAP	IPI	232					
GTP-EFTu D2																
HseEIF2 γ	225	SAQLKYNIEV	VCEYLVKK--	IPVPRDFTS	EPRLIVRSEF	DVNKPGC--	EVDDLKGGVA	GGSIILKGVLK	VGQEI	EVVRCG	299					
LiEIF2 γ 1	228	SAQLKRNVDY	LLEYLH--	IPVPRQLKA	PVRMTVVRSEF	DVNKPGEG--	DIESLKGGVA	GGIVTQGVVK	VGQVVE	IRPG	302					
TbEIF2 γ	232	SAQLKRNVDY	LLEYLH--	IPVPRQLKA	PVRMTVVRSEF	DVNKPGEG--	DIESLKGGVA	GGIVTQGVVK	VGQVVE	IRPG	306					
BsEIF2 γ	149	SAALHCGIDV	VLQYITH--	LPNPQRFLN	PFRYLVVRSF	DVNKPG--	GDLPPHGRH	RGSIEQGVV	LDGDI	IRPG	222					
PerEIF2 γ	228	SAQLGYNEDI	LCEYLVNETF	IPVPRHQFTL	PAKMYVVRSEF	DVNKPG--	ESSQLKGGVA	GGITQGVV	VGQEI	IRPG	304					
LiEIF2 γ 2	230	SAQLKYNIEV	VCEYLVKK--	IPVPRQLKA	PVRMTVVRSEF	DVNKPG--	EVDDLKGGVA	GGSIILKGVLK	VGQEI	IRPG	306					
TgEIF2 γ 2	233	SAQALANDY	LLEYLH--	IPVPRQLKA	PVRMTVVRSEF	DVNKPG--	EVDDLKGGVA	GGSIILKGVLK	VGQEI	IRPG	306					
GTP-EFTu D2																
HseEIF2 γ	300	TVSK--	-----	DS	EKLIMCKPIF	SKIIVSLFABH	NDLQMAAPGG	LIVGVTKIDP	TLGRADRMVG	355						
LiEIF2 γ 1	303	LVRSR	-----	SNEQ	GERFTYSPLK	TRVTILKAEN	NVLQMAIPGG	LIVGVTLDP	TLTRQDKMG	362						
TbEIF2 γ	307	QVHTQT	-----	---	GTFNSTPLR	TRALTLKABD	NSLQMAIPGG	LIVGVTLDP	TLTRQDKMG	362						
BsEIF2 γ	223	LNNTSD	-----	---	NNNVWSPER	TRVTSRTEG	ASLPEAVPGG	LIVGVTLDP	SLTRNQRCG	279						
PerEIF2 γ	305	LMIKPQ	-----	NG	SERMHCQPIH	TRVVSICABD	NLQMAIPGG	LIVGVTMDDP	SLTRQDKMG	362						
LiEIF2 γ 2	307	LLVLRRRDGV	LPRRADVLSG	RELAYVASPP	PEELYAVPLR	ITVVRLQABK	NELOMAIPGG	LIVGVTLDP	SLTRQDKMG	386						
TgEIF2 γ 2	307	LALTRRRDGV	LPSRTGMLSG	QELSFISGPP	PEVHMVPLR	ITVRSILMSEA	NLQMAIPGG	LIVGVTLDP	SLTRQDKMG	386						

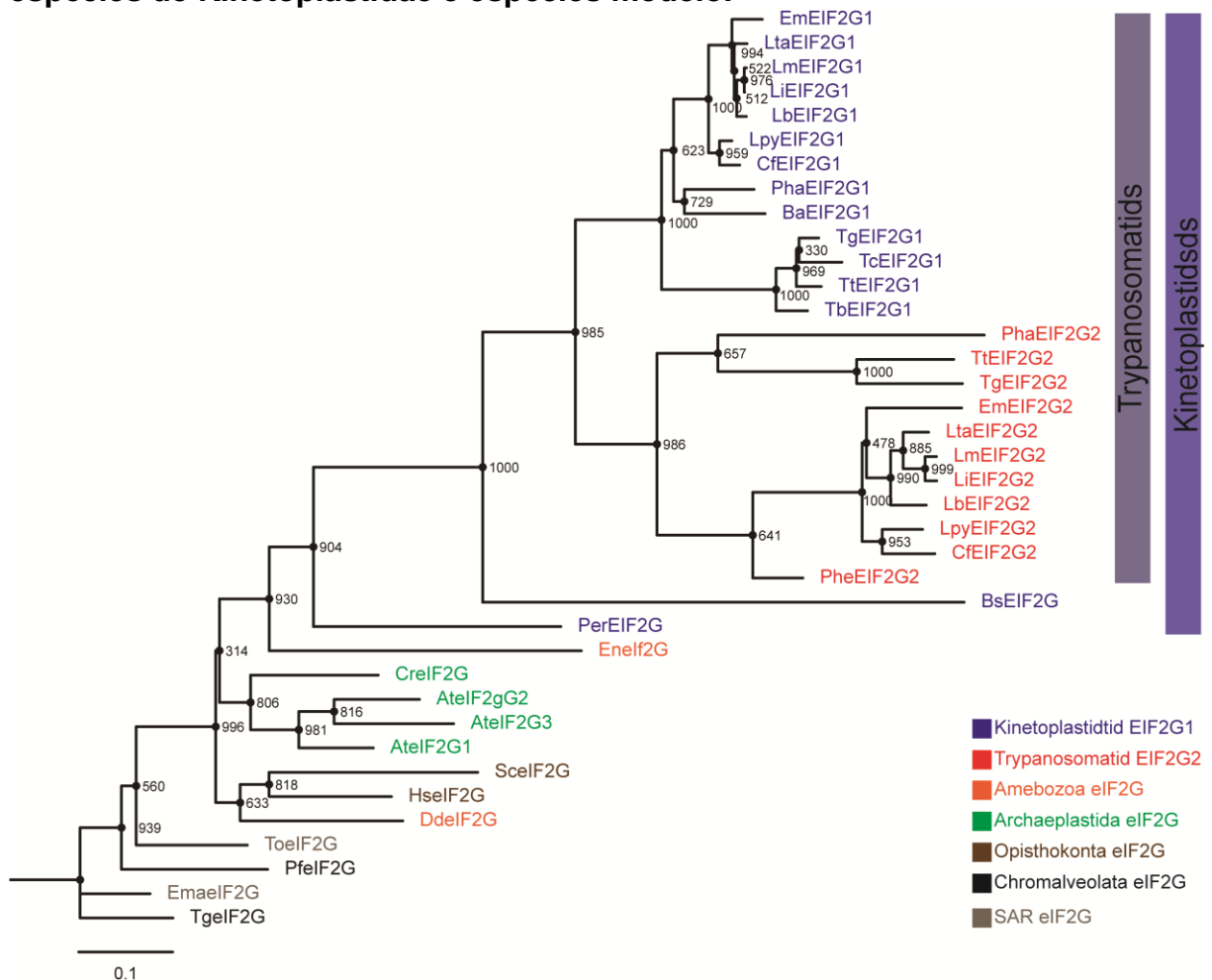
HseEIF2 γ	356	QVL	---	---	---	---	GAVGA	LP	---	EIFTE	370					
LiEIF2 γ 1	363	NML	---	---	---	---	GEET	LP	---	BVYCE	377					
TbEIF2 γ	363	HMI	---	---	---	---	GEET	LP	---	QVYAE	377					
BsEIF2 γ	280	SIC	---	---	---	---	PPSS	LP	---	DVFIE	294					
PerEIF2 γ	363	QVL	---	---	---	---	GDVGT	LP	---	DIFDE	377					
LiEIF2 γ 2	387	QVVRVSRGP	VWRSRGTRCS	LRRCDSPPAV	AEAASASTTP	ATTTTGDATP	TPSGAVLPCQ	SPAPAPVAVE	CHTGMQVYQE	466						
TgEIF2 γ 2	387	QVLR	---	---	---	---	ATTHDGDGSS	ADNN	---	DDGMQVYQE	414					
eIF2-CT																
HseEIF2 γ	371	LEISYFLLR	LLGW-RTEGD	-----	-----	-----	KKAARKVQK	LSKNEVMVNV	IGSLS	TGGRV	417					
LiEIF2 γ 1	378	LEVQYLFSE	MVGAKSVDGK	-----	-----	-----	STAKRVQK	LNVLESIQIN	VGILT	TAGATV	425					
TbEIF2 γ	378	LEVQYLFSE	MVGAKSVDGK	-----	-----	-----	RNAKRVQK	LNVELESQIN	VGILT	TAGATV	424					
BsEIF2 γ	295	LEVQYFLTH	VVGI-PRSPI	GPSTSSSSG--	SSGAAQVG	PSSRFLPG--	AFLQHRVQK	LIVGETHERIT	CRAYT	TAAATV	366					
PerEIF2 γ	378	IDVSFFLMRQ	YVGV-KTDAS	-----	-----	---	EHVRVQK	LQEGELIQVN	IGALS	TGGRV	424					
LiEIF2 γ 2	467	VVLEFFLRE	LLGL-AAPRP	GRAKPHRRDR	RDFNGGAD	SYRVHPDRV	VSSLHRVVP	DREDESILS	VGILT	TAAATV	542					
TgEIF2 γ 2	415	VELQFLRPA	LLGW-SRDRS	GSAVTTTT--	NTAAN--	AR--NAGVM	VPSRYRVQK	LKDGENVVLS	INILT	TAAATV	483					

HseEIF2 γ	418	SAVKA--DLG	KIVLTNPVCT	E-VGEKIALS	RRVEKHWRLI	GWGQIRRGVT	IKPTVDDD-	472								
LiEIF2 γ 1	426	LNITQDPEIA	KLEIVTPVCC	S-TGEQVAIS	RMVDKTRFLI	GWGTIRRGVP	MKQSS----	479								
TbEIF2 γ	425	VSIKSPDIA	KLIVTPVCC	T-MDEQIATS	RIVEKNFRLI	GWGTIRRGVP	VKLN----	477								
BsEIF2 γ	367	LVRRG--ELA	KLRLRNPICC	S-IDDKITIS	RMVERKHWRLI	GWGTIRRGVP	AAELK----	418								
PerEIF2 γ	425	VKTKA--NVA	RLILTQPVCT	S-VGEKIALS	RRVDKHFRLI	GWGEVKKGMA	LST----	474								
LiEIF2 γ 2	543	LRTSRHAGRA	ICRDESPLSA	DPPHQRIVLA	RVVDRKRVVY	GWGTIRRGVP	VKLLPESDQ	601								
TgEIF2 γ 2	484	LRTAH--GKA	KVRLECQVYC	A-CGDRVALA	RFFHRRKIRLI	GWGVVVRGVP	VRAVD----	535								

Através das análises de domínios entre as proteínas e de árvore filogenética é possível concluir que anterior aos tripanosomatídeos só é encontrado nas espécies analisadas apenas um homólogo de eIF2 γ com similaridade ao EIF2 γ -1 (Figura 19). O EIF2 γ -2 parece ter surgido nos tripanosomatídeos, mas após sua separação (aproximadamente 450 milhões de anos) a maioria das espécies do gênero

Trypanosoma sp. teria perdido o EIF2 γ -2, sendo mantido apenas no *Trypanosoma grayi* e no gênero *Leishmania* e espécies próximas.

Figura 19 - Arvore filogenética dos homólogos EIF2 γ 1 e EIF2 γ 2 com outras espécies de Kinetoplastidae e espécies modelo.



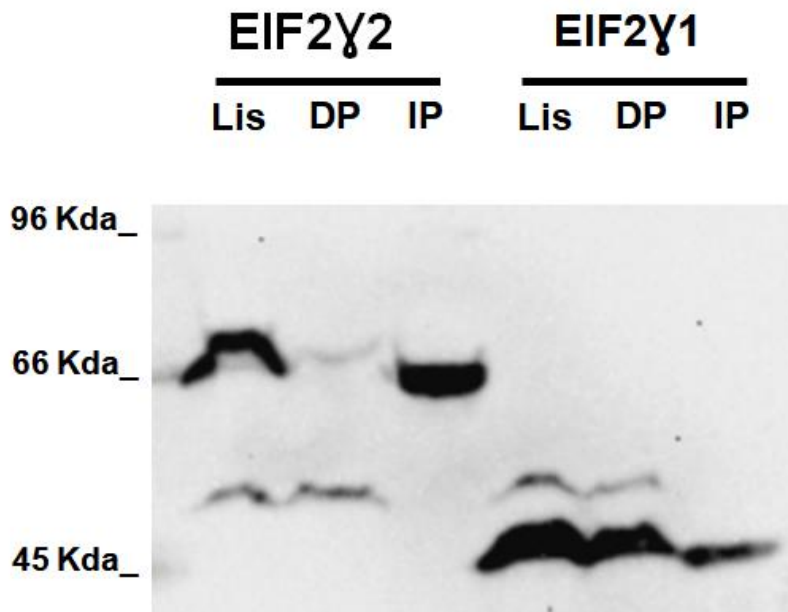
Fonte: Autoria própria (2021)

Nota: Arvore filogenética dos homólogos EIF2 γ 1 e EIF2 γ 2 com identificação da família Trypanosomatidae, classe Kinetoplastida e por coloração da fonte outros grupos de importância evolutiva.

6.1.2 Identificação de parceiros funcionais dos EIF2 γ 1 e EIF2 γ 2

A partir de células expressando o EIF2 γ 1-HA e EIF2 γ 2-HA foi possível imunoprecipitar as proteínas recombinantes EIF2 γ 1-HA e EIF2 γ 2-HA usando a resina de microesferas magnéticas anti-HA. Amostras das diferentes etapas do experimento, incluindo o lisado celular, o imunoprecipitado e o lisado depletado (lisado após incubação com a resina) foram submetidos a ensaios de *western blot* para observar a eficiência da imunoprecipitação (Figura 20).

Figura 20 - Western blot de Imunoprecipitação de fatores EIF2 γ 1 e EIF2 γ 2.



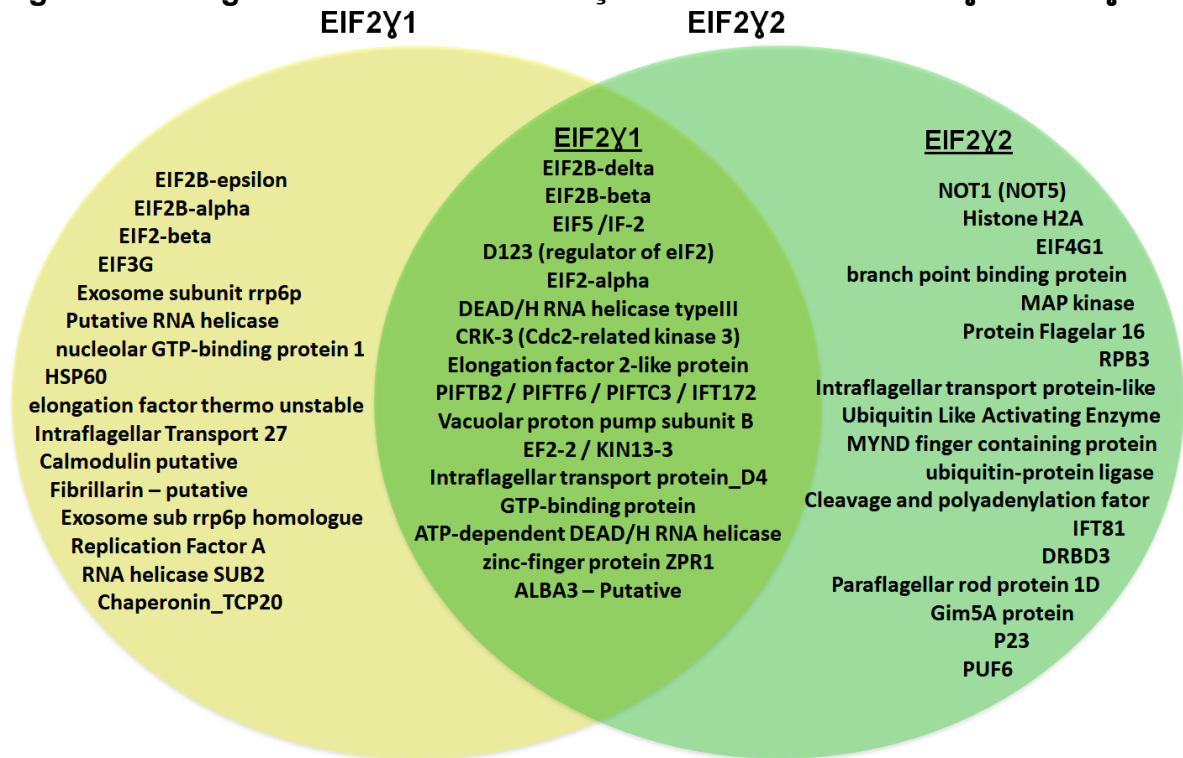
Fonte: Autoria própria (2021)

Nota: *Western blot* dos fatores EIF2 γ 1 e EIF2 γ 2 utilizando anticorpo monoclonal comercial Anti-HA na proporção de 1:3000 e anticorpo secundário na mesma proporção. Evidenciando as etapas da imunoprecipitação, Lis= Lisado celular, DP= Lisado depletado e IP= Imunoprecipitação.

Através das análises de espectrometria de massas foi possível identificar proteínas que interagem de forma direta ou indireta com os fatores EIF2 γ 1 e EIF2 γ 2 (Figura 21). Assim, os resultados de imunoprecipitação dos EIF2 γ -1 e EIF2 γ -2 expressos ectopicamente em células promastigotas, seguido de identificação proteica dos parceiros co-precipitados por espectrometria de massa, demonstraram que ambas as proteínas são capazes de co-precipitar com um dos principais parceiros do complexo eIF2, o fator eIF5, de forma similar ao descrito para o eIF2 eucariótico. Ambas as proteínas co-precipitam com outras subunidades do eIF2 (EIF2 α) e também do complexo eIF2B (EIF2B β e EIF2B δ). Entretanto, apenas o EIF2 γ -1 co-precipita com a terceira subunidade do EIF2 (EIF2 β) e com as subunidades EIF2B α e EIF2B ϵ do EIF2B. Além destas a proteína EIF2B γ é observada apenas com o EIF2 γ -1, o que é uma informação interessantes, mas que não consta na análise devido ao baixo número de peptídeos encontrado na espectrometria de massas. Apenas o EIF2 γ -1 co-precipitou com uma única subunidade do complexo eIF3, o EIF3G, além de proteínas relacionadas a exosomas e proteínas helicases. A proteína EIF2 γ 2 co-precipitou ainda com Histonas, proteínas do complexo eIF4F, o EIF4G1, MAP kinases e proteínas de

ubiquitinação e degradação celular além de algumas proteínas de ligação ao RNA. Ambas as proteínas co-precipitaram com a kinase CRK3 e o regulador de eIF2 denominado de D123.

Figura 21 - Diagrama de Venn de interação com os fatores EIF2Y1 e EIF2Y2.



Fonte: Autoria própria (2021)

Nota: Diagrama de Venn evidenciando os grupos de proteínas que interagem com os fatores EIF2Y1 e EIF2Y2, dispostos de forma decrescente as proteínas de maior razão de interação com os fatores de estudo.

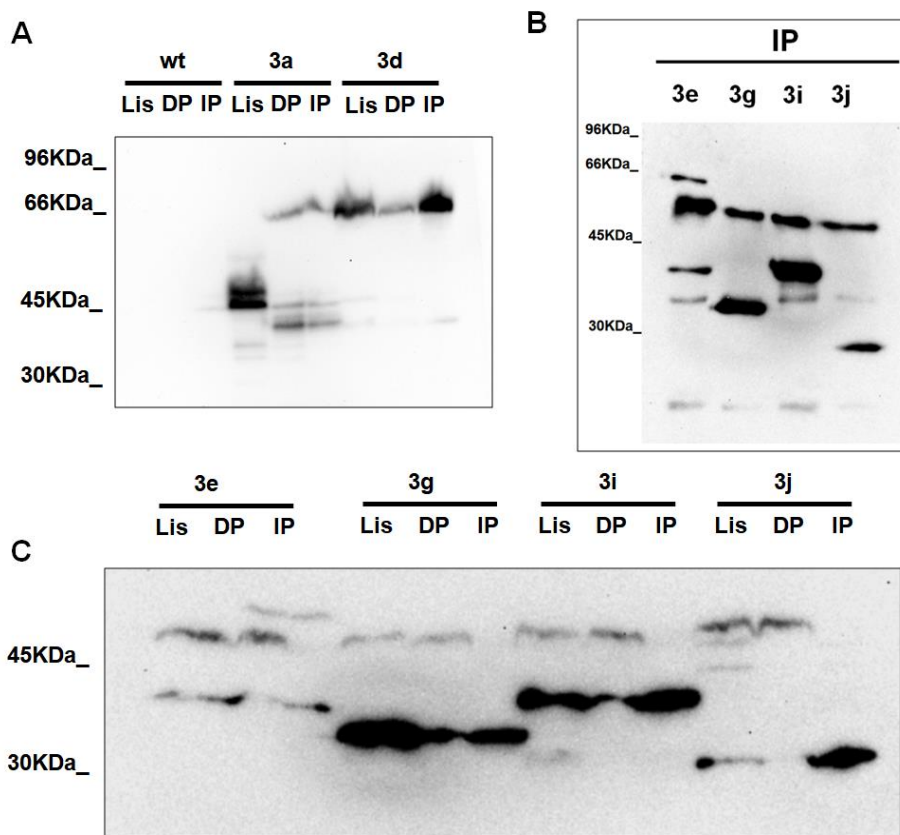
6.2 Análises de interação com parceiros funcionais do complexo eIF3

6.2.1 Expressão heteróloga de subunidades do eIF3 de *Leishmania infantum*

Assim como foi realizado com a subunidade eIF2 γ do complexo eIF2, foi realizado a expressão celular em *Leishmania infantum* de subunidades selecionadas do complexo eIF3, geradas a partir de construções plasmidiais produzidas em trabalhos prévios. No caso, seis subunidades foram expressas no parasita de forma ectópica: EIF3A, EIF3D, EIF3E, EIF3G, EIF3I e EIF3J. A presença do epítipo HA na extremidade C-terminal de cada uma destas proteínas recombinantes nos permitiu

realizar sua imunoprecipitação. Amostras das diferentes etapas do experimento, incluindo o lisado celular, o imunoprecipitado e o lisado depletado foram submetidos a ensaios de *western blot* para observar a eficiência da imunoprecipitação (Figura 22). É possível observar que a proteína EIF3A durante o processo de lise e imunoprecipitação é clivada por fatores como a lise mecânica ou ação de enzimas proteolíticas após a lise, mas que apesar da clivagem, a região C-terminal, com o peptídeo HA é mantida e pode nos trazer informações valiosas sobre parceiros que possam interagir com esta região.

Figura 22 - Western blot de Imunoprecipitação de fatores EIF3.



Fonte: Autoria própria (2021)

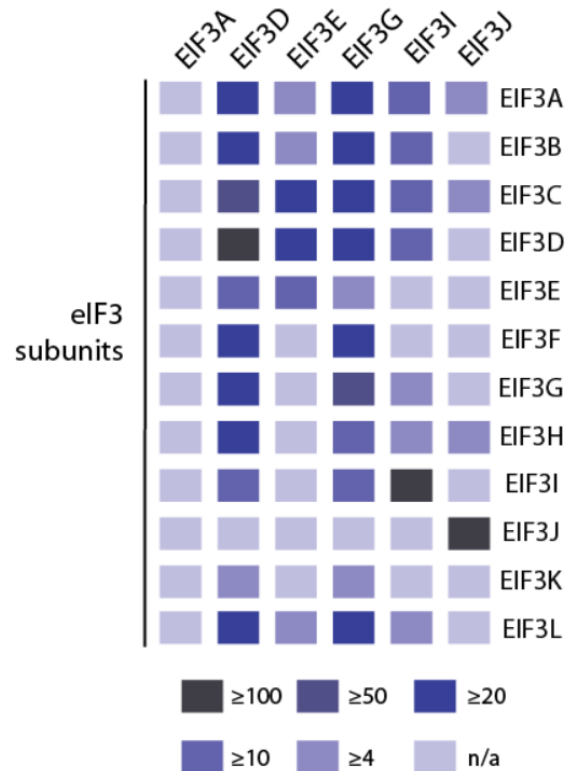
Nota: *Western blot* das subunidades de eIF3 selecionadas para este estudo: EIF3A, EIF3D em A, EIF3E, EIF3G, EIF3I e EIF3J em B e C utilizando anticorpo monoclonal comercial Anti-HA na proporção de 1:3000 e anticorpo secundário na mesma proporção. Evidenciando as etapas da imunoprecipitação, Lis= Lisado celular, DP= Lisado depletado e IP= Imunoprecipitação.

6.2.2 Análise de espectrometria de massas de proteínas do complexo eIF3

Por meio das análises dos dados da espectrometria de massas foi possível observar quais subunidades do complexo eIF3 estava interagindo entre si e com grupos distintos de proteínas, como proteínas de ligação ao RNA e as proteínas que

participam de outros complexos de iniciação da tradução. Em primeiro lugar, analisando os dados das proteínas que co-precipitaram com cada um dos fatores, podemos analisar as interações das subunidades que co-precipitam com elas mesmas (Figura 23). Com relação a imunoprecipitação do EIF3A podemos observar uma co-precipitação baixa se comparada com as das demais subunidades do complexo eIF3, mas como na preparação dos lisados e ensaio de imunoprecipitação, foi observado, que a proteína EIF3A completa não estava mais presente, dando lugar a truncagens da proteína que não possuem sua região N-terminal, estas truncagens podem explicar a falta de interação com as outras subunidades do eIF3. Já a subunidade EIF3E co-precipita principalmente com os fatores EIF3D e EIF3C, e em menor proporção com as subunidades EIF3A, EIF3B e EIF3L. A subunidade EIF3I co-precipita com um maior número de proteínas como as EIF3A, EIF3B, EIF3C e EIF3D, com maior intensidade, e ainda as EIF3G, EIF3H e EIF3L, com menor intensidade. As subunidades EIF3D e EIF3G são as que co-precipitam com o maior número de outras subunidades, mas ainda assim não co-precipitam com a EIF3J. Por outro lado, a própria EIF3J co-precipita com menos intensidade com várias subunidades, mais especificamente interage com as EIF3A, EIF3C e EIF3H, vindo muito fracamente com as demais subunidades de eIF3.

Figura 23 - Mapa de co-precipitação de subunidades do complexo eIF3 com outras subunidades desse mesmo complexo.



Fonte: Autoria própria (2021)

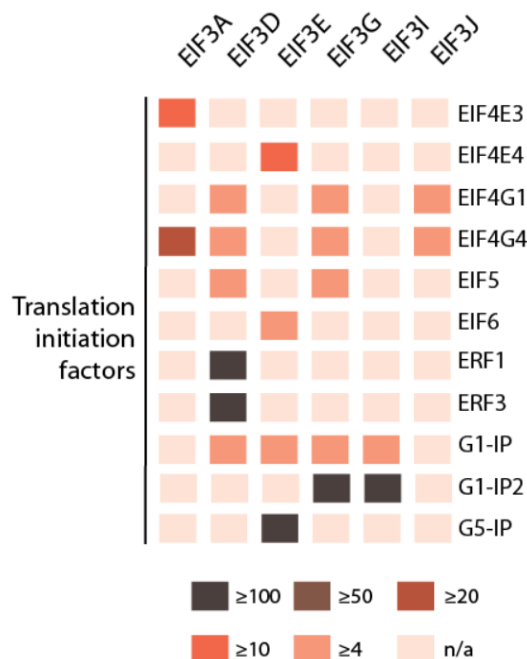
Nota: Mapa resumindo os resultados de co-precipitação das proteínas EIF3A, EIF3D, EIF3E, EIF3G, EIF3I e EIF3J, contra outras subunidades do complexo eIF3 de *L. infantum*. A intensidade da coloração aumenta na medida em que se aumenta o número de vezes que a proteína alvo é enriquecida se comparada com o controle do ensaio.

A partir dos dados de imunoprecipitação, partiu-se para avaliar como as diferentes subunidades de eIF3 co-precipitavam com outros fatores envolvidos na iniciação da tradução (Figura 24). É interessante observar que embora a proteína imunoprecipitada EIF3A tenha tido um rendimento baixo ou quase nulo na co-precipitação com as outras subunidades de eIF3, ainda assim co-precipita bem com os EIF4E3 e EIF4G4, que juntos formam um dos principais complexos do tipo eIF4F dos tripanosomatídeos. Este resultado pode indicar uma interação mais direta da região C-terminal do EIF3A e o complexo EIF4E3/EIF4G4. Já a proteína EIF3E parece ter mais afinidades com proteínas dos complexos EIF4E4/EIF4G3 e EIF4E6/EIF4G5, sendo observado uma alta interação com a proteína G5-IP, parceira conhecida do EIF4E6. A subunidade EIF3G co-precipita bem com o EIF4G4, embora não com seu parceiro EIF4E3, e com proteínas associadas ao complexo EIF4E5/EIF4G1, inclusive com suas parceiras G1-IP e G1-IP2. Por outro lado, a EIF3I co-precipita com as G1-IP e G1-IP2, embora não seja observado

associação com o fator EIF4G1 ou com o EIF4E5. No caso da subunidade EIF3D, esta co-precipita com várias proteínas em comum com o EIF3G, a exceção da ausente G1-IP2, e co-precipita fortemente ainda e de forma específica com dois fatores de terminação da tradução, os ERF1 e ERF3. Por fim a EIF3J co-precipita apenas com os EIF4G1 e EIF4G4.

Analisando de forma isolada os EIF3E e EIF3G, ambos co-precipitam com o fator EIF4E5 e com proteínas parceiras deste, como o EIF4G1, EIF4G1-IP e EIF4G1-IP2. Estes resultados reforçam a associação do complexo do tipo eIF4F baseado no EIF4E5 com o complexo eIF3 de *Leishmania*, indicando um papel direto na iniciação da tradução, motivo pelo qual estes resultados foram incluídos no artigo publicado no qual se descreve o EIF4E5 de *Leishmania* (DE LIMA; DE LIMA CAVALCANTI; DE BRITO; DE ASSIS *et al.*, 2021).

Figura 24 - Mapa de co-precipitação de subunidades do complexo eIF3 com homólogos selecionados fatores de iniciação da tradução.

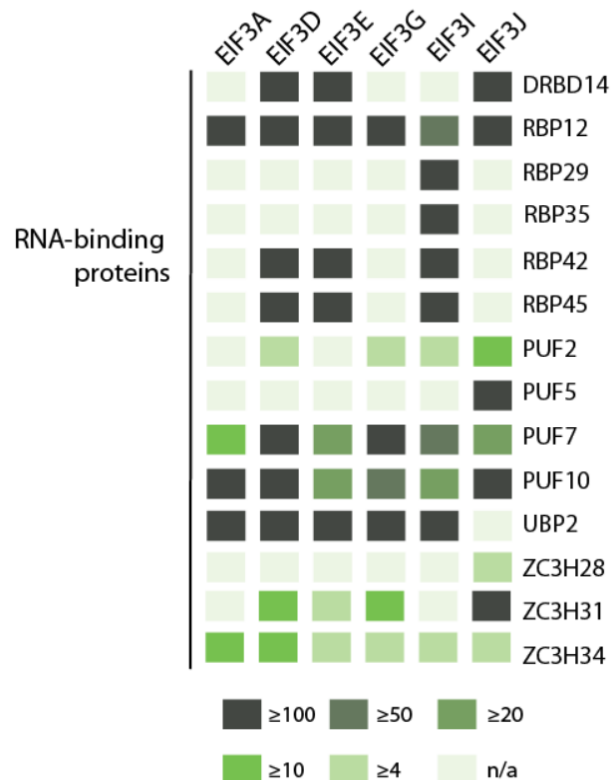


Fonte: Autoria própria (2021)

Nota: Mapa de co-precipitação das proteínas EIF3A, EIF3D, EIF3E, EIF3G, EIF3I e EIF3J, com fatores de iniciação da tradução selecionados de *L. infantum*.

Sobre a interação das subunidades do complexo eIF3 com proteínas de ligação ao RNA, vários pontos precisam ser considerados (Figura 25). É interessante observar que proteínas específicas co-precipitam diretamente com quase todas as subunidades que foram alvo dos ensaios de imunoprecipitação, como as RBP12, PUF7, PUF10 e ZC3H34. Outras proteínas só precipitaram com o EIF3J (PUF5 e ZC3H28) ou especificamente deixaram de co-precipitar só com esta subunidade, como a UBP2. É observado que algumas proteínas interagem fortemente e de forma exclusiva com o EIF3I, as RBP29 e RBP35, enquanto as RBP42 e RBP45 também co-precipitam com as subunidades EIF3D e EIF3E. Não foram observados enriquecimento específicos para proteínas de ligação ao RNA mais abundantes como PABPs e ALBAs.

Figura 25 - Mapa de co-precipitação de subunidades do complexo eIF3 com proteínas de ligação ao RNA.

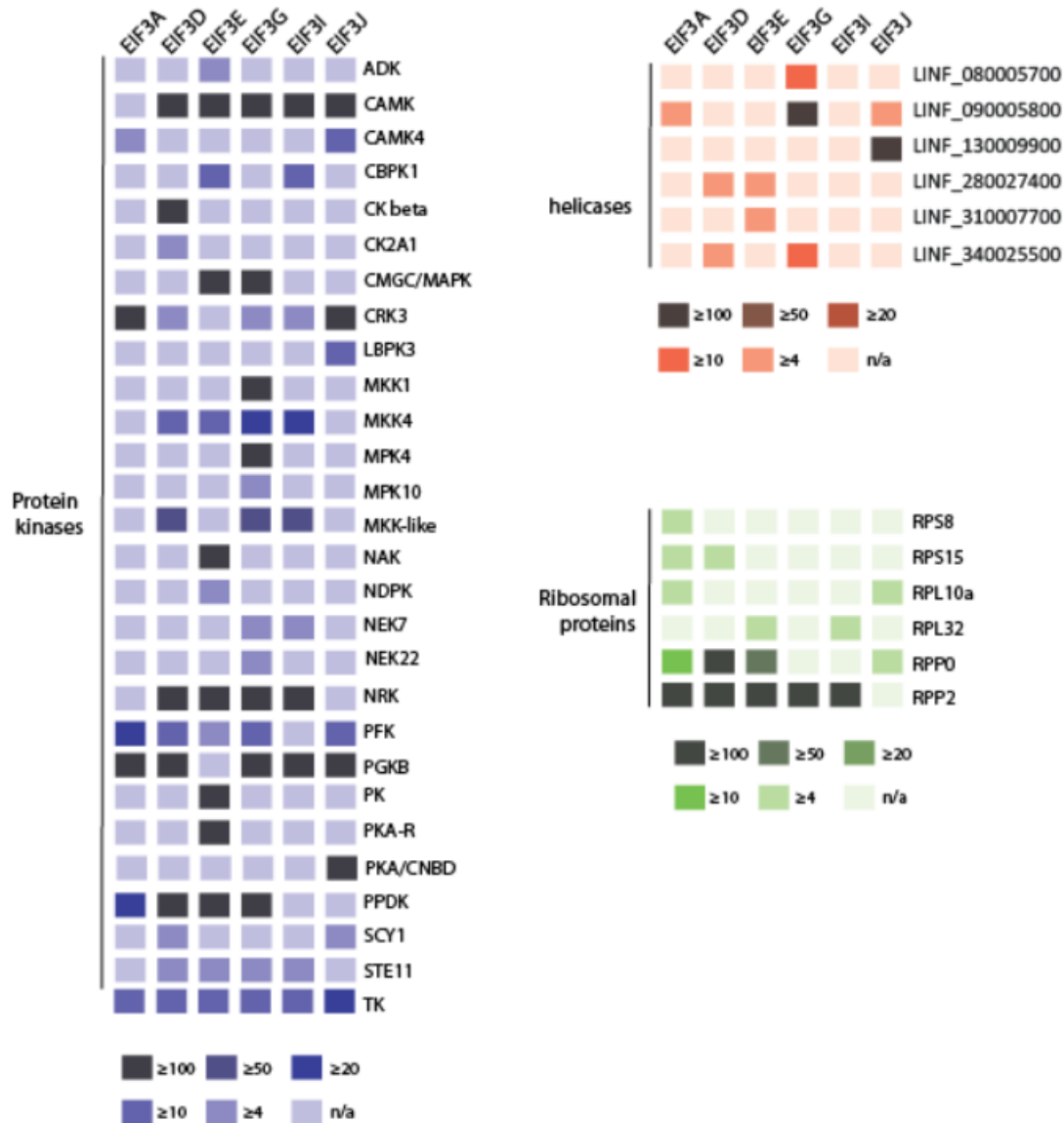


Fonte: Autoria própria (2021)

Nota: Mapa de co-precipitação das proteínas EIF3A, EIF3D, EIF3E, EIF3G, EIF3I e EIF3J, contra proteínas de ligação ao RNA de *L. infantum*.

Sobre outras proteínas que co-precipitam com as subunidades do complexo eIF3 é interessante observar que entre as kinases é observado um número maior destas co-precipitando com a EIF3E e EIF3G, sendo que ambos apresentam várias kinases com as quais co-precipitam especificamente. Algumas kinases co-precipitam com a maioria das subunidades de eIF3 avaliadas, como a CAMK, NRK e PGKB. Chama a atenção o fato de que os EIF3J e EIF3A, parceiras diretas, têm em comum a forte co-precipitação com a CRK3 que também co-precipita com os homólogos de eIF2 γ . Uma outra classe de proteínas avaliadas foi a das helicase, onde uma associação clara foi observada entre as subunidades EIF3E, EIF3G e o EIF3J com helicase específicas. Em relação às proteínas ribossomais foi observado uma alta associação de todo o complexo com a proteína RPP2, menos com o fator EIF3J, mas de forma geral, a presença de proteínas ribossomais reforça o papel do complexo eIF3 no início da tradução (Figura 26).

Figura 26 - Mapa de co-precipitação de subunidades do complexo eIF3 com proteínas kinases, RNA helicases e proteínas ribossomais.



Fonte: Autoria própria (2021)

Nota: Mapa de co-precipitação das proteínas EIF3A, EIF3D, EIF3E, EIF3G, EIF3I e EIF3J, contra proteínas kinases, RNA helicases e proteínas ribossomais de *L. infantum*.

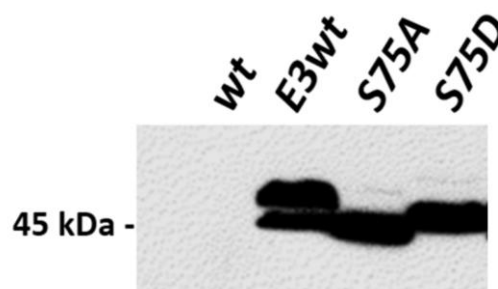
6.3 Análises de interação com parceiros funcionais do fator EIF4E3

6.3.1 Expressão heteróloga do EIF4E3 e seus mutantos

Também a partir de construções obtidas em trabalhos anteriores, foi realizado a transfecção de construção com o gene do fator EIF4E3 fusionado ao epítipo HA a fim de produzir amostras de imunoprecipitação visando a caracterização e identificação de parceiros funcionais deste fator. Neste caso o modelo usado foi a

Leishmania amazonensis tendo em vista o perfil previamente descrito para esta proteína de expressão de isoforma de maior peso molecular em fase estacionária de curvas de crescimento (PEREIRA; MALVEZZI; NASCIMENTO; LIMA *et al.*, 2013). Ao se iniciar este trabalho, os genes codificando o EIF4E3 nativo e dois mutantes importantes associados a expressão desta isoformas já se encontravam disponíveis em vetor plasmidial para a expressão em *Leishmania* (RÔMULO; MORAES, 2015). A transfecção ocorreu em uma linhagem clonal de *L. amazonensis* onde já havia sido deletada umas das cópias endógenas do gene *EIF4E3*. A expressão foi confirmada com a visualização de bandas mostrando um tamanho predito de ~46 kDa, porém duas ou mais bandas foram geralmente detectadas, sugerindo mudanças pós-traducionais, como fosforilação. É claro observar que no mutante em que a serina 75 é substituída por uma alanina ocorre a supressão da banda superior (Figura 27, S75A) indicativo da importância deste resíduo como sítio de fosforilação. Como contraprova, a serina 75 também foi substituída por um aspartato, para simular a carga de um resíduo fosforilado. Observou-se então um aumento de tamanho gerando uma proteína de peso intermediário (S75D). Estes resultados evidenciam o papel da serina 75 no fator EIF4E3 de *L. amazonensis*, visto que este resíduo não se encontra em espécies mais próximas de *Leishmania*, como a *L. infantum*.

Figura 27 - Ensaio de *western-blot* avaliando a expressão em *Leishmania amazonensis* da proteína EIF4E3 fusionada ao epítipo HA e dois mutantes.



Fonte: Autoria própria (2021)

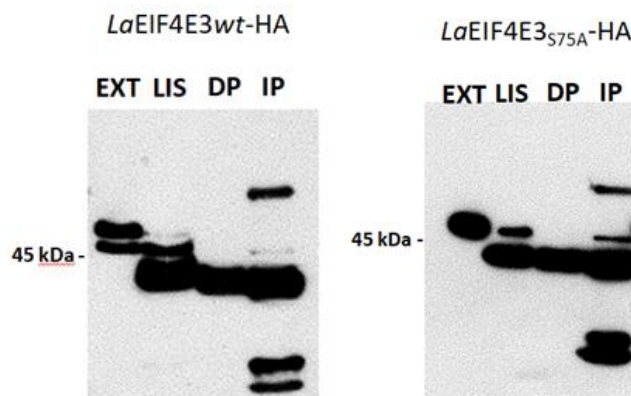
Nota: Experimento realizado com anticorpo monoclonal comercial anti-HA usando formas promastigotas de *L. amazonensis* em crescimento exponencial. Células de *L. amazonensis* não transfectadas (*wt*) foram utilizadas como controle negativo. EIF4E3wt representa a proteína nativa, com os seus mutantes S75A e S75D indicados. Foram aplicados em cada poço a quantidade de extrato total da célula de equivalente a 1×10^7 células.

6.3.2 Lisados celulares e imunoprecipitação do EIF4E3 nativo e mutante S75A.

A partir das células expressando o EIF4E3 selvagem e o mutante S75A, foram produzidos lisados celulares e realizados ensaios de imunoprecipitação, utilizando a resina ligada ao anticorpo anti-HA. Amostras das diferentes etapas do experimento, incluindo o lisado celular, o imunoprecipitado e o lisado depletado (lisado após incubação com a resina) foram submetidos a ensaios de *western blot* para observar a eficiência da imunoprecipitação (Figura 28).

Com as imunoprecipitações foi possível observar que a proteína após a lise celular sofre uma clivagem específica, e em lugar da banda com o tamanho previsto de 47kDa observa-se um segundo fragmento mais abundante de 40kDa. Esta clivagem é observada tanto no lisado (LIS), como no lisado depletado (DP) e na imunoprecipitação (IP), mas, não no extrato total (EXT). Com isto conclui-se que existe algo na técnica que leva a proteína a ser clivada rapidamente logo após a lise. Para evitar este tipo de clivagem foi utilizado uma concentração maior de inibidor de protease, que não alterou este padrão. As amostras de imunoprecipitação foram então enviadas para espectrometria de massas, para análise das proteínas que co-precipitam juntamente com o EIF4E3 e seu mutante.

Figura 28 - Ensaio de *western-blot* para avaliação de imunoprecipitação do fator EIF4E3 de *Leishmania amazonensis* expressando o EIF4E3 e seu mutante S75A, ambos fusionados ao epítipo HA.



Fonte: Autoria própria (2021)

Nota: Experimento utilizando o anticorpo monoclonal comercial anti-HA da marca sigma® diluído 1:3000. EXT- extrato proteico total, LIS- lisado celular, DP- lisado depletado e IP- imunoprecipitado. Pode ser observado nas IPs uma banda superior a todas as outras e uma inferior, essas bandas corresponde as cadeias das imunoglobulinas utilizadas para este ensaio.

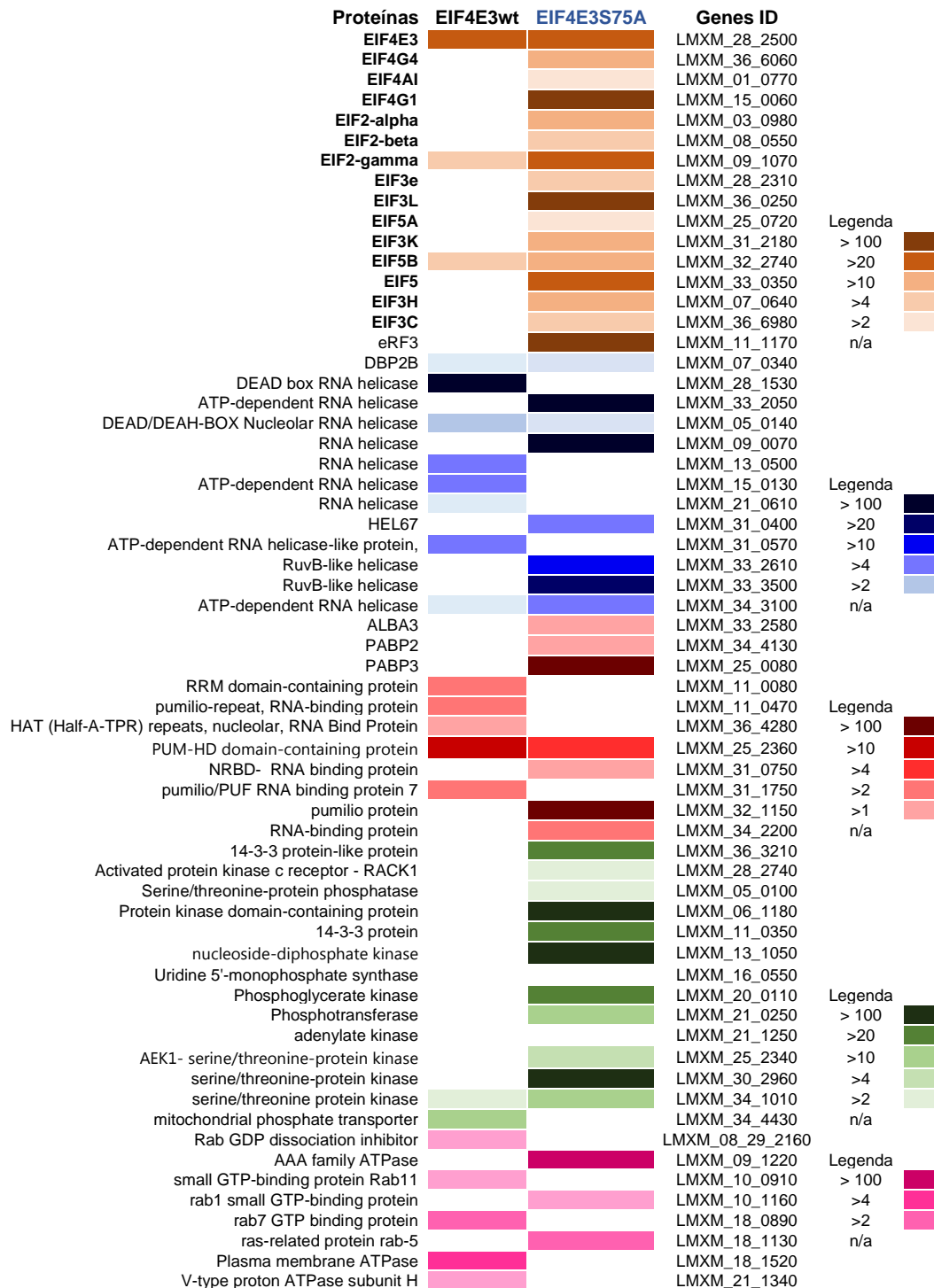
6.3.3 Identificação de parceiros do EIF4E3 e avaliação do impacto da mutação S75A sobre sua capacidade de interação a parceiros selecionados

Através da espectrometria de massas foi possível analisar de forma independente a proteína EIF4E3 nativa e seu gene mutado com a mutação S75A, com os resultados obtidos resumidos nas Figura 29 a Figura 33. Com estas análises encontramos algo peculiar, onde a proteína selvagem a princípio formaria um complexo do tipo eIF4F juntamente com o EIF4G4 além de outras proteínas parceiras, o que não é observado. Este padrão, contudo, é observado na proteína com a mutação na serina 75 que co-precipitou tanto com o EIF4G4 como com o EIF4G1. Além destes, a proteína mutante co-precipitou com todo um conjunto de proteínas que estariam associadas a uma função na iniciação da tradução, como as PABP3 (fortemente) e PABP2 (fracamente), EIF4AI, proteínas do complexo eIF2 (EIF2 α , EIF2 β e EIF2 γ), EIF5, EIF5A, EIF5B e proteínas do complexo EIF3 (EIF3E, EIF3L e EIF3K). Além destes, várias proteínas ribossomais também foram observadas com a proteína mutante, indicando uma associação com a tradução.

Na co-precipitação do fator EIF4E3 selvagem foi interessante observar em alguns grupos de proteínas a interação de forma distinta do fator mutado. Por exemplo, algumas helicases que foram encontradas co-precipitando com o fator EIF4E3 selvagem não foram sequer detectadas na proteína mutante, e vice-versa. É importante notar algumas proteínas que interagem exclusivamente com o EIF4E3 selvagem, como proteínas com domínio PUF (pumilio) que estão relacionadas a inibição da tradução de RNAs com os quais estão ligadas, proteína ribossomal L7, algumas proteínas nucleares e histonas, que nos dão indícios de que este fator enquanto fosforilado está de alguma forma inativo para um papel de tradução.

As análises de espectrometria de massa mostraram ainda que o fator EIF4E3 selvagem interagiu com um número maior e com maior intensidade com proteínas do proteasoma ou proteínas relacionadas a estas como proteínas regulatórias do proteasoma ou ubiquitinas, nos indicando uma possível maior facilidade de via de degradação, embora não sendo observado por *western blot* um grau de degradação aparente.

Figura 29 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com fatores de iniciação da tradução, helicases, proteínas de ligação ao RNA, quinases e NTPases.



Fonte: Autoria própria (2021)

Nota: Gráfico ilustrativo em estilo *heatmap*, evidenciando as intensidades em número de vezes que a proteína está frequente se comparado ao controle negativo.

Figura 30 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com proteínas ribossomais, ubiquitinas, proteínas do proteasoma e nucleares.

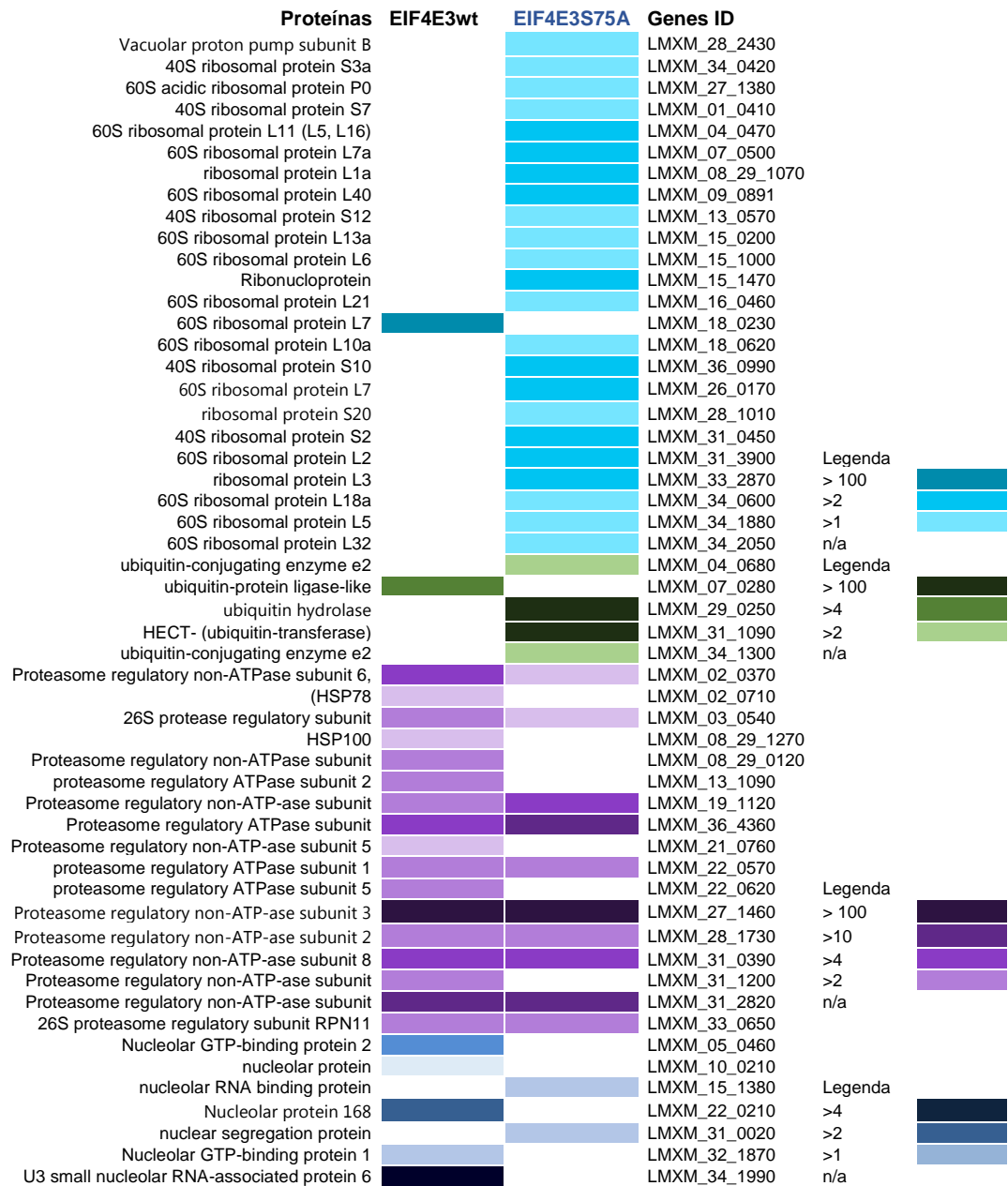


Figura 31 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com proteínas transportadoras, mitocondriais, de atividade metabólica e hipotéticas.

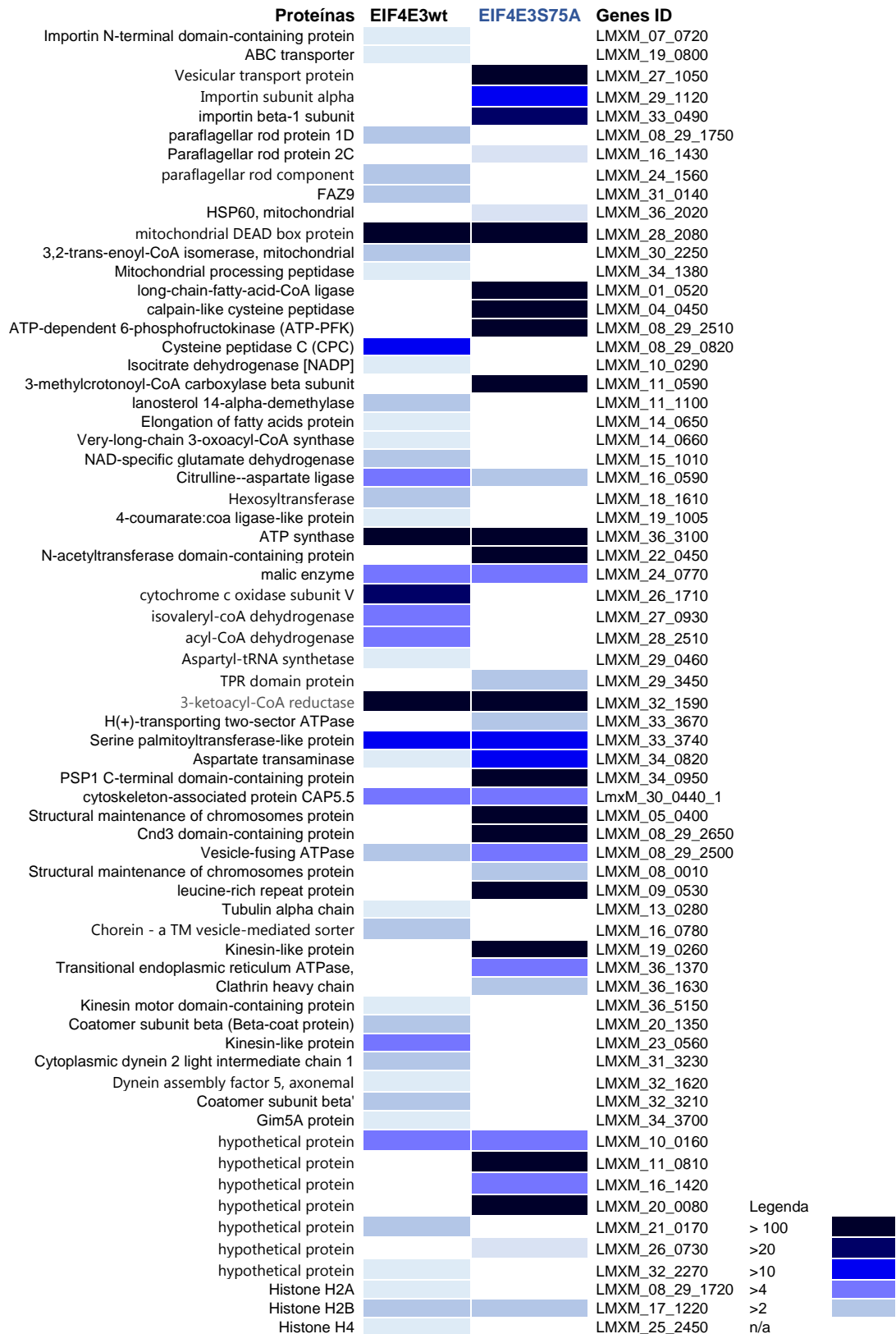


Figura 32 - Análise de interação do fator EIF4E3wt e EIF4E3S75A com proteínas diversas.

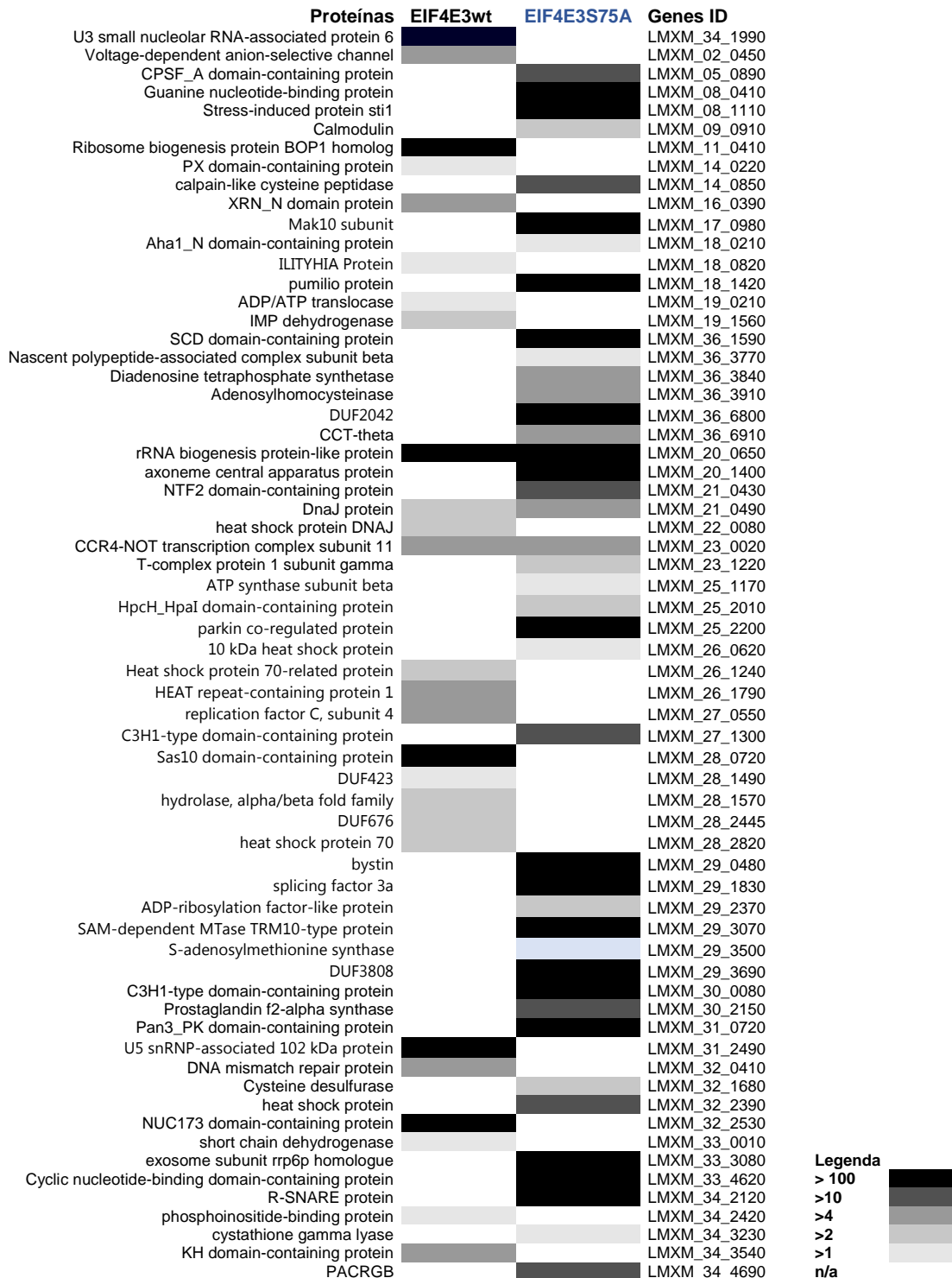
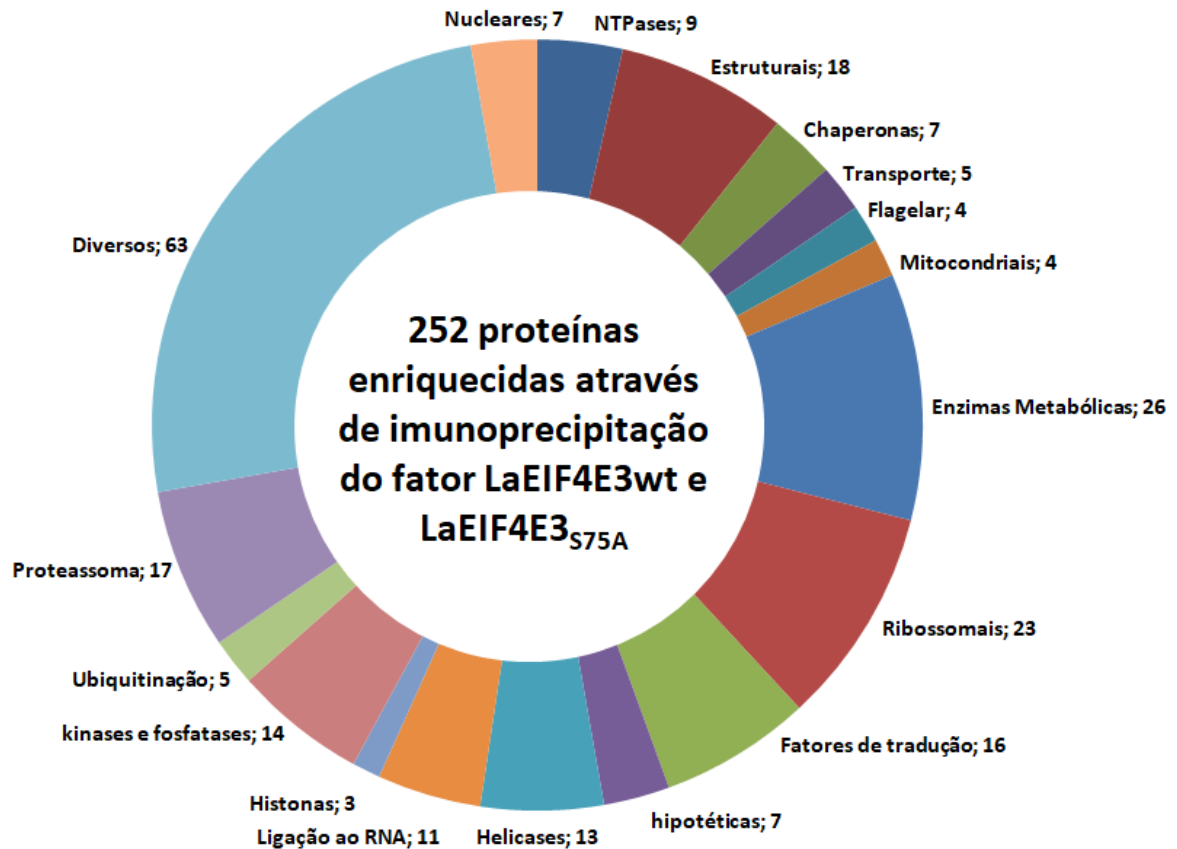


Figura 33 - Distribuição das proteínas que co-precipitaram com o fator EIF4E3.



Fonte: Autoria própria (2021)

Nota: Gráfico em pizza evidenciando os grupos de proteínas e a sua quantificação, todas enriquecidas pela imunoprecipitação do fator EIF4E3 em *L. amazonensis*.

Todas as proteínas citadas nas Figura 29 à Figura 32 estavam relacionadas à co-precipitação de alguns fatores que consigo trazem outros, mas é evidente o papel que a mutação na serina 75 e seu papel na ativação/inibição por fosforilação atua de alguma forma inibindo a interação deste fator em *L. amazonensis*. Como contraprova, para garantir que não houve qualquer troca no processamento das amostras, foi realizado o sequenciamento dos peptídeos destas proteínas a fim de identificar as mutações em suas respectivas linhagens (Tabela 2).

Tabela 2 - Sequências de peptídeos da proteína alvo por Espectrometria de massas do fator EIF4E3.

Protein IDs	Fasta headers	Peptide sequences
E9B072	tr E9B072 E9B072_LEIM U Eukaryotic translation initiation factor, putative OS= <i>Leishmania mexicana</i> (strain MHOM/GT/2001/U1103) OX=929439 GN=LMXM_28_2500 PE=3 SV=1	DIKPEWEHPR; GGTISIVVFDR; KDIKPEWEHPR; KGGTISIVVFDR; KGGTISIVVFDNRDR; LEYYSHHQK; LSAAAEFPVPGGPK; TTV <u>S</u> PLTERPASSLPNSAAAAGAAK; YLPAPSGLPTVYTYSWFR
LEIMU (proteína mutada)	LEIMU Mutant Eukaryotic translation initiation factor	DIKPEWEHPR; GGTISIVVFDR; KDIKPEWEHPR; KGGTISIVVFDR; KGGTISIVVFDNRDR; LEYYSHHQK; SAAAEFPVPGGPK; TATEDGKTTV <u>A</u> PLTERPASSLPNSAAAA GAAK; TTV <u>A</u> PLTERPASSLPNSAAAAGAAK; YLPAPSGLPTVYTYSWFR

Pode ser observado pela análise que foi sequenciado um peptídeo que tem a região da serina 75 na proteína selvagem e na proteína mutante teve a presença da mutação (alanina) em dois peptídeos sequenciados pela espectrometria de massas.

6.4 Caracterização estrutural do fator EIF4E5

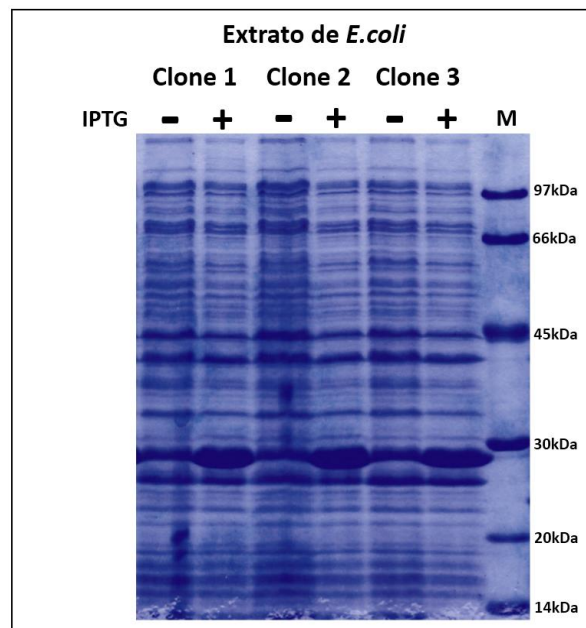
Para compreender as interações, assim como obter uma estrutura tridimensional do fator EIF4E5 de *Leishmania major*, foi realizado em conjunto com o laboratório de Proteômica e Engenharia de Proteínas do Instituto Carlos Chagas (Fiocruz), sob supervisão da pesquisadora Dra Beatriz Gomes Guimarães,

experimentos para se obter a cristalografia do fator EIF4E5 de *Leishmania*. Este foi escolhido entre os fatores utilizados neste trabalho por ter sido observado uma alta expressão em sistema de expressão bacteriano com alta solubilidade, o que facilita a sua purificação.

6.4.1 Expressão e purificação do fator EIF4E5 com cauda de Histidina

Para realização da cristalografia foi necessária se obter a proteína purificada em estudo em uma concentração de 5 a 10 mg/ml. Para isto, foi feita a indução para se expressar em *E. coli* (BL21) a proteína EIF4E5 que contém uma sequência de seis histidinas na sua extremidade C-terminal, utilizando se para isso a construção pET21d-EIF4E5 (Figura 34).

Figura 34 - Expressão do fator EIF4E5 em *Escherichia coli* após indução.

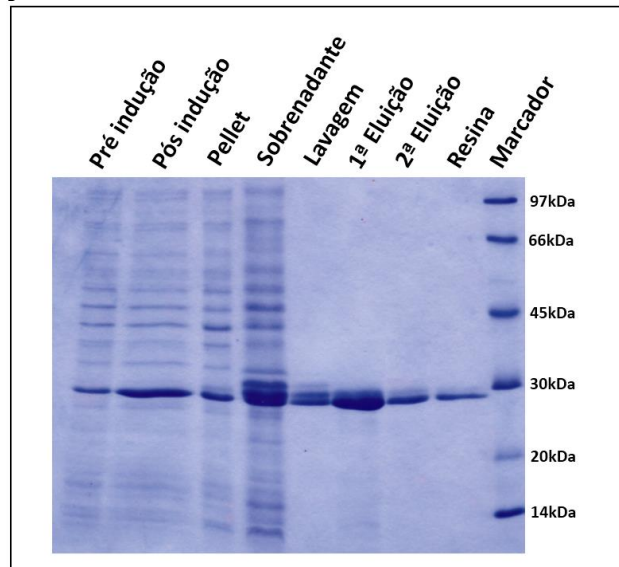


Fonte: Autoria própria (2021)

Nota: Separação por eletroforese em gel SDS-PAGE, onde foi observada a indução da proteína EIF4E5-His de valor aproximado a 28kDa. Foram avaliados três clones bacterianos.

Após confirmada a expressão do fator EIF4E5-6xHis, foi realizada a purificação utilizando cromatografia líquida, através de uma resina de Níquel, separando as proteínas por afinidade iônica (Figura 35). Com isto foi visto que a proteína não tem dificuldade de ser expressa e purificada, sendo um bom alvo para a formação de cristal e de podermos caracterizá-lo estruturalmente.

Figura 35 - Purificação do fator EIF4E5.



Fonte: Autoria própria (2021)

Nota: Separação por eletroforese em gel SDS-PAGE, onde foi observada a purificação da proteína EIF4E5-His.

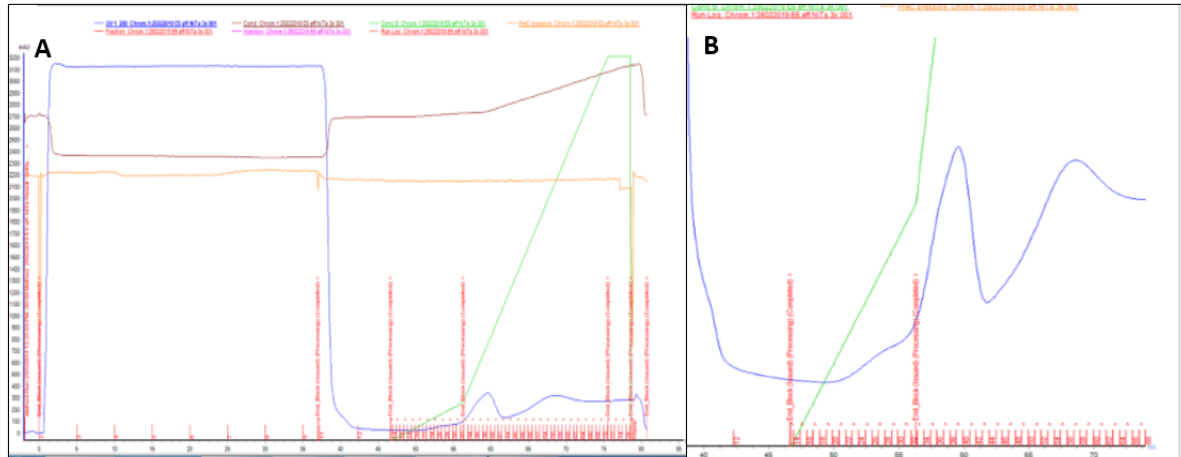
6.4.1.1 Expressão e purificação do EIF4E5 por FPLC por afinidade, troca iônica e filtração em gel

A expressão e purificação da proteína EIF4E5 para cristalografia foi realizada otimizando cada etapa do processo para uma melhor obtenção da proteína purificada. Para expressão em vetor de expressão bacteriano foi utilizado a *E. coli* (BL21 – Δ SlyD).

6.4.1.2 Purificação por afinidade a histidina

Após confirmar a expressão em pequena escala, foi realizada a indução em larga escala para purificação diretamente em coluna cromatográfica de afinidade a Histidina (HisTrap 1ml FF) (Figura 36).

Figura 36 - Cromatograma de purificação da proteína EIF4E5 por cauda de histidina.

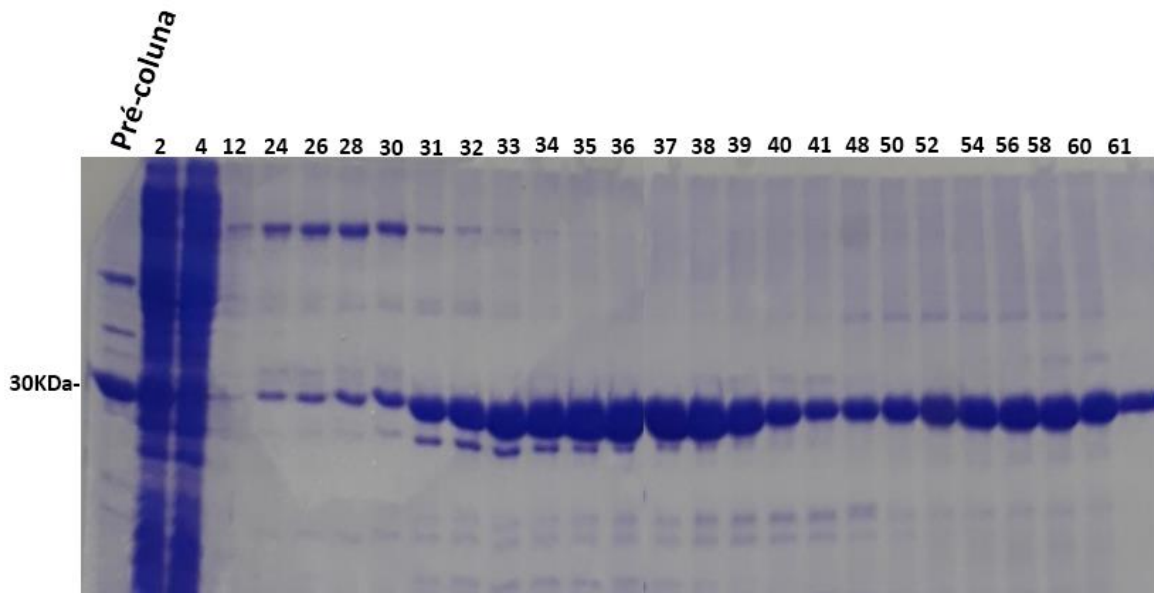


Fonte: Autoria própria (2021)

Nota: A) Cromatograma da purificação do fator EIF4E5 com cauda de 6x Histidina, B) zoom na região de eluição, onde pode ser observado dois pico específico da fração 34 à 61.

Foram coletadas várias frações do cromatograma, em específico a região com maior absorvância no comprimento de luz UV de 280, e resolvidos em gel SDS-PAGE para observar o conteúdo de cada fração, a presença da proteína alvo e de prováveis contaminantes (Figura 37).

Figura 37 - Eletroforese das frações do cromatograma de purificação por afinidade a cauda de histidina do fator EIF4E5.



Fonte: Autoria própria (2021)

Nota: Eletroforese da purificação por FPLC em SDS-PAGE, pré coluna é o extrato total bacteriano após indução, e as numerações indicam as frações da cromatografia (Figura 36).

A purificação resultou em dois picos equivalentes a proteína de interesse EIF4E5. Por terem saído em dois picos é certo de que há duas populações com

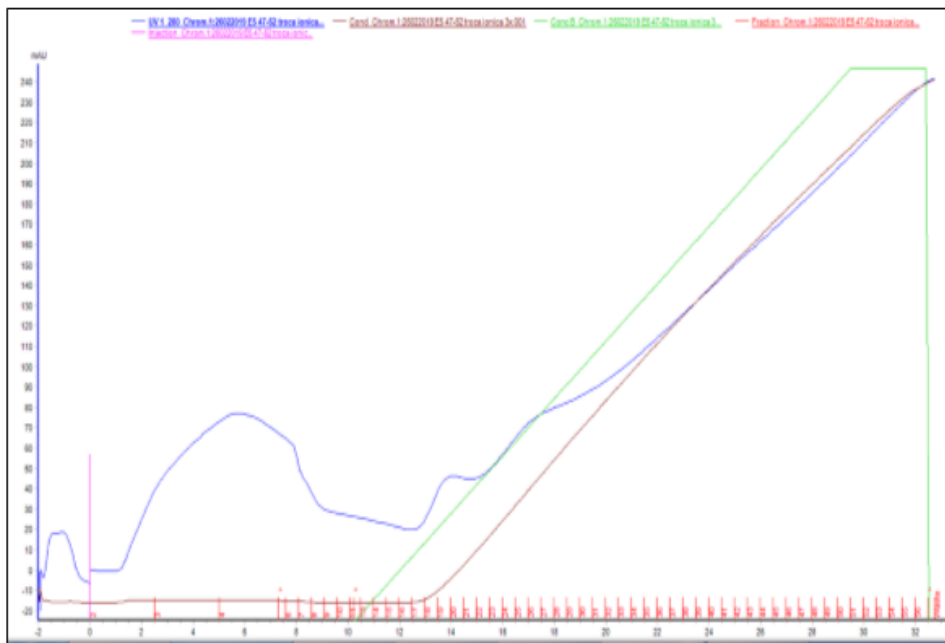
diferenças estruturais do fator EIF4E5, onde no primeiro pico é necessária uma concentração menor de imidazol para eluição, e na segunda uma concentração maior.

Foi utilizado então o primeiro pico da cromatografia de afinidade de histidina para realizar a etapa de purificação seguinte. Para isso foram misturadas todas as frações do pico 1 e o conjunto dialisado com o tampão A da cromatografia de troca iônica.

6.4.1.3 Purificação por troca iônica

Após realizar a diálise da proteína foi realizado uma purificação diferente da purificação por cauda de histidina. Nesta as proteínas purificadas são separadas pela sua carga, utilizando uma coluna de troca iônica, especificamente carga aniônica (HiTrap Q HP 1ml) (Figura 38), Nesta purificação o objetivo é separar a proteína alvo dos contaminantes.

Figura 38 - Cromatograma de purificação da proteína EIF4E5 por troca iônica.

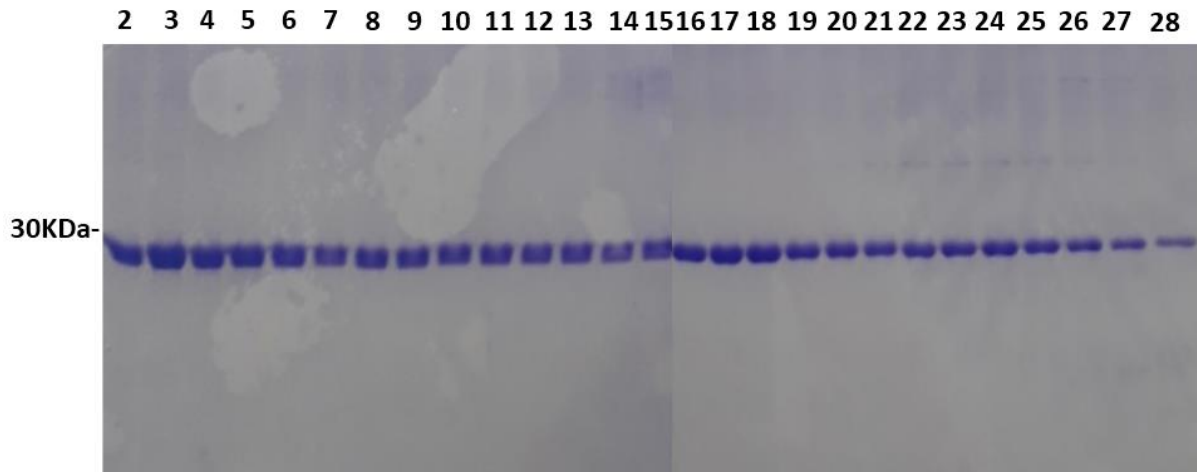


Fonte: Autoria própria (2021)

Nota: Cromatograma da purificação do fator EIF4E5 por troca iônica, onde pode ser observado vários picos, mas é possível observar uma maior absorbância de UV nas primeiras frações.

Foram coletadas várias frações do cromatograma, em específico a região com maior absorbância no comprimento de luz UV de 280, e resolvidos em gel SDS-PAGE para observar o conteúdo de cada fração, a presença da proteína alvo e de prováveis contaminantes (Figura 39).

Figura 39 - Eletroforese das frações do cromatograma de purificação por troca iônica do fator EIF4E5.



Fonte: Autoria própria (2021)

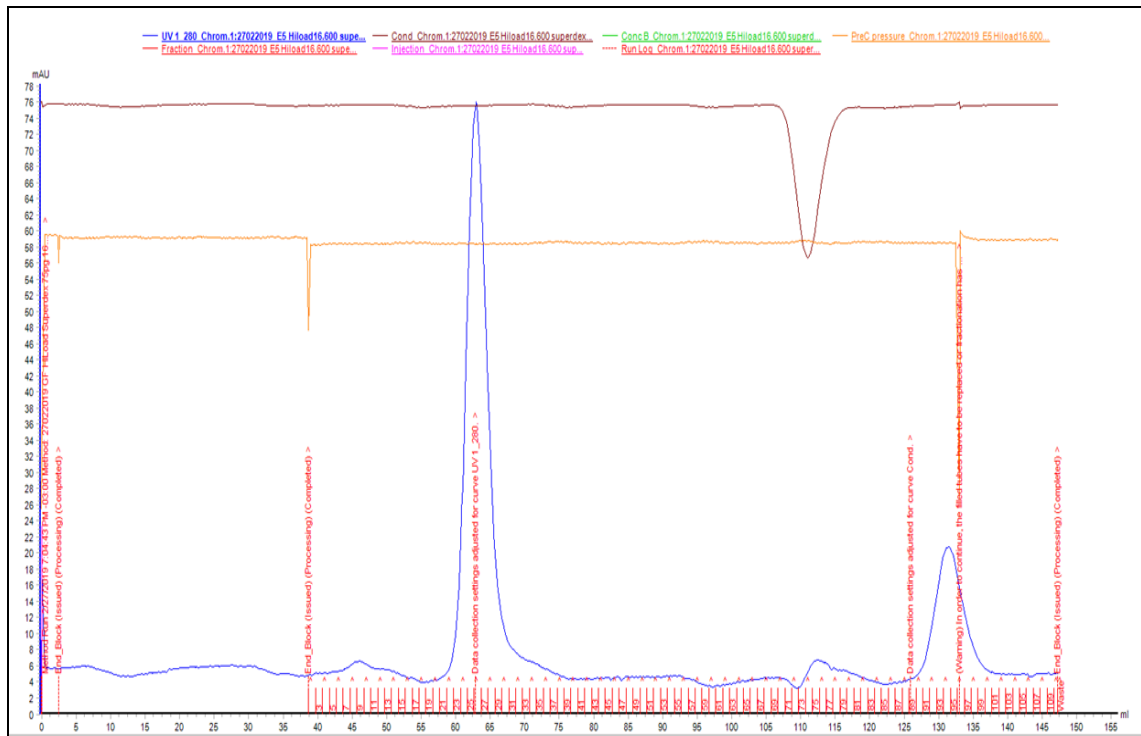
Nota: Eletroforese da purificação por FPLC em SDS-PAGE. As numerações indicam as frações da cromatografia apresentada na Figura 38.

A purificação resultou em vários picos, só não foi esperado que a proteína não fosse liberada em um pico específico, nos levando a crer que a proteína recombinante EIF4E5 não se liga a coluna de troca iônica aniônica, mas serviu para remover proteínas inespecíficas que aparecem a partir da fração 21. Foi utilizado da fração número 2 até a fração 20, estas frações foram misturadas e dialisadas em tampão de gel filtração.

6.4.1.4 Purificação por gel filtração

Após purificação por troca iônica e diálise, a preparação da amostra para purificação por exclusão de tamanho (gel filtração) necessitou que esta fosse concentrada para o volume de aplicação compatível com a coluna utilizada, no caso em um máximo de 2 ml (superdex 16/600 75pg). Nesta purificação o objetivo foi o de separar a proteína alvo de contaminantes e agregados através da separação por tamanho (massa globular) (Figura 40).

Figura 40 - Cromatograma de purificação da proteína EIF4E5 por exclusão de tamanho.

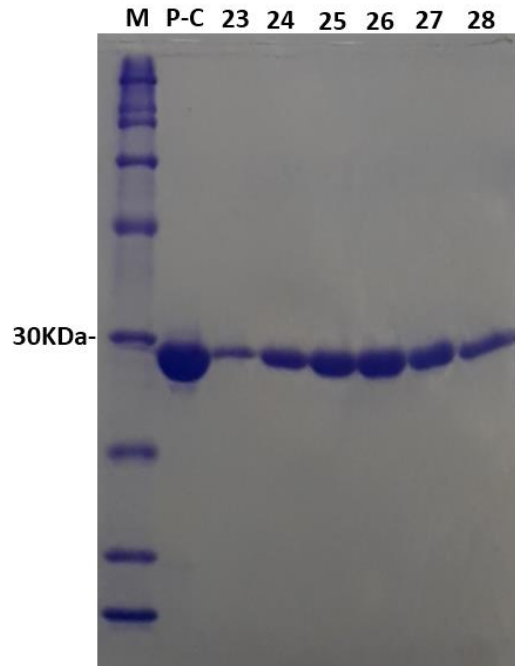


Fonte: Autoria própria (2021)

Nota: Cromatograma da purificação do fator EIF4E5 por exclusão de tamanho, onde pode ser observado um único pico nas frações iniciais, onde se espera uma proteína de tamanho aproximado entre 30 e 25KDa.

Foram coletadas as frações do cromatograma, em que se encontra o pico nas fases iniciais do cromatograma, em específico a região com maior absorvância no comprimento de luz UV de 280, e resolvidos em gel SDS-PAGE para observar o conteúdo de cada fração, a presença da proteína alvo e de prováveis contaminantes (Figura 41).

Figura 41 - Eletroforese das frações do cromatograma de purificação por exclusão de tamanho do fator EIF4E5.



Fonte: Autoria própria (2021)

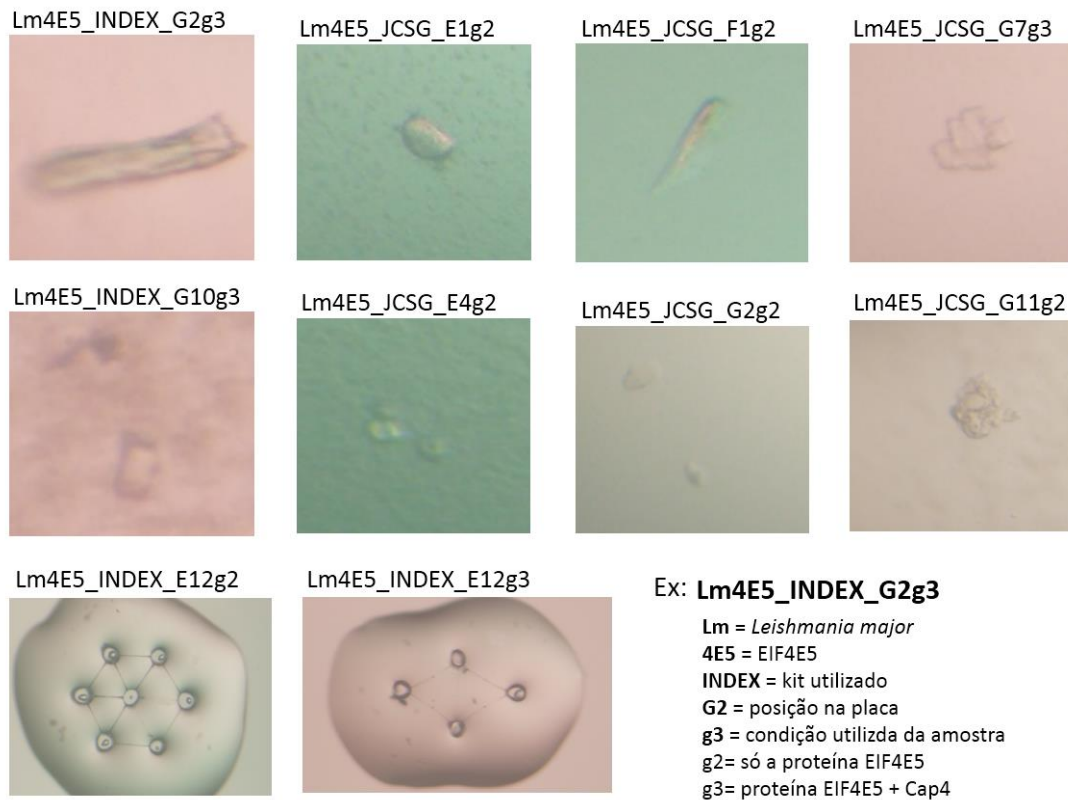
Nota: Eletroforese da purificação por FPLC em SDS-PAGE, M- marcador de peso molecular, P-C – Pré coluna e as numerações indicam as frações da cromatografia por exclusão de tamanho (Figura 40).

A purificação resultou em um único pico, cada fração desta cromatografia foi feita em 1ml para cada fração, resultando em 6ml totais de proteína altamente purificada, a proteína foi quantificada e armazenada a 4°C.

6.4.2 Cristalização do EIF4E5 recombinante

Para triagem de cristais do fator EIF4E5 foram utilizadas as frações recentemente purificadas por exclusão de tamanho. Foi necessário inicialmente a quantificação, pois caso a concentração da proteína estivesse abaixo da desejada, de 5 a 10 mg/ml, seria necessário realizar a concentração da amostra proteica. Foi utilizado soluções providas pelos sistemas comerciais (Hampton Research) PEG/Ion, INDEX e JCSG I. A condição de cristalização encontrada e reprodutível foi usando 5,2mg/mL da proteína de interesse em 0.2 M de Citrato de potássio tribásico monoidratado, 20% (p/v) de polietilenoglicol (PEG) 3350 incubado a 20°C (Figura 42).

Figura 42 - Formação dos cristais através de soluções precipitantes.

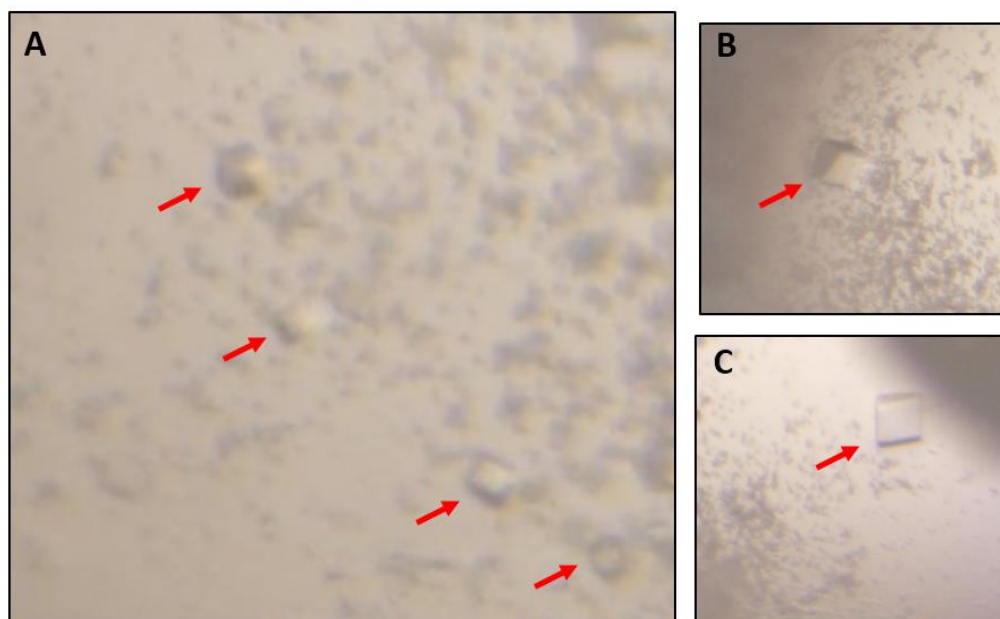


Fonte: Autoria própria (2021)

Nota: Estruturas formadas pela precipitação controlada da proteína EIF4E5 e EIF4E5+Cap4.

A reprodução dos cristais foi realizada sob método manual de pipetagem. O experimento de otimização de cristais foi realizado modificando os parâmetros de concentração de PEG 3533 (Figura 43).

Figura 43 - Refinamento de cristais da proteína recombinante EIF4E5 através de soluções precipitantes específicas.



Fonte: Autoria própria (2021)

Nota: Estruturas formadas pela precipitação controlada da proteína EIF4E5, especificados nas setas em vermelho, (A) proteína EIF4E5 com *cap4* e em (B) e (C) cristais formados apenas do EIF4E5. Todos os cristais foram obtidos sob mesma solução precipitante de 1,25M de Citrato de sódio, 0,1M de cacodilato de sódio em pH 7.0.

Análises de difração dos cristais obtidos foram realizadas em difratômetro Rigaku. A coleta de dados foi realizada utilizando síncrotron APS-U; França em colaboração com o laboratório de proteômica e engenharia de proteínas do instituto Carlos Chagas, FIOCRUZ-PR.

6.4.3 Refinamento estatístico e dados cristalográficos

Para a obtenção de uma caracterização estrutural e informações sobre os mecanismos moleculares envolvido na interação entre o fator EIF4E5 e seus parceiros, nós fomos capazes de cristalizar a partir da proteína obtida da sequência de *Leishmania major* em colaboração com o Instituto Carlos Chagas, FIOCRUZ-PR, que compartilha 98% de identidade com EIF4E5 de *Leishmania infantum*. A estrutura cristalina foi determinada por substituição molecular usando um modelo gerado pelo servidor Phyre2 como modelo de pesquisa. Apesar da obtenção de pequenos cristais da proteína ligada ao *cap*, dados de difração adequados foram obtidos apenas a partir de cristais crescidos na ausência do ligante (apo) (Tabela 3). O conjunto de resultados obtidos constitui parte fundamental do artigo publicado ligado a esta tese (DE LIMA; DE LIMA CAVALCANTI; DE BRITO; DE ASSIS *et al.*, 2021).

Tabela 3 - Dados cristalográficos e estatísticas de refinamento.

Dados estatísticos:	
Fonte	SOLEIL-PX1
Comprimento de onda (Å)	0,978565
Resolução (Å)	50–2,0 (2,12–2,00)
Grupo de Espaço	P3221
Célula unitária (Å)	a = b = 61,89 c = 173,03
Número de Observações	470.983 (61.054)
Número de reflexões únicas	26.851 (4162)
Completeness (%)	99,8 (98,6)
Redundância	17.5 (14.7)
I/σ	9.1 (0.96)
CC ½	99.8 (44.6)
Estatísticas de Refinamento:	
Rwork	0,201

Rfree	0,230
Comprimento da ligação RMSD (Å) / Ângulo (°)	0,008 / 0,96
Valor médio de B (proteína geral, Å [2])	41,9
Átomos de proteína	2807
Número de moléculas de água	204
Ramachandran Plot:	
Favorecido (%)	98,0
Discrepantes (%)	0

Fonte: Autoria própria (2021)

Nota: Os valores entre parênteses são para a camada de maior resolução.

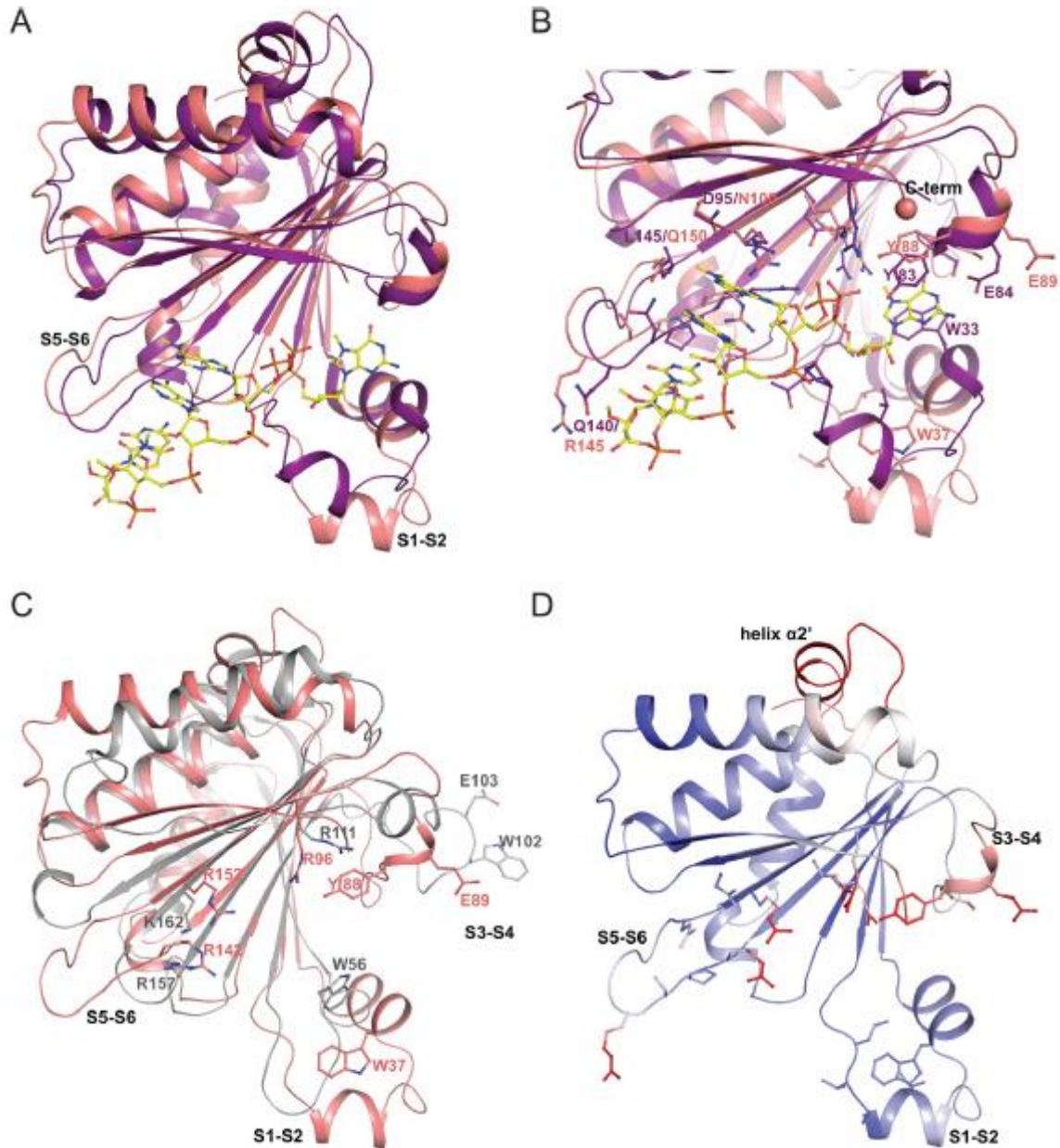
Com os dados obtidos foi realizada a superposição da estrutura do EIF4E5 de *Leishmania* com o EIF4E5 de *T. cruzi* em complexo com o *cap4* (PDB6O7Y) resultou em um desvio da média da raiz quadrada (RMSD- *Root Mean Square Deviation*) de 2,3 Å para 172 C-alfa átomos alinhados, com 43% de identidade de sequência, valores que indicam o grau de semelhança entre as estruturas do EIF4E5 de *L. major* e *T. cruzi*. A comparação estrutural entre as formas holo e apo evidenciou rearranjos significativos (Figura 43 A). O resíduo de triptofano que participa da interação clássica de empilhamento com a porção m7G na estrutura de holo TcEIF4E5 (W33) se move longe do bolso m⁷G em LmEIF4E5 (W37), como uma consequência do rearranjo de toda a região de fitas β1-β2 (S1-S2) (Figura 44B).

6.4.4 Análise do fator EIF4E5 com interface ao EIF4G

Com a estrutura do fator EIF4E5 de *L. major*, nós podemos investigar diferenças na superfície de interação ao eIF4G quando comparada com outros homólogos de eIF4E com estrutura já descrita na literatura. A sobreposição estrutural de LmEIF4E5 com eIF4E1 humano e eIF4E de levedura em complexo com peptídeos eIF4G (Figura 44) mostra que, assumindo uma interface de ligação semelhante, o resíduo W58 do LmEIF4E5 (equivalente a W53 em L_hEIF4E5) mantém a conformação altamente conservada do triptofano equivalente de humano e levedura (W73 e W75, respectivamente) e estaria em posição de interagir com a hélice canônica do eIF4G. O W50 do LmEIF4E5 (W45 em L_hEIF4E5) não reflete na interação quando comparado com o complexo humano (Figura 45 A, painel direito). Um cenário diferente é observado no complexo de levedura, onde F68, estruturalmente alinhado com o W50 do LmEIF4E5, está envolvido na interação com

eIF4G (Figura 45 B, painel direito) por meio de regiões hidrofóbicas com a hélice C-terminal de eIF4G.

Figura 44 - Análise estrutural do EIF4E5 de *Leishmania major* na forma apo.

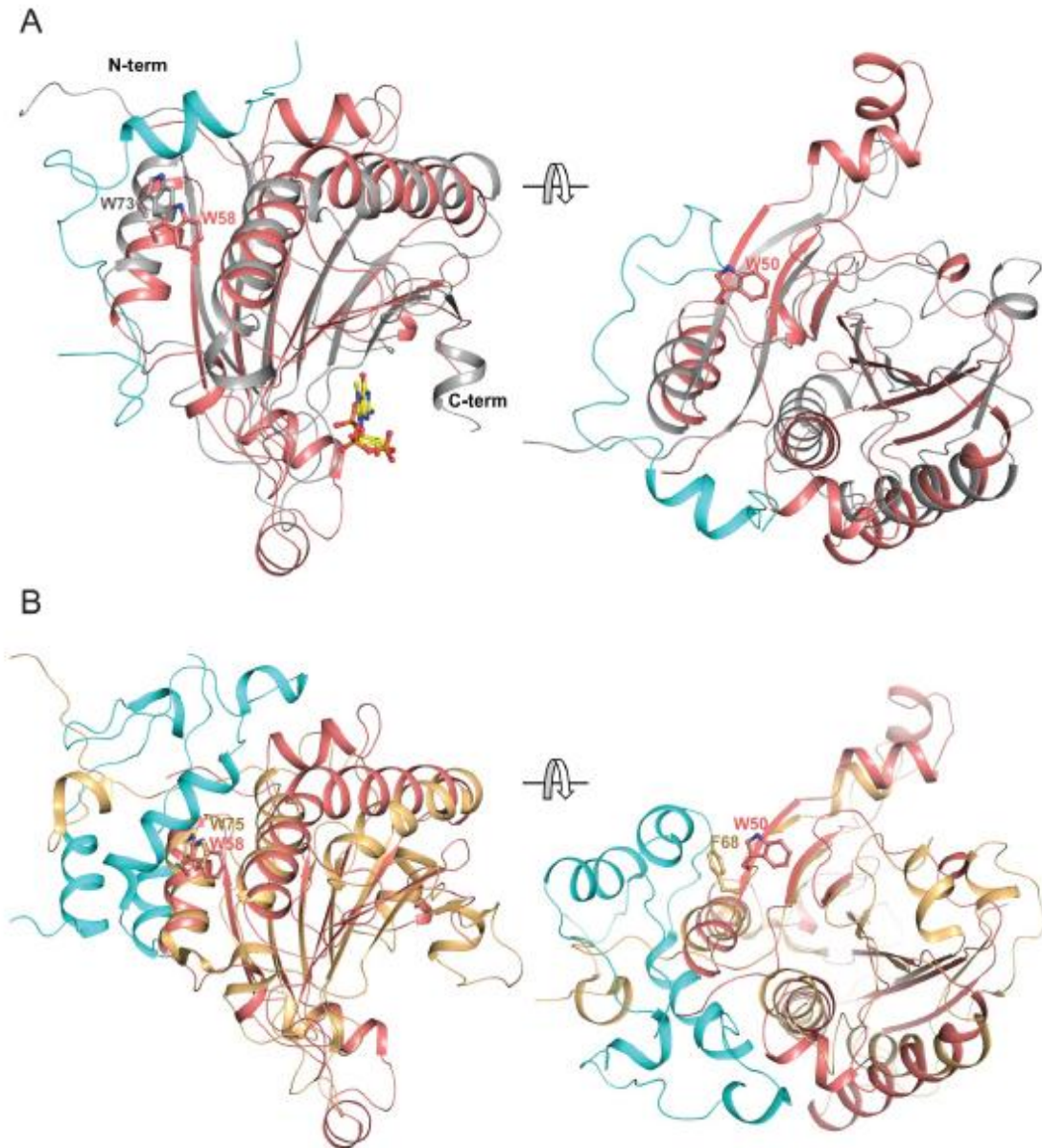


Fonte: Autoria própria (2021)

Nota: **A)** Superposição geral de LmEIF4E5 (salmão) e *cap4* ligado ao EIF4E5 de *T. cruzi* (PDB 6O7Y, roxa). A superposição evidencia rearranjos estruturais das regiões S1-S2 e S5-S6 após a ligação de *cap4*. O *cap4* é representado em bastões. **B)** Análise detalhada dos resíduos que interagem com *cap4* em TcEIF4E5 mostrando uma mudança conformacional significativa de W37, Y88 e E89 (numeração de LmEIF4E5, correspondendo aos W33, Y83 e E84 do TcEIF4E5) na ausência do ligante. Substituições não conservadas de resíduos que interagem com *cap4* também são indicadas. O último resíduo LmEIF4E5 que foi modelado na estrutura 3D é destacado por uma esfera. **C)** Superposição geral de LmEIF4E5 (salmão) e eIF4E humano na forma apo (PDB 2GPQ, cinza). Os resíduos de eIF4E humano conhecidos por participar do sanduíche m7G (W56 e W102) e os resíduos básicos que interagem com os grupos fosfato *cap*, juntamente com os resíduos correspondentes em LmEIF4E5 são representados em bastões. **D)** Análise da distribuição do fator B na estrutura LmEIF4E. A estrutura é colorida de acordo com os valores do fator B, de azul a vermelho, indicando

aumento do fator B. Os resíduos que interagem com o *cap* (em comparação com TcEIF4E5) são representados em bastões.

Figura 45 - Análise da interface eIF4E – eIF4G.



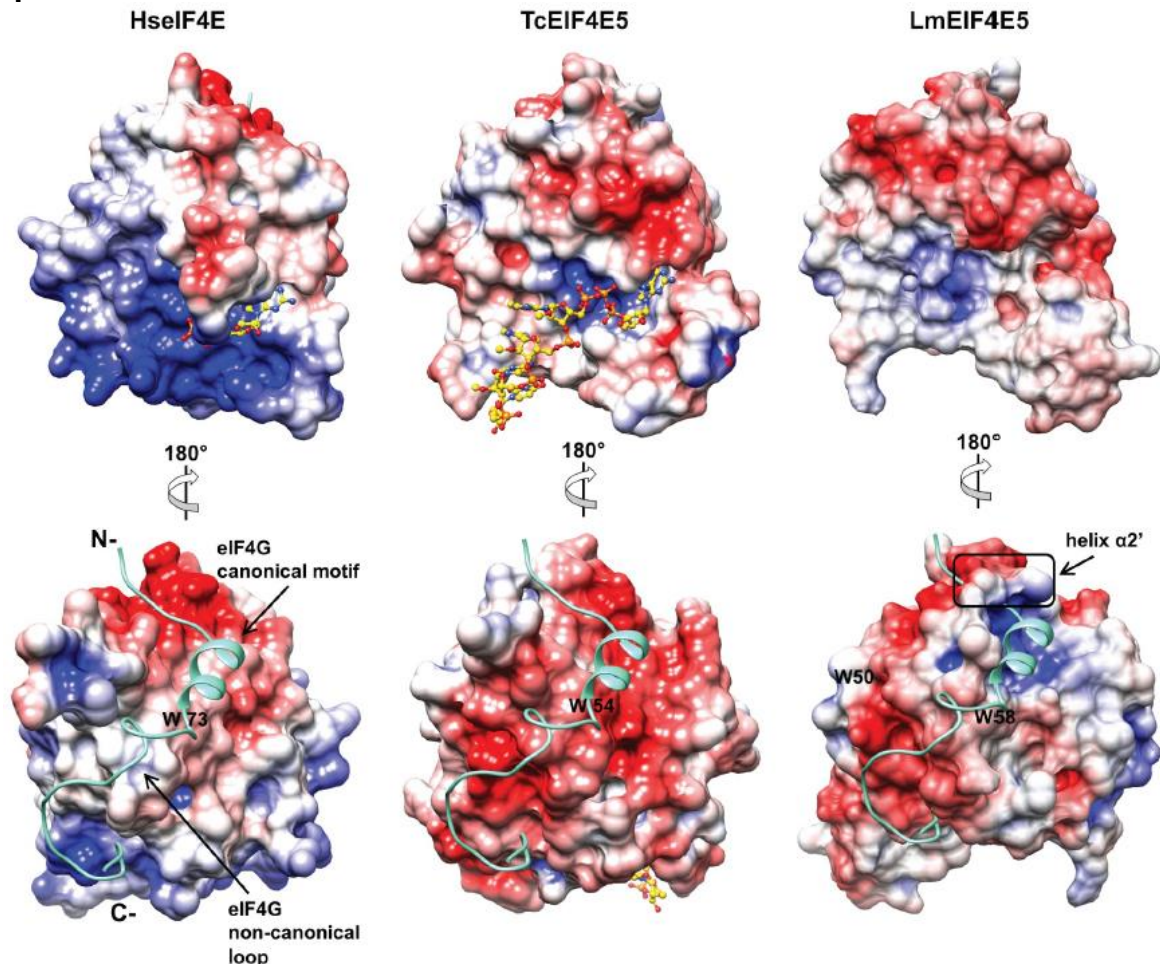
Fonte: Autoria própria (2021)

Nota: **A)** Superposição estrutural de EIF4E5 de *L. major* (salmão) e eIF4E1 de *H. sapiens* (cinza) em complexo com o peptídeo derivado do eIF4G (ciano) e m7GTP (bastões). As extremidades N e C-terminais, localizadas nos lados convexo e côncavo de eIF4E, respectivamente, são indicadas. À esquerda, o triptofano altamente conservado (W73 em HseIF4E1) envolvido na interação com a hélice canônica eIF4G é mostrado em bastão. O W58 correspondente em LmEIF4E5 também é mostrado. No painel direito, o W50 do LmEIF4E5 está destacado. O resíduo correspondente em HseIF4E1 (não mostrado) não está envolvido na interação eIF4G. **B)** Sobreposição estrutural de *LmEIF4E5* (salmão) e *S. cerevisiae* eIF4E (ouro) em complexo com um peptídeo eIF4G (ciano) (PDB 1RF8). À esquerda, o W75 de ScEIF4E, que participa da interação com a hélice canônica eIF4G e o W58 de LmEIF4E5 correspondente são mostrados em bastões. À direita, o W50 de *LmEIF4E5*, o resíduo correspondente ao F68 de ScEIF4E que participa da interação com o eIF4G, está destacado em bastão.

6.4.5 Análise comparativa da superfície de interação do fator EIF4E5

Em análise comparativa do potencial eletrostático de superfície do *LmEIF4E5* com o EIF4E5 de *T. cruzi* e eIF4E1 humano, foram reveladas diferenças interessantes (Figura 46). Também se observou uma distribuição de cargas distintas entre EIF4E5 e eIF4E1 humano no sítio de ligação ao *cap*. A sobreposição das estruturas de *T. cruzi* e *L. major* EIF4E5 com eIF4E1 humano em complexo com o peptídeo de interação eIF4G também evidencia uma diferença marcante na distribuição de carga na região correspondente à superfície de interação do eIF4G (Figura 46, painéis inferiores).

Figura 46 - Análise das superfícies de interação putativas EIF4E5 dos tripanosomatídeos



Fonte: Autoria própria (2021)

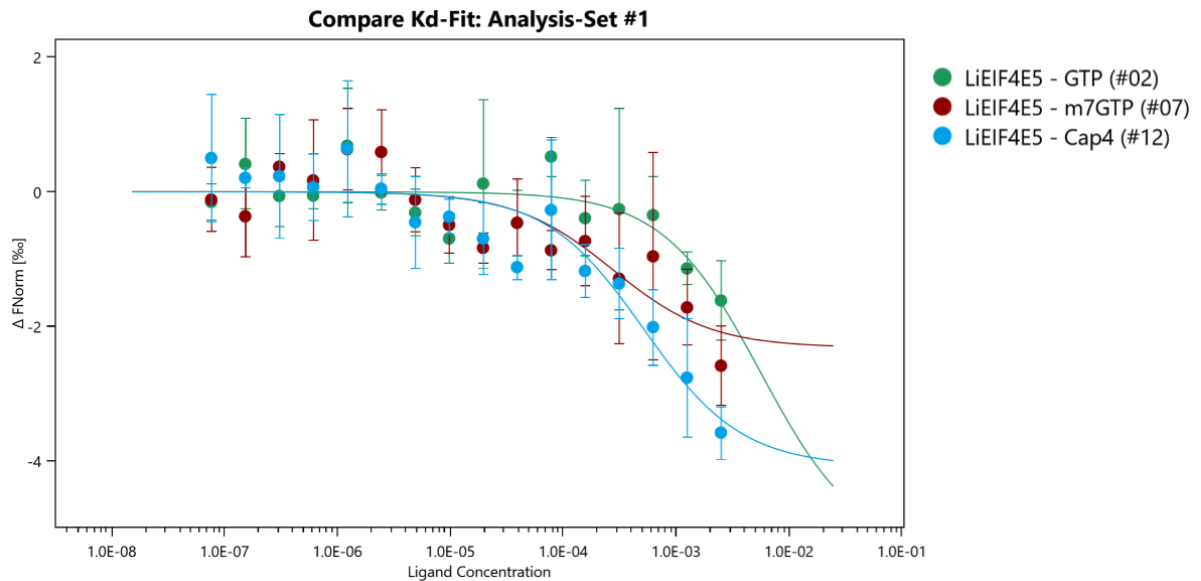
Nota: Comparação das superfícies de potencial eletrostático de *H. sapiens* eIF4E (à esquerda), *T. cruzi* EIF4E5 (meio) e *L. major* EIF4E5 (direita). Os limites para a visualização do mapa de contorno potencial são ± 5 kT / e. m7GTP e *cap4*, ligado a eIF4E1 humano e TcEIF4E5, respectivamente, são representados em bastões (painéis superiores). O peptídeo eIF4G ligado ao eIF4E humano é representado em ciano na parte inferior esquerda. As extremidades dos N e C terminais, bem como a hélice canônica e a alça de interação não canônica são indicadas. Com base na superposição estrutural, o peptídeo eIF4G do complexo humano é posicionado nas estruturas dos EIF4E5 dos

tripanosomatídeos, para fins comparativos. As posições do W73 de HseIF4E, W54 de TcEIF4E5, W58 e W50 de LmEIF4E5 são indicados. A posição da hélice $\alpha 2'$ de *LmEIF4E5* é indicada no painel inferior direito.

6.5 Ensaio de interação entre o fator EIF4E5 e o *cap4* através de termoforese em microescala (MST)

A análise da interação entre o fator EIF4E5 e o *cap4* foi realizada utilizando a proteína purificada para cristalização e o *cap4* sintetizado quimicamente. As amostras utilizadas foram diluídas três vezes, para serem utilizadas nos ensaios. Foi realizado um pré-teste com o fator EIF4E5 fusionado a histidina que foi incubado com uma molécula fluorescente de ligação a cauda de histidina para verificar se ela era compatível com o tampão de eluição diluído três vezes, e foi confirmada ausência de agregação entre as moléculas e os componentes da solução. A interação entre a proteína EIF4E5 e dois tipos de *cap* (*cap0* e *cap4*) foi analisada, e as curvas geradas avaliadas. Também foi realizado um experimento controle utilizando o fator EIF4E5 e GTP. A curva gerada é resultante da fluorescência emitida dos quatro capilares contendo apenas a proteína alvo, e quatro capilares contendo a proteína alvo com adição dos ligantes (*cap0* e *cap4*) ou do controle (GTP). Observa-se que as interações do fator EIF4E5 e o *cap0* e o *cap4* geraram uma curva diferente em relação a curva da proteína sem a adição de ligantes, indicando que os ligantes conseguem interagir com essa proteína (Gráfico 1).

Gráfico 1 - Gráfico da diferença de ligação do Fator EIF4E5 e o *cap0* e *cap4*.



Fonte: Autoria própria (2021)

Nota: Termoforese em Micro Escala que evidencia a diferença de ligação entre os ligantes estudados.

Os dados indicam que não houve interação com a molécula de GTP (controle negativo) Kda de 0,0148 enquanto o *cap0* o Kd foi de 0,000209 e o *cap4* de 0,000254. Como quanto maior o valor da constante de dissociação, menor é a afinidade entre o fator EIF4E5 e seus ligantes, logo o *cap0* se liga com maior afinidade que o *cap4*. Este dado corrobora com o encontrado para este mesmo homólogo em *T. cruzi*. O fator EIF4E5 é então uma proteína de ligação ao *cap* que, talvez por não interagir da mesma forma que outros homólogos, possui uma menor afinidade ao *cap4*, embora este seja o *cap* que compõe a estrutura dos tripanosomatídeos.

7 DISCUSSÃO

O estudo de fatores de tradução vem crescendo a cada ano, através de técnicas e novas tecnologias que estão evoluindo a cada dia, descobrimos novas informações sobre a síntese protéica, e como podemos aprender sobre o estudo da biologia molecular através dela (DE TARAFDER; PARAJULI; MAJUMDAR; KAÇAR *et al.*, 2021).

A tradução, a última etapa da expressão gênica propriamente, é um processo extremamente dinâmico que envolve 4 fases: iniciação, alongamento, término e reciclagem dos ribossomos. Em cada um dessas fases os ribossomos formam complexo transitórios juntamente com fatores de tradução auxiliares que facilitam a síntese protéica. A dinâmica de complexos de tradução com ribossomos, as mudanças conformacionais e os fatores de tradução que são diversos, desempenha um papel no direcionamento do que vai ser sintetizado pela célula (MERRICK; PAVITT, 2018; RODNINA, 2018).

Um grande desafio é entender como as interações fortemente relacionadas ou fracamente acopladas desses fatores leva a uma síntese de proteína rápida e precisa para a necessidade da célula (ROSS-KASCHITZA; ALTMANN, 2020) Aqui nesta discussão tentamos avaliar novas interações, que ainda não foram descritas, e reforçar o conhecimento adquirido com outros organismos a respeito dos fatores estudados e como a dinâmica entre elas permite regulação da síntese proteica.

eIF2 γ

Através de análise filogenética, determinar os domínios que estão presentes em ambos os homólogos de eIF2 γ , assim como descrito por Krauss, 2005, nos mostra onde a conservação é mais forte entre as subunidades, indicando a conservação ou não de propriedades como a sua ligação a substratos como tRNA e GTP/GDP (KRAUSS; PECYNA; KURZ; SASS, 2005). As informações encontradas neste trabalho são de uma importância evolutiva, comparável às obtidas para outros fatores que não são encontrados em todas as espécies de tripanosomatídeos, como por exemplo a PABP3, originalmente identificado no gênero *Leishmania* (DA COSTA LIMA; MOURA; REIS; VASCONCELOS *et al.*, 2010; KRAMER; BANNERMAN-

CHUKUALIM; ELLIS; BOULDEN *et al.*, 2013). Os genes de eIF2 γ 2 se encontra a jusante do eIF2 γ 1 com pequena região intergênica menor que 500pb entre os genes, indicando um evento de duplicação gênica em *tandem*. A taxa na qual novas cópias de genes aparecem nas populações influencia muito sua dinâmica evolutiva e a variação do número de cópias de genes, mas isto não indica que estas apenas aumentem a taxa de expressão deste gene ou que a função seja a mesma se houver mudanças de sequência (KATJU; BERGTHORSSON, 2013).

As características dos dois genes eIF2 γ apontam para a dualidade funcional dessas proteínas e demonstram que os genes não apresentam redundância de função (NASCIMENTO, 2016). Das interações encontradas o EIF5 se liga tanto ao EIF2 γ 1 como ao EIF2 γ 2 sendo esta uma das ligações mais críticas para o inibidor de dissociação de GTP GDI (JENNINGS; KERSHAW; WHITE; HOYLE *et al.*, 2016). Já o D123, também denominado Cdc123 é crítico para a atividade de eIF2. Cdc123 se liga mais fortemente a subunidade eIF2 γ de forma separada das outras subunidades do complexo EIF2. A região C-terminal do domínio III de eIF2 γ foi necessária e suficiente para a ligação de Cdc123, sendo que este participa do controle do ciclo celular em eucariotos superiores (PERZLMAIER; RICHTER; SEUFERT, 2013), podendo atuar também de alguma forma na diferenciação e replicação celular de tripanosomatídeos. Outra evidência apontando uma associação mais importante dos homólogos de eIF2 γ com o ciclo celular foi sua co-precipitação com a CRK3. Em estudos em *Leishmania mexicana* foi observada a essencialidade dessa kinase (HASSAN; FERGUSON; GRANT; MOTTRAM, 2001), além de sua associação com a divisão celular e interação com histonas (GRANT; HASSAN; ANDERSON; MOTTRAM, 1998), o que pode ainda corroborar com a presença de histonas no fator EIF2 γ 2.

Com base nos resultados de espectrometria de massas e as informações obtidas pela análise filogenética, estes achados encontrados nos leva a acreditar que os dois homólogos de eIF2 γ são potencialmente funcionais. Já as diferenças encontradas nas sequências gênicas e nas interações que divergem entre os fatores, indicaria semelhanças, pela sua região N-terminal, e divergências funcionais. O eIF2 γ 1 parece ter um papel mais voltado na tradução propriamente dita, enquanto o eIF2 γ 2, embora possa ter um papel na tradução, apresenta interações importantes com proteínas nucleares e proteínas que atuam na

diferenciação celular. É possível então que possa assumir algum papel a acrescentar que não seja na tradução, ressaltando que foi observado em trabalhos anteriores que o eIF2 γ 2 não consegue substituir na ausência do eIF2 γ 1 no seu papel na síntese protéica, sugerindo uma divergência de função fisiológica (NASCIMENTO, 2016).

EIF3

Os resultados obtidos com as subunidades do complexo eIF3 de *Leishmania infantum* são consistentes com a funcionalidade descrita para o complexo eIF3 em eucariotos em geral. Assim há uma predominância de maior interação das proteínas que se ligam diretamente, como a interação do EIF3E com os EIF3C e EIF3D, por exemplo, proteínas com as quais ele sabidamente interage diretamente, ou com os EIF3A e EIF3L, próximas a essas (WAGNER; HERRMANNOVÁ; MALÍK; PECLINOVSKÁ *et al.*, 2014). Da mesma forma, o EIF3I parece também estar associado com as proteínas que são reconhecidamente suas parceiras diretas (EIF3B, EIF3G e EIF3I) (KHOSHNEVIS; HAUER; MILÓN; STARK *et al.*, 2012; REZENDE; ASSIS; NUNES; DA COSTA LIMA *et al.*, 2014), mas também co-precipita com outras associadas a estas como os EIF3H e EIF3L. Para os EIF3D e EIF3G, ambos parecem co-precipitar mais com o complexo eIF3 completo sem muita distinção de subunidades, com destaque para a ausência do EIF3J. Isso quer dizer também que estas duas subunidades só ligam com o complexo inteiro já pronto, o que faz sentido visto que já foi observado, pela baixa interação com as outras subunidades, que o eIF3J se liga mais facilmente após todas as outras subunidades estarem formadas (SHA; BRILL; CABRERA; KLEIFELD *et al.*, 2009).

A subunidade EIF3J traz mais fracamente todas as subunidades do eIF3, mais especificamente o EIF3A (sua parceira direta) e os parceiros destas (EIF3C e EIF3H), mas não está vindo com nenhum dos demais. Em *Schizosaccharomyces pombe* foi relatado que o EIF3J se liga fracamente a outras proteínas do complexo eIF3, e apenas se liga a todo o complexo quando este está praticamente montado com todas as outras subunidades (SHA; BRILL; CABRERA; KLEIFELD *et al.*, 2009). Com base em nossos resultados isso pode estar ocorrendo também em *Leishmania*,

ou então estar havendo alguma expressão estágio específica do EIF3J ainda não detectada.

Em relação as RBPs que co-precipitaram com as subunidades do eIF3, chama a atenção também o perfil contrastante de RBPs vindo com o EIF3I mas ausentes do EIF3G (RBP29, RBP35, RBP42 e RBP45), em especial sabendo que essas duas subunidades de eIF3 são parceiras diretas. No caso da RBP42, esta também já foi observada em *L. braziliensis* coprecipitando com o EIF3I, mas não com o EIF3G, sugerindo um papel diferenciado para o EIF3I na formação de subcomplexos mRNA-proteína com funções diferenciadas para tradução (NOCUA; REQUENA; PUERTA, 2021). Como hipótese, o EIF3I pode então interagir quando não está associado ao complexo eIF3? Em estudos foi observado a interação de proteínas de ligação ao RNA com complexo eIF3 agindo de forma seletiva na tradução de RNAs mensageiros (FRANCISCO-VELILLA; AZMAN; MARTINEZ-SALAS, 2019). Além destes foi observado um número interessante de RBPs que interagem de forma seletiva com diferentes subunidades do complexo eIF3, indicando a ocorrência de outros subcomplexos específicos envolvendo diferentes subunidades de eIF3.

Um resultado relevante foi a co-precipitação da kinase CRK3 com o EIF3A. Esta co-precipitação mesmo em uma condição em que o EIF3A não baixa com as outras subunidades do eIF3 pode sugerir uma interação direta com o CRK3 mediada pela sua região C-terminal. A CRK3 é uma proteína que atua na divisão celular e que já foi relacionada anteriormente com o EIF2 γ e posteriormente com o fator EIF4E5. Além dessa, diversas outras kinases demonstram vir pela co-precipitação com as subunidades do complexo EIF3, demonstrando um alto grau de regulação que pode estar ocorrendo neste complexo, com impacto direto na tradução. As interações vistas nestas proteínas são similares ao que acontece com a CRK1, proteína esta que interage especificamente com o complexo EIF4E4/EIF4G3/PABP1 em *T. brucei* (AN; LIU; GOURGUECHON; WANG *et al.*, 2018).

A proteína EIF3A também apresentou relação de interação direta com os fatores EIF4E3 e EIF4G4, sendo estes parceiros em um complexo de iniciação da tradução em tripanosomatídeos. Este resultado foi observado apesar da ausência da região N-Terminal da proteína EIF3A (parcialmente clivada no processo de lise

celular) e corrobora com achados de Mellepattu,2015 com *L. amazonensis*. Este estudo identificou através ensaios de duplo híbrido que na região C-Terminal do EIF3A (especificamente os resíduos de 537 à 774) se encontra uma região de interação com outras proteínas, no caso o EIF4E1 (MELEPPATTU; KAMUS-ELIMELEH; ZINOVIEV; COHEN-MOR *et al.*, 2015).

Em relação às proteínas ribossomais, a P0 e principalmente a P2 podem possuir um papel diferencial devido a intensidade de co-precipitação com quase todas as subunidades do complexo eIF3 testadas. Em levedura foram observados os mesmos achados, através de ensaios de polissomas e de imunoprecipitação, onde foram identificados motivos específicos podem estar associados à interação destas subunidades a proteínas ribossomais (CHIU; WAGNER; HERRMANNOVÁ; BURELA *et al.*, 2010), mas ainda é necessário mais estudo neste campo de atuação pouco estudado.

Com os resultados de espectrometria de massas com as seis subunidades descritos aqui, foi possível então se avançar na análise de interações conservadas do complexo eIF3, como por exemplo a interação do EIF3E com o EIF3C e EIF3D, além da identificação de interações diferenciadas, talvez específicas de tripanosomatídeos. Também foi possível corroborar um maior papel da região C-terminal do EIF3A na iniciação da tradução além de se identificar diversas proteínas de ligação ao RNA e outras com possível papel auxiliar na síntese proteica.

EIF4E3

O fator eIF4E é conhecido como um fator de iniciação da tradução geral que permite a interação mRNA-ribossomo e a tradução dependente de *cap* em células eucarióticas, sendo bem estudado em diversas espécies (ROSS-KASCHITZA; ALTMANN, 2020). Este parece também ser o caso do EIF4E3, atuando através da sua interação específica com o eIF4G4, como observado em outros estudos (MOURA; REIS; XAVIER; DA COSTA LIMA *et al.*, 2015; YOFFE; LÉGER; ZINOVIEV; ZUBEREK *et al.*, 2009). Neste trabalho, foi observado um grande número de proteínas co-precipitando com o EIF4E3: ribossomais, RBPs, kinases, proteínas do proteasoma, enzimas metabólicas entre outras. Interessante e crucial

nesta investigação, foi observado que a maioria das proteínas que possuem um papel importante na tradução co-precipitaram apenas com o EIF4E3 com a mutação S75A. É importante ressaltar, contudo, que apenas em *L. amazonensis* e espécies mais próximas se observa a presença de uma serina na posição 75 (PEREIRA; MALVEZZI; NASCIMENTO; LIMA *et al.*, 2013),.

Foi observado que entre os fatores de iniciação tradução que co-precipitaram com o fator LaEIF4E3_{S75A}, além do EIF4G4 já comentado anteriormente, se destacam o EIF4AI e outras proteínas já esperadas pra formação do complexo de pré-iniciação da tradução, como subunidades do complexo eIF2 (alfa, beta e gamma), subunidades de EIF3 (EIF3E, EIF3L, EIF3K, EIF3H e EIF3C) e também os EIF5, EIF5A e EIF5B. A presença destes fatores demonstra o papel que o EIF4E3 realmente tem como um fator canônico na tradução dos tripanosomatídeos (ROSS-KASCHITZA; ALTMANN, 2020).

Através das análises realizadas aqui, um grupo de enzimas helicases do grupo DEAD-box foi observado com as duas proteínas EIF4E3, mas cada helicase interagia apenas com uma das versões do fator, sendo isto um fato curioso. Dentre as proteínas de ligação ao RNA, nós temos como uma proteína que veio em grande intensidade a PABP3, está que pode ser a proteína chave para confirmar com qual PABP este fator interage, visto que o EIF4E4 interage apenas com a PABP1 (FREIRE; STURM; CAMPBELL; DE MELO NETO, 2017). Além desta, outras proteínas RBP foram identificadas sendo interessante ressaltar as proteínas pumilio que são proteínas de ligação ao RNA de sequências específicas e que podem atuar como repressores pós-transcricionais, ligando-se a 3'-UTR de mRNA alvos (FILIPOVSKA; RAZIF; NYGÅRD; RACKHAM, 2011) e justificando uma forma de selecionar os mRNAs que são traduzidos pelo EIF4E3.

Dentre as kinases e proteína relacionadas que co-precipitam com o fator EIF4E3, as 14-3-3 foram observadas em trabalhos anteriores. Estas são proteínas de ligação de fosfoserina / fosfotreonina que se ligam a uma infinidade de diversos fatores de sinalização e regulam as interações proteína-proteína (FREIRE; STURM; CAMPBELL; DE MELO NETO, 2017). Um grande numero de proteínas ribossomais também co-precipitaram com o mutante LaEIF4E3_{S75A}, enquanto que apenas uma

proteína ribossomal baixou com a proteína selvagem, a proteína ribossomal L7, sendo necessário mais estudos para se tentar compreender essas observações.

O fator EIF4E3 já vem sendo estudado por diversos grupos que acreditam ter este um papel importante para a fisiologia celular. Foi observado que, após a diminuição deste fator por deleção gênica de uma das cópias, houve alteração de morfologia celular em curvas crescimento em cultura. Foi observado também que a substituição da serina 75 por uma alanina em *L. amazonensis* reduziu a sua capacidade de se concentrar em grânulos (SHRIVASTAVA; DRORY-RETWITZER; SHAPIRA, 2019), o que pode indicar que na presença da proteína fosforilada neste resíduo o EIF4E3 é estocado em grânulos. É então necessário sua desfosforilação para que se ligue a fatores voltados a síntese proteica. Em outro estudo foi sugerido que a proteína EIF4E3 pode ser encontrada em dois estados dentro da célula, ou ligado ao parceiro EIF4G4 ou encaminhado a grânulos de estresse como local de armazenamento em grânulos (ZINOVIEV; MANOR; SHAPIRA, 2012) corroborando com o trabalho anterior. Um fato interessante foi que em artigo publicado foi observado que a presença da mutação no resíduo de serina realmente afeta o perfil de fosforilação, mas que este diminuía parcialmente a ligação ao EIF4G4 (SHRIVASTAVA; TUPPERWAR; DRORY-RETWITZER; SHAPIRA, 2019), contrário ao que foi observado aqui.

EIF4E5

Através da determinação da estrutura de *L. major* EIF4E5, nós pudemos investigar semelhanças e diferenças na superfície de interação com eIF4G quando comparada com outros homólogos de eIF4E. A sobreposição estrutural de LmEIF4E5 com eIF4E1 humano (GRÜNER; PETER; WEBER; WOHLBOLD *et al.*, 2016) e eIF4E de levedura (GROSS; MOERKE; VON DER HAAR; LUGOVSKOY *et al.*, 2003) em complexo com peptídeos eIF4G mostra que, assumindo uma interface de ligação semelhante, o W58 de LmEIF4E5 (equivalente ao W53 em LiEIF4E5) mantém uma conformação altamente conservada e estaria em posição de interagir com a hélice conservada do eIF4G, como o W73 e W75 dos eIF4E humano e de levedura, respectivamente. Em contraponto, o W50 do LmEIF4E5 (W45 em LiEIF4E5), estruturalmente alinhado com o F68 de levedura, poderia estar envolvido

na interação com eIF4G por meio de regiões hidrofóbicas com a hélice C-terminal de eIF4G (GROSS; MOERKE; VON DER HAAR; LUGOVSKOY *et al.*, 2003).

O estudo do potencial eletrostático de superfície do LmEIF4E5 e a comparação com o EIF4E5 de *T. cruzi* (REOLON; VICHIER-GUERRE; DE MATOS; DUGUÉ *et al.*, 2019) e eIF4E1 humano (GRÜNER; PETER; WEBER; WOHLBOLD *et al.*, 2016) revelam diferenças interessantes. Em trabalho anterior já foi observada uma distribuição de cargas distintas entre o EIF4E5 de *T. cruzi* e o eIF4E1 humano no sítio de ligação do *cap* (REOLON; VICHIER-GUERRE; DE MATOS; DUGUÉ *et al.*, 2019). A sobreposição das estruturas do EIF4E5 de *T. cruzi* e *L. major* com o eIF4E1 humano, em complexo com o peptídeo de interação eIF4G, também demonstra que há uma diferença marcante na distribuição de cargas na região correspondente à superfície de interação de eIF4G. Na verdade, o alinhamento da sequência mostra que, embora vários dos resíduos do eIF4E humano envolvidos na interação com o eIF4G sejam conservados no EIF4E5 de tripanosomatídeos, regiões não conservadas em algumas posições coincidem com diferenças nas superfícies de potencial eletrostático. Essas substituições modificam a superfície em torno destes resíduos de triptofano altamente conservados e muito provavelmente têm um impacto no mecanismo de interação EIF4E5 com proteínas parceiras de. Notavelmente, a forma molecular e distribuição de cargas desta região é bastante distinta entre as três proteínas, em particular no EIF4E5 de *L. major*.

Em relação a interação do EIF4E5 e o *cap4* observada por termoforese em microescala, não houve interação com a molécula de GTP utilizado como controle, mas tanto com o *cap0* quanto com o *cap4* foi observada uma interação. Esta interação, entretanto, é de baixa força, assim como a encontrada em trabalhos que observou através de estrutura a ligação do EIF4E5 de *T. cruzi* (REOLON; VICHIER-GUERRE; DE MATOS; DUGUÉ *et al.*, 2019). É possível então que o fator EIF4E5, ainda que uma proteína de ligação ao *cap*, talvez não interaja da mesma forma que outros homólogos e possua uma menor afinidade ao *cap4*. Ou então necessita de uma ligação alostérica, talvez com homólogos de eIF4G, para ter uma maior afinidade ao *cap4*, tanto em *L. major* como em *T. cruzi*.

A cristalografia assim como o ensaio de termoforese em microescala apontam o grau de semelhanças que os ortólogos de EIF4E5 de *L. major* e *T. cruzi* possuem.

A forma de ligação ao *cap*, baseada na estrutura de *T. cruzi*, evidencia como é o encaixe EIF4E5-*cap* e simula a interação com a região da proteína EIF4G, evidenciando ainda quais as similaridades com outras proteínas de ligação ao *cap*. Ainda é necessário entender o baixo grau de interação ao *cap* encontrado e se a intensidade da interação se dá por interações alostéricas que aumentam sua afinidade. Essas comparações estruturais fornecem *insights* sobre diferentes características que podem ter um impacto na interação do EIF4E5 de tripanosomatídeos com seus parceiros de ligação.

8 CONCLUSÕES

- Ambos os homólogos de EIF2 γ de *Leishmania* apresentam interações com fatores de iniciação da tradução característicos deste grupo, mas diferem em parceiros específicos, que os indicam funcionais ainda que com papéis distintos na síntese proteica.
- As seis subunidades do eIF3 estudadas interagem principalmente com seus parceiros funcionais pré-estabelecidos, como em outros eucariotos, embora subunidades específicas apresentem interações diferenciadas com os complexos do tipo eIF4F de *Leishmania*, sendo que ao menos em relação ao complexo EIF4E3/EIF4G4, a região C-terminal da subunidade EIF3A parece estar envolvida nesta interação.
- Diferentes subunidades do eIF3 apresentam interações específicas com proteínas de ligação a RNA e outros grupos proteicos, incluindo quinases e helicases, sugerindo um papel regulatório dinâmico destas subunidades com proteínas que atuam de forma auxiliar na síntese proteica.
- O fator EIF4E3 em *Leishmania amazonenses* tem a capacidade de interagir com diversos fatores que atuam na síntese proteica, sendo sua atuação regulada pela fosforilação em um resíduo de serina único que, quando fosforilado, parece impedir a participação do EIF4E3 no processo de tradução.
- O EIF4E5 de *Leishmania* possui uma estrutura similar com a estrutura de seu ortólogo em *T. cruzi*, mostrando uma interface semelhante com outros homólogos de eIF4E na sua região de ligação a homólogos de eIF4G, porém com diferenças marcantes no seu potencial eletrostático de superfície e baixa afinidade de ligação ao *cap4* de tripanosomatídeos.

9 REFERENCIAS

ADDGENE. Watertown. Plasmid: pSP72. 2018.

AITKEN, C. E.; BEZNOSKOVÁ, P.; VLČKOVA, V.; CHIU, W. L. *et al.* Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. **Elife**, Cambridge. 5, 10 26 2016.

ALGIRE, M. A.; MAAG, D.; LORSCH, J. R. Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. **Mol Cell**, Washington 20, n. 2, p. 251-262, Oct 28 2005.

ALSFORD, S.; TURNER, D. J.; OBADO, S. O.; SANCHEZ-FLORES, A. *et al.* High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. **Genome Res**, Baltimore, 21, n. 6, p. 915-924, Jun 2011.

AN, T.; LIU, Y.; GOURGUECHON, S.; WANG, C. C. *et al.* CDK Phosphorylation of Translation Initiation Factors Couples Protein Translation with Cell-Cycle Transition. **Cell Rep**, Cambridge, 25, n. 11, p. 3204-3214.e3205, 12 11 2018.

ASSIS, L. A. D. **Caracterização do complexo de iniciação da tradução eIF3 e investigação de sua associação a complexos do tipo eIF4F em Leishmania infantum**. Recife, 2015. .

BASTIN, P.; PULLEN, T. J.; MOREIRA-LEITE, F. F.; GULL, K. Inside and outside of the trypanosome flagellum: a multifunctional organelle. **Microbes Infect**, Paris, 2, n. 15, p. 1865-1874, Dec 2000.

BEILSTEN-EDMANDS, V.; GORDIYENKO, Y.; KUNG, J. C.; MOHAMMED, S. *et al.* eIF2 interactions with initiator tRNA and eIF2B are regulated by post-translational modifications and conformational dynamics. **Cell Discov**, Yueyang , 1, p. 15020, 2015.

BOCHLER, A.; QUERIDO, J. B.; PRILEPSKAJA, T.; SOUFARI, H. *et al.* Structural Differences in Translation Initiation between Pathogenic Trypanosomatids and Their Mammalian Hosts. **Cell Rep**, Cambridge 33, n. 12, p. 108534, 12 22 2020.

BORCK, G.; SHIN, B. S.; STILLER, B.; MIMOUNI-BLOCH, A. *et al.* eIF2 γ mutation that disrupts eIF2 complex integrity links intellectual disability to impaired translation initiation. **Mol Cell**, Washington, 48, n. 4, p. 641-646, Nov 30 2012.

BORDEN, K.; CULJKOVIC-KRALJACIC, B.; VOLPON, L. EIF4E (eukaryotic translation initiation factor 4E), Austin. 2011.

BRINGAUD, F.; MÜLLER, M.; CERQUEIRA, G. C.; SMITH, M. *et al.* Members of a large retroposon family are determinants of post-transcriptional gene expression in Leishmania. **PLoS Pathog**, San Francisco 3, n. 9, p. 1291-1307, Sep 2007.

BROWNING, K. S.; BAILEY-SERRES, J. Mechanism of cytoplasmic mRNA translation. **Arabidopsis Book**, Washington, 13, p. e0176, 2015.

BUBIS, J.; MARTÍNEZ, J. C.; CALABOKIS, M.; FERREIRA, J. *et al.* The gene product of a *Trypanosoma equiperdum* ortholog of the cAMP-dependent protein kinase regulatory subunit is a monomeric protein that is not capable of binding cyclic nucleotides. **Biochimie**, Amsterdã, 146, p. 166-180, Mar 2018.

CAMPBELL, D. A.; THOMAS, S.; STURM, N. R. Transcription in kinetoplastid protozoa: why be normal? **Microbes Infect**, Paris, 5, n. 13, p. 1231-1240, Nov 2003.

CASTRO MACHADO, F.; BITTENCOURT-CUNHA, P.; MALVEZZI, A. M.; ARICO, M. *et al.* EIF2 α phosphorylation is regulated in intracellular amastigotes for the generation of infective *Trypanosoma cruzi* trypomastigote forms. **Cell Microbiol**, Londres, 22, n. 11, p. e13243, 11 2020.

CATE, J. H. Human eIF3: from 'blobology' to biological insight. **Philos Trans R Soc Lond B Biol Sci**, Londres, 372, n. 1716, 03 19 2017.

CENCIC, R.; DESFORGES, M.; HALL, D. R.; KOZAKOV, D. *et al.* Blocking eIF4E-eIF4G interaction as a strategy to impair coronavirus replication. **J Virol**, Washington, 85, n. 13, p. 6381-6389, Jul 2011.

CHIU, W. L.; WAGNER, S.; HERRMANNOVÁ, A.; BURELA, L. *et al.* The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. **Mol Cell Biol**, Washington, 30, n. 18, p. 4415-4434, Sep 2010.

CHOW, C.; CLOUTIER, S.; DUMAS, C.; CHOU, M. N. *et al.* Promastigote to amastigote differentiation of *Leishmania* is markedly delayed in the absence of PERK eIF2 α kinase-dependent eIF2 α phosphorylation. **Cell Microbiol**, New Jersey, 13, n. 7, p. 1059-1077, Jul 2011.

DA COSTA LIMA, T. D.; MOURA, D. M.; REIS, C. R.; VASCONCELOS, J. R. *et al.* Functional characterization of three *leishmania* poly(a) binding protein homologues with distinct binding properties to RNA and protein partners. **Eukaryot Cell**, Chicago, 9, n. 10, p. 1484-1494, Oct 2010.

DE LIMA, G. B.; DE LIMA CAVALCANTI, T. Y. V.; DE BRITO, A. N. A. L.; DE ASSIS, L. A. *et al.* The translation initiation factor EIF4E5 from *Leishmania*: crystal structure and interacting partners. **RNA Biol**, Vienna, p. 1-17, May 04 2021.

DE MELO NETO, O. P.; DA COSTA LIMA, T. D.; XAVIER, C. C.; NASCIMENTO, L. M. *et al.* The unique *Leishmania* EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. **RNA Biol**, Vienna, 12, n. 11, p. 1209-1221, 2015.

DE TARAFDER, A.; PARAJULI, N. P.; MAJUMDAR, S.; KAÇAR, B. *et al.* Kinetic Analysis Suggests Evolution of Ribosome Specificity in Modern Elongation Factor-Tus from "Generalist" Ancestors. **Mol Biol Evol**, Oxford, 38, n. 8, p. 3436-3444, 07 29 2021.

DES GEORGES, A.; DHOTE, V.; KUHN, L.; HELLEN, C. U. *et al.* Structure of mammalian eIF3 in the context of the 43S preinitiation complex. **Nature**, 525, n. 7570, p. 491-495, Sep 24 2015.

DHALIA, R.; MARINSEK, N.; REIS, C. R.; KATZ, R. *et al.* The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. **Nucleic Acids Res**, 34, n. 9, p. 2495-2507, 2006.

DHALIA, R.; REIS, C. R.; FREIRE, E. R.; ROCHA, P. O. *et al.* Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. **Mol Biochem Parasitol**, 140, n. 1, p. 23-41, Mar 2005.

DONNELLY, N.; GORMAN, A. M.; GUPTA, S.; SAMALI, A. The eIF2 α kinases: their structures and functions. **Cell Mol Life Sci**, 70, n. 19, p. 3493-3511, Oct 2013.

DOS SANTOS RODRIGUES, F. H.; FIRZUK, H.; BREEZE, A. L.; CAMERON, A. D. *et al.* The *Leishmania* PABP1-eIF4E4 interface: a novel 5'-3' interaction architecture for trans-spliced mRNAs. **Nucleic Acids Res**, 47, n. 3, p. 1493-1504, 02 20 2019.

ELLISDON, A. M.; STEWART, M. Structural biology of the PCI-protein fold. **Bioarchitecture**, 2, n. 4, p. 118-123, 2012 Jul-Aug 2012.

FILIPOVSKA, A.; RAZIF, M. F.; NYGÅRD, K. K.; RACKHAM, O. A universal code for RNA recognition by PUF proteins. **Nat Chem Biol**, 7, n. 7, p. 425-427, May 15 2011.

FRANCISCO-VELILLA, R.; AZMAN, E. B.; MARTINEZ-SALAS, E. Impact of RNA-Protein Interaction Modes on Translation Control: The Versatile Multidomain Protein Gemin5. **Bioessays**, 41, n. 4, p. e1800241, 04 2019.

FRASER, C. S. Quantitative studies of mRNA recruitment to the eukaryotic ribosome. **Biochimie**, 114, p. 58-71, Jul 2015.

FREIRE, E. R.; DHALIA, R.; MOURA, D. M.; DA COSTA LIMA, T. D. *et al.* The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. **Mol Biochem Parasitol**, 176, n. 1, p. 25-36, Mar 2011.

FREIRE, E. R.; MALVEZZI, A. M.; VASHISHT, A. A.; ZUBEREK, J. *et al.* *Trypanosoma brucei* translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog. **Eukaryot Cell**, Chicago, 13, n. 7, p. 896-908, Jul 2014.

FREIRE, E. R.; MOURA, D. M. N.; BEZERRA, M. J. R.; XAVIER, C. C. *et al.* *Trypanosoma brucei* EIF4E2 cap-binding protein binds a homolog of the histone-mRNA stem-loop-binding protein. **Curr Genet**, 64, n. 4, p. 821-839, Aug 2018.

FREIRE, E. R.; STURM, N. R.; CAMPBELL, D. A.; DE MELO NETO, O. P. The Role of Cytoplasmic mRNA Cap-Binding Protein Complexes in *Trypanosoma brucei* and Other Trypanosomatids. **Pathogens**, 6, n. 4, Oct 27 2017.

FREIRE, E. R.; VASHISHT, A. A.; MALVEZZI, A. M.; ZUBEREK, J. *et al.* eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in *Trypanosoma brucei*. **RNA**, 20, n. 8, p. 1272-1286, Aug 2014.

GAO, M.; RYCHLIK, W.; RHOADS, R. E. Cloning and characterization of human eIF4E genes. **J Biol Chem**, Baltimore, 273, n. 8, p. 4622-4628, Feb 1998.

GENUTH, N. R.; BARNA, M. Heterogeneity and specialized functions of translation machinery: from genes to organisms. **Nat Rev Genet**, 19, n. 7, p. 431-452, 07 2018.

GHOSH, A.; LIMA, C. D. Enzymology of RNA cap synthesis. **Wiley Interdiscip Rev RNA**, 1, n. 1, p. 152-172, 2010 Jul-Aug 2010.

GINGRAS, A. C.; RAUGHT, B.; SONENBERG, N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. **Annu Rev Biochem**, 68, p. 913-963, 1999.

GOMES-DUARTE, A.; LACERDA, R.; MENEZES, J.; ROMÃO, L. eIF3: a factor for human health and disease. **RNA Biol**, Vienna, 15, n. 1, p. 26-34, 01 02 2018.

GRANT, K. M.; HASSAN, P.; ANDERSON, J. S.; MOTTRAM, J. C. The *crk3* gene of *Leishmania mexicana* encodes a stage-regulated *cdc2*-related histone H1 kinase that associates with p12. **J Biol Chem**, Baltimore, 273, n. 17, p. 10153-10159, Apr 24 1998.

GROSS, J. D.; MOERKE, N. J.; VON DER HAAR, T.; LUGOVSKOY, A. A. *et al.* Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. **Cell**, 115, n. 6, p. 739-750, Dec 12 2003.

GRÜNER, S.; PETER, D.; WEBER, R.; WOHLBOLD, L. *et al.* The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation. **Mol Cell**, 64, n. 3, p. 467-479, 11 03 2016.

GUARNER, J. Chagas disease as example of a reemerging parasite. **Semin Diagn Pathol**, 36, n. 3, p. 164-169, May 2019.

HAO, P.; YU, J.; WARD, R.; LIU, Y. *et al.* Eukaryotic translation initiation factors as promising targets in cancer therapy. **Cell Commun Signal**, 18, n. 1, p. 175, 11 04 2020.

HASSAN, P.; FERGUSSON, D.; GRANT, K. M.; MOTTRAM, J. C. The CRK3 protein kinase is essential for cell cycle progression of *Leishmania mexicana*. **Mol Biochem Parasitol**, 113, n. 2, p. 189-198, Apr 06 2001.

JACKSON, R. J.; HELLEN, C. U.; PESTOVA, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. **Nat Rev Mol Cell Biol**, 11, n. 2, p. 113-127, Feb 2010.

JAGUS, R.; BACHVAROFF, T. R.; JOSHI, B.; PLACE, A. R. Diversity of Eukaryotic Translational Initiation Factor eIF4E in Protists. **Comp Funct Genomics**, 2012, p. 134839, 2012.

JENNINGS, M. D.; KERSHAW, C. J.; WHITE, C.; HOYLE, D. *et al.* eIF2 β is critical for eIF5-mediated GDP-dissociation inhibitor activity and translational control. **Nucleic Acids Res**, 44, n. 20, p. 9698-9709, Nov 16 2016.

JOSHI, B.; CAMERON, A.; JAGUS, R. Characterization of mammalian eIF4E-family members. **Eur J Biochem**, 271, n. 11, p. 2189-2203, Jun 2004.

JOSHI, B.; LEE, K.; MAEDER, D. L.; JAGUS, R. Phylogenetic analysis of eIF4E-family members. **BMC Evol Biol**, 5, p. 48, Sep 2005.

KAMENSKA, A.; SIMPSON, C.; STANDART, N. eIF4E-binding proteins: new factors, new locations, new roles. **Biochem Soc Trans**, 42, n. 4, p. 1238-1245, Aug 2014.

KASHIWAGI, K.; ITO, T.; YOKOYAMA, S. Crystal structure of eIF2B and insights into eIF2-eIF2B interactions. **FEBS J**, 284, n. 6, p. 868-874, 03 2017.

KASHIWAGI, K.; TAKAHASHI, M.; NISHIMOTO, M.; HIYAMA, T. B. *et al.* Crystal structure of eukaryotic translation initiation factor 2B. **Nature**, 531, n. 7592, p. 122-125, Mar 03 2016.

KATJU, V.; BERGTHORSSON, U. Copy-number changes in evolution: rates, fitness effects and adaptive significance. **Front Genet**, 4, p. 273, Dec 10 2013.

KAUFER, A.; ELLIS, J.; STARK, D.; BARRATT, J. The evolution of trypanosomatid taxonomy. **Parasit Vectors**, 10, n. 1, p. 287, Jun 2017.

KEIPER, B. D.; LAMPHEAR, B. J.; DESHPANDE, A. M.; JANKOWSKA-ANYSZKA, M. *et al.* Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. **J Biol Chem, Baltimore**, 275, n. 14, p. 10590-10596, Apr 2000.

KENNEDY, P. G. E. Update on human African trypanosomiasis (sleeping sickness). **J Neurol**, 266, n. 9, p. 2334-2337, Sep 2019.

KHOSHNEVIS, S.; HAUER, F.; MILÓN, P.; STARK, H. *et al.* Novel insights into the architecture and protein interaction network of yeast eIF3. **RNA**, 18, n. 12, p. 2306-2319, Dec 2012.

KIMBALL, S. R. Eukaryotic initiation factor eIF2. **Int J Biochem Cell Biol**, 31, n. 1, p. 25-29, Jan 1999.

KISSINGER, J. C. A tale of three genomes: the kinetoplastids have arrived. **Trends Parasitol**, 22, n. 6, p. 240-243, Jun 2006.

KOMAR, A. A.; MERRICK, W. C. A Retrospective on eIF2A-and Not the Alpha Subunit of eIF2. **Int J Mol Sci**, 21, n. 6, Mar 17 2020.

KRAMER, S. Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids. **Mol Biochem Parasitol**, 181, n. 2, p. 61-72, Feb 2012.

KRAMER, S.; BANNERMAN-CHUKUALIM, B.; ELLIS, L.; BOULDEN, E. A. *et al.* Differential localization of the two *T. brucei* poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. **PLoS One**, 8, n. 1, p. e54004, 2013.

KRAUSS, V.; PECYNA, M.; KURZ, K.; SASS, H. Phylogenetic mapping of intron positions: a case study of translation initiation factor eIF2gamma. **Mol Biol Evol**, Oxford, 22, n. 1, p. 74-84, Jan 2005.

LAMA, D.; PRADHAN, M. R.; BROWN, C. J.; EAPEN, R. S. *et al.* Water-Bridge Mediates Recognition of mRNA Cap in eIF4E. **Structure**, 25, n. 1, p. 188-194, 01 2017.

LEE, A. S.; KRANZUSCH, P. J.; DOUDNA, J. A.; CATE, J. H. eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. **Nature**, 536, n. 7614, p. 96-99, 08 04 2016.

LI, K.; ZHOU, S.; GUO, Q.; CHEN, X. *et al.* The eIF3 complex of. **RNA**, 23, n. 3, p. 333-345, 03 2017.

LUKES, J.; GUILBRIDE, D. L.; VOTÝPKA, J.; ZÍKOVÁ, A. *et al.* Kinetoplast DNA network: evolution of an improbable structure. **Eukaryot Cell**, Chicago, 1, n. 4, p. 495-502, Aug 2002.

LUKEŠ, J.; SKALICKÝ, T.; TÝČ, J.; VOTÝPKA, J. *et al.* Evolution of parasitism in kinetoplastid flagellates. **Mol Biochem Parasitol**, 195, n. 2, p. 115-122, Jul 2014.

MALINA, A.; CENCIC, R.; PELLETIER, J. Targeting translation dependence in cancer. **Oncotarget**, 2, n. 1-2, p. 76-88, 2011 Jan-Feb 2011.

MARCOTRIGIANO, J.; GINGRAS, A. C.; SONENBERG, N.; BURLEY, S. K. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. **Mol Cell**, 3, n. 6, p. 707-716, Jun 1999.

MATSUO, H.; LI, H.; MCGUIRE, A. M.; FLETCHER, C. M. *et al.* Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. **Nat Struct Biol**, 4, n. 9, p. 717-724, Sep 1997.

MCCALL, L. I.; MCKERROW, J. H. Determinants of disease phenotype in trypanosomatid parasites. **Trends Parasitol**, 30, n. 7, p. 342-349, Jul 2014.

MELEPPATTU, S.; KAMUS-ELIMELEH, D.; ZINOVIEV, A.; COHEN-MOR, S. *et al.* The eIF3 complex of *Leishmania*-subunit composition and mode of recruitment to different cap-binding complexes. **Nucleic Acids Res**, 43, n. 13, p. 6222-6235, Jul 27 2015.

MERRICK, W. C.; PAVITT, G. D. Protein Synthesis Initiation in Eukaryotic Cells. **Cold Spring Harb Perspect Biol**, 10, n. 12, 12 03 2018.

MOERKE, N. J.; AKTAS, H.; CHEN, H.; CANTEL, S. *et al.* Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. **Cell**, 128, n. 2, p. 257-267, Jan 26 2007.

MOURA, D. M.; REIS, C. R.; XAVIER, C. C.; DA COSTA LIMA, T. D. *et al.* Two related trypanosomatid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. **RNA Biol**, Vienna, 12, n. 3, p. 305-319, 2015.

NARASIMHAN, J.; JOYCE, B. R.; NAGULESWARAN, A.; SMITH, A. T. *et al.* Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in *Toxoplasma gondii*. **J Biol Chem**, Baltimore, 283, n. 24, p. 16591-16601, Jun 13 2008.

NASCIMENTO, L. M. D. **Complexo eIF2 em Leishmania sp.: expressão proteica, função biológica, interações moleculares e descrição de novo fator de iniciação da tradução.** 2016. - Disponível em: <https://repositorio.ufpe.br/handle/123456789/24444>.

NCBI. Disponível em: <https://www.ncbi.nlm.nih.gov/>.

NCBI. Map view. 2016.

NETO, O. P. D. M.; REIS, C. R. D. S.; MOURA, D. M. D. N.; CARRINGTON, M. Unique and Conserved Features of the Protein Synthesis Apparatus in Parasitic Trypanosomatid (*Trypanosoma* and *Leishmania*) Species. *In*. Evolution of the Protein Synthesis Machinery and Its Regulation, 2016. p. 435-475.

NOCUA, P. A.; REQUENA, J. M.; PUERTA, C. J. Identification of the interactomes associated with SCD6 and RBP42 proteins in *Leishmania braziliensis*. **J Proteomics**, 233, p. 104066, 02 20 2021.

PACE, D. Leishmaniasis. **J Infect**, 69 Suppl 1, p. S10-18, Nov 2014.

PELLETIER, J.; GRAFF, J.; RUGGERO, D.; SONENBERG, N. Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. **Cancer Res**, 75, n. 2, p. 250-263, Jan 2015.

PEREIRA, M. M.; MALVEZZI, A. M.; NASCIMENTO, L. M.; LIMA, T. D. *et al.* The eIF4E subunits of two distinct trypanosomatid eIF4F complexes are subjected to differential post-translational modifications associated to distinct growth phases in culture. **Mol Biochem Parasitol**, 190, n. 2, p. 82-86, Aug 2013.

PERZLMAIER, A. F.; RICHTER, F.; SEUFERT, W. Translation initiation requires cell division cycle 123 (Cdc123) to facilitate biogenesis of the eukaryotic initiation factor 2 (eIF2). **J Biol Chem**, Baltimore, 288, n. 30, p. 21537-21546, Jul 26 2013.

PETTERSSON, F.; DEL RINCON, S. V.; MILLER, W. H. Eukaryotic translation initiation factor 4E as a novel therapeutic target in hematological malignancies and beyond. **Expert Opin Ther Targets**, 18, n. 9, p. 1035-1048, Sep 2014.

PTUSHKINA, M.; BERTHELOT, K.; VON DER HAAR, T.; GEFFERS, L. *et al.* A second eIF4E protein in *Schizosaccharomyces pombe* has distinct eIF4G-binding properties. **Nucleic Acids Res**, 29, n. 22, p. 4561-4569, Nov 2001.

REN, B.; GUPTA, N. Taming Parasites by Tailoring Them. **Front Cell Infect Microbiol**, 7, p. 292, 2017.

REOLON, L. W.; VICHIER-GUERRE, S.; DE MATOS, B. M.; DUGUÉ, L. *et al.* Crystal structure of the *Trypanosoma cruzi* EIF4E5 translation factor homologue in complex with mRNA cap-4. **Nucleic Acids Res**, 47, n. 11, p. 5973-5987, 06 20 2019.

REZENDE, A. M.; ASSIS, L. A.; NUNES, E. C.; DA COSTA LIMA, T. D. *et al.* The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates-identification of conserved and divergent features based on orthologue analysis. **BMC Genomics**, 15, p. 1175, Dec 23 2014.

RHOADS, R. E. eIF4E: new family members, new binding partners, new roles. **J Biol Chem, Baltimore**, 284, n. 25, p. 16711-16715, Jun 2009.

RODNINA, M. V. Translation in Prokaryotes. **Cold Spring Harb Perspect Biol**, 10, n. 9, 09 04 2018.

ROM, E.; KIM, H. C.; GINGRAS, A. C.; MARCOTRIGIANO, J. *et al.* Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. **J Biol Chem, Baltimore**, 273, n. 21, p. 13104-13109, May 1998.

ROSS-KASCHITZA, D.; ALTMANN, M. eIF4E and Interactors from Unicellular Eukaryotes. **Int J Mol Sci**, 21, n. 6, Mar 21 2020.

RUUD, K. A.; KUHLOW, C.; GOSS, D. J.; BROWNING, K. S. Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*. **J Biol Chem, Baltimore**, 273, n. 17, p. 10325-10330, Apr 1998.

RÔMULO; MORAES, M. D. N. **Identificação de motivos relevantes para a função do fator de iniciação da tradução EIF4E3 de *Leishmania sp.*** Orientador: NETO, O. P. D. M. 2015. (Master) - BIOCÊNCIAS E BIOTECNOLOGIA EM SAÚDE.

SENOHRABKOVA, L.; MALCOVA, I.; HASEK, J. An aggregation-prone mutant of eIF3a forms reversible assemblies escaping spatial control in exponentially growing yeast cells. **Curr Genet**, 65, n. 4, p. 919-940, Aug 2019.

SHA, Z.; BRILL, L. M.; CABRERA, R.; KLEIFELD, O. *et al.* The eIF3 interactome reveals the translasome, a supercomplex linking protein synthesis and degradation machineries. **Mol Cell**, 36, n. 1, p. 141-152, Oct 09 2009.

SHIROKIKH, N. E.; PREISS, T. Translation initiation by cap-dependent ribosome recruitment: Recent insights and open questions. **Wiley Interdiscip Rev RNA**, 9, n. 4, p. e1473, 07 2018.

SHRIVASTAVA, R.; DRORY-RETWITZER, M.; SHAPIRA, M. Nutritional stress targets LeishIF4E-3 to storage granules that contain RNA and ribosome components in Leishmania. **PLoS Negl Trop Dis**, 13, n. 3, p. e0007237, 03 2019.

SHRIVASTAVA, R.; TUPPERWAR, N.; DRORY-RETWITZER, M.; SHAPIRA, M. Deletion of a Single LeishIF4E-3 Allele by the CRISPR-Cas9 System Alters Cell Morphology and Infectivity of. **mSphere**, 4, n. 5, 09 04 2019.

SIMONETTI, A.; BRITO QUERIDO, J.; MYASNIKOV, A. G.; MANCERA-MARTINEZ, E. *et al.* eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. **Mol Cell**, 63, n. 2, p. 206-217, 07 21 2016.

SIMPSON, A. G.; STEVENS, J. R.; LUKES, J. The evolution and diversity of kinetoplastid flagellates. **Trends Parasitol**, 22, n. 4, p. 168-174, Apr 2006.

SONENBERG, N.; HINNEBUSCH, A. G. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. **Cell**, 136, n. 4, p. 731-745, Feb 2009.

STRUDWICK, S.; BORDEN, K. L. The emerging roles of translation factor eIF4E in the nucleus. **Differentiation**, 70, n. 1, p. 10-22, Mar 2002.

STUART, K.; BRUN, R.; CROFT, S.; FAIRLAMB, A. *et al.* Kinetoplastids: related protozoan pathogens, different diseases. **J Clin Invest**, 118, n. 4, p. 1301-1310, Apr 2008.

TASLIMI, Y.; ZAHEDIFARD, F.; RAFATI, S. Leishmaniasis and various immunotherapeutic approaches. **Parasitology**, 145, n. 4, p. 497-507, 04 2018.

THAKOR, N.; SMITH, M. D.; ROBERTS, L.; FAYE, M. D. *et al.* Cellular mRNA recruits the ribosome via eIF3-PABP bridge to initiate internal translation. **RNA Biol**, Vienna, 14, n. 5, p. 553-567, 05 04 2017.

TONELLI, R. R.; AUGUSTO, L. A. S.; CASTILHO, B. A.; SCHENKMAN, S. Protein synthesis attenuation by phosphorylation of eIF2 α is required for the differentiation of *Trypanosoma cruzi* into infective forms. **PLoS One**, 6, n. 11, p. e27904, 2011.

TRITRYPDB.ORG.

VILLA, N.; DO, A.; HERSHEY, J. W.; FRASER, C. S. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. **J Biol Chem, Baltimore**, 288, n. 46, p. 32932-32940, Nov 15 2013.

WAGNER, S.; HERRMANNOVÁ, A.; MALÍK, R.; PECLINOVSKÁ, L. *et al.* Functional and biochemical characterization of human eukaryotic translation initiation factor 3 in living cells. **Mol Cell Biol**, 34, n. 16, p. 3041-3052, Aug 2014.

WHO. **Leishmaniasis, the vector.** 2018. Disponível em: <http://www.who.int/leishmaniasis/disease/vector/en/>.

WHO. Global leishmaniasis surveillance, 2017–2018, and first report on 5 additional indicators. 2020.

WOLF, D. A.; LIN, Y.; DUAN, H.; CHENG, Y. eIF-Three to Tango: emerging functions of translation initiation factor eIF3 in protein synthesis and disease. **J Mol Cell Biol**, 12, n. 6, p. 403-409, 07 03 2020.

WORTHAM, N. C.; PROUD, C. G. eIF2B: recent structural and functional insights into a key regulator of translation. **Biochem Soc Trans**, 43, n. 6, p. 1234-1240, Dec 2015.

YAZAKI, E.; ISHIKAWA, S. A.; KUME, K.; KUMAGAI, A. *et al.* Global Kinetoplastea phylogeny inferred from a large-scale multigene alignment including parasitic species for better understanding transitions from a free-living to a parasitic lifestyle. **Genes Genet Syst**, 92, n. 1, p. 35-42, Sep 12 2017.

YOFFE, Y.; LÉGER, M.; ZINOVIEV, A.; ZUBEREK, J. *et al.* Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions. **Nucleic Acids Res**, 37, n. 10, p. 3243-3253, Jun 2009.

YOFFE, Y.; ZUBEREK, J.; LERER, A.; LEWDOROWICZ, M. *et al.* Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania. **Eukaryot Cell**, Chicago, 5, n. 12, p. 1969-1979, Dec 2006.

YOFFE, Y.; ZUBEREK, J.; LEWDOROWICZ, M.; ZEIRA, Z. *et al.* Cap-binding activity of an eIF4E homolog from Leishmania. **RNA**, 10, n. 11, p. 1764-1775, Nov 2004.

ZHANG, M.; FENNELL, C.; RANFORD-CARTWRIGHT, L.; SAKTHIVEL, R. *et al.* The Plasmodium eukaryotic initiation factor-2alpha kinase IK2 controls the latency of sporozoites in the mosquito salivary glands. **J Exp Med**, 207, n. 7, p. 1465-1474, Jul 05 2010.

ZINOVIEV, A.; LÉGER, M.; WAGNER, G.; SHAPIRA, M. A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways. **Nucleic Acids Res**, 39, n. 19, p. 8404-8415, Oct 2011.

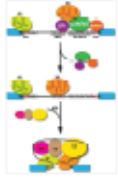
ZINOVIEV, A.; MANOR, S.; SHAPIRA, M. Nutritional stress affects an atypical cap-binding protein in Leishmania. **RNA Biol**, Vienna, 9, n. 12, p. 1450-1460, Dec 2012.

APÊNDICE A – ARTIGO PUBLICADO

The translation initiation factor EIF4E5 from *Leishmania*: crystal structure and interacting partners.

Artigo anexo na página seguinte

O doutorando e primeiro autor deste artigo publicado na *RNA Biology* foi responsável pela participação na produção de construções plasmidiais análises de sequencias, manipulação do DNA, ensaios de *western blot*, Expressão, purificação por FPLC e cristalização da proteína EIF4E5 de *L. major*.





The translation initiation factor EIF4E5 from *Leishmania*: crystal structure and interacting partners

Gustavo Barbosa de Lima, Thaíse Yasmine Vasconcelos de Lima Cavalcanti, Adriana Neuman Albuquerque Lins Moura de Brito, Ludmilla Arruda de Assis, Rafaela Paiva Andrade-Vieira, Eden Ribeiro Freire, Tatiana Reichert da Silva Assunção, Christian Robson de Souza Reis, Nilson Ivo Tonin Zanchin, Beatriz Gomes Guimarães & Osvaldo Pompílio de-Melo-Neto


To cite this article: Gustavo Barbosa de Lima, Thaíse Yasmine Vasconcelos de Lima Cavalcanti, Adriana Neuman Albuquerque Lins Moura de Brito, Ludmilla Arruda de Assis, Rafaela Paiva Andrade-Vieira, Eden Ribeiro Freire, Tatiana Reichert da Silva Assunção, Christian Robson de Souza Reis, Nilson Ivo Tonin Zanchin, Beatriz Gomes Guimarães & Osvaldo Pompílio de-Melo-Neto (2021): The translation initiation factor EIF4E5 from *Leishmania*: crystal structure and interacting partners, RNA Biology, DOI: [10.1080/15476286.2021.1918919](https://doi.org/10.1080/15476286.2021.1918919)


To link to this article: <https://doi.org/10.1080/15476286.2021.1918919>


 View supplementary material [↗](#)

 Published online: 04 May 2021.

 Submit your article to this journal [↗](#)

 Article views: 135

 View related articles [↗](#)

 View Crossmark data [↗](#)

RESEARCH PAPER



The translation initiation factor EIF4E5 from *Leishmania*: crystal structure and interacting partners

Gustavo Barbosa de Lima^a, Thaíse Yasmine Vasconcelos de Lima Cavalcanti^{a,b}, Adriana Neuman Albuquerque Lins Moura de Brito^{a,b}, Ludmilla Arruda de Assis^a, Rafaela Paiva Andrade-Vieira^a, Eden Ribeiro Freire^{a,c}, Tatiana Reichert da Silva Assunção^c, Christian Robson de Souza Reis^a, Nilson Ivo Tonin Zanchin^{b,c}, Beatriz Gomes Guimarães^{b,c}, and Osvaldo Pompílio de-Melo-Neto^b

^aDepartamento de Microbiologia, Instituto Aggeu Magalhães, FIOCRUZ-PE, Av. Moraes Rego s/n, Recife-PE, Brazil; ^bPrograma de Pós-Graduação em Genética, Universidade Federal de Pernambuco, Recife-PE, Brazil; ^cLaboratório de Biologia Estrutural e Engenharia de Proteínas, Instituto Carlos Chagas, FIOCRUZ-PR, Curitiba-PR, Brazil

ABSTRACT

The mRNA cap-binding protein, eIF4E, mediates the recognition of the mRNA 5' end and, as part of the heterotrimeric eIF4F complex, facilitates the recruitment of the ribosomal subunits to initiate eukaryotic translation. Various regulatory events involving eIF4E and a second eIF4F subunit, eIF4G, are required for proper control of translation initiation. In pathogenic trypanosomatids, six eIF4Es and five eIF4Gs have been described, several forming different eIF4F-like complexes with yet unresolved roles. EIF4E5 is one of the least known of the trypanosomatid eIF4Es and has not been characterized in *Leishmania* species. Here, we used immunoprecipitation assays, combined with mass-spectrometry, to identify major EIF4E5 interacting proteins in *L. infantum*. A constitutively expressed, HA-tagged, EIF4E5 co-precipitated mainly with EIF4G1 and binding partners previously described in *Trypanosoma brucei*, EIF4G1-IP, RBP43 and the 14-3-3 proteins. In contrast, no clear co-precipitation with EIF4G2, also previously reported, was observed. EIF4E5 also co-precipitated with protein kinases, possibly associated with cell-cycle regulation, selected RNA binding proteins and histones. Phosphorylated residues were identified and mapped to the *Leishmania*-specific C-terminal end. Mutagenesis of the tryptophan residue (W53) postulated to mediate interactions with protein partners or of a neighbouring tryptophan conserved in *Leishmania* (W45) did not substantially impair the identified interactions. Finally, the crystal structure of *Leishmania* EIF4E5 evidences remarkable differences in the eIF4G interfacing region, when compared with human eIF4E-1 and with its *Trypanosoma* orthologue. Mapping of its C-terminal end near the cap-binding site also imply relevant differences in cap-binding function and/or regulation.

ARTICLE HISTORY

Received 13 October 2020
Revised 15 March 2021
Accepted 14 April 2021

KEYWORDS




Translation Initiation;
Leishmania; EIF4E5; protein
Interaction; crystal structure


Introduction

Trypanosomatids are early branching protozoan parasites belonging to the class Kinetoplastida and characterized by complex life cycles that can alternate between two different hosts. Many species are pathogenic to a wide variety of animals, humans and plants, with *Trypanosoma cruzi*, *Trypanosoma brucei* and *Leishmania sp.* being the aetiologic agents of neglected human diseases [1]. These organisms are further characterized by their unique molecular biology. For instance, trypanosomatid genomes reveal functionally unrelated genes organized in large polycistronic clusters [2]. These clusters are constitutively transcribed into long polycistronic precursors which are subsequently processed by *trans*-splicing and polyadenylation into mature monocistronic mRNAs. As a result, all mature mRNAs have the *trans*-spliced SL (spliced leader) sequence plus the modified cap-4 on their 5' ends [3]. Regulation of gene expression in trypanosomatids is thus largely dependent on post-transcriptional mechanisms mainly

acting upon the different aspects of mRNA metabolism, as well as during their translation [4–6].

In eukaryotes, the heterotrimeric eIF4F complex is essential for the cap-dependent mRNA translation and is a main target of translation regulation. eIF4F is responsible for the mRNA 5' cap recognition and for the recruitment of the 40S ribosomal subunit to the mRNAs, in order to properly initiate translation. It is formed by the cap-binding protein eIF4E, the RNA helicase eIF4A and the large modular protein eIF4G. During translation initiation, its activity is enhanced by a direct interaction with the poly(A) binding protein (PABP), bound to the 3' end of the mRNAs [7–9]. The eIF4E subunit is responsible for recognizing the methylated guanosine residue of the cap structure and its interaction with eIF4G is required for ribosomal recruitment [10–12]. In unicellular eukaryotes, multiple homologues for the subunits of the eIF4F complex were identified in dinoflagellates and trypanosomatids [13], indicating independent events of genetic

CONTACT Osvaldo Pompílio de-Melo-Neto  osvaldo.pompilio@fiocruz.br  Instituto Aggeu Magalhães, FIOCRUZ-PE, Av. Moraes Rego S/n, Cidade Universitária, Recife-PE, Brazil; Beatriz Gomes Guimarães  beatriz.gulmaraes@fiocruz.br  Instituto Carlos Chagas, FIOCRUZ-PR, Rua Prof. Algacyr Munhoz Mader 3775, Curitiba, PR 81350-010, Brazil

 Supplemental data for this article can be accessed here.

© 2021 Informa UK Limited, trading as Taylor & Francis Group

duplication in each lineage of major functional relevance. In trypanosomatids there are six eIF4E homologues, classified into three groups based on structural and molecular properties (reviewed in [14]). Group 1 is formed by EIF4E1 and EIF4E2, less abundant proteins which do not seem to form eIF4F-like complexes and are probably not required for general translation [15–18]. Group 2, based on EIF4E3 and EIF4E4, comprises the most abundant eIF4Es, forming typical eIF4F complexes that are required for translation and are essential for cell survival [16,19–21]. These eIF4E homologues are atypical as they are characterized by long N-terminal extensions which have been shown, at least for EIF4E4, to directly interact with PABP homologues, a mode of interaction not seen so far in other eukaryotes [21,22]. The third trypanosomatid eIF4E group is based on EIF4E5 and EIF4E6, both small proteins capable of binding to eIF4G homologues and forming eIF4F-like complexes with yet undefined roles in translation [14,23,24].

eIF4Es of most eukaryotic species have a conserved core of 160–170 amino acids characterized by the presence of eight tryptophan residues generally conserved (or mostly replaced by other aromatic residues) [12]. Three of those, W56, W102 and W166 (human eIF4E-1 numbering), are involved in cap recognition, while W73 is involved in the interaction with eIF4G [25]. The mammalian eIF4E is regulated by the eIF4E-binding proteins, a family of repressor proteins, named 4E-BP1, 4E-BP2, and 4E-BP3 [26,27]. The consensus eIF4E binding motif is represented by YX₄LΦ (where X is any amino acid and Φ is a hydrophobic amino acid residue) and is found in both mammalian eIF4G and 4E-BPs. As these factors bind to eIF4E at the same site, their mutual binding is not observed [28].

The three-dimensional structure of the so-called eIF4E domain is highly conserved, consisting of a cupped-hand structure formed by seven beta-strands arranged to form a pocket for the cap interaction, with three flanking α-helices forming the interface regions with eIF4E-interacting proteins [29]. Structural studies of eIF4E from several organisms, including human, revealed details of the cap-binding mechanism [30–32] and the interaction of eIF4G and 4E-BP at the convex side of eIF4E, as well as the motifs modulating these interactions to regulate translation initiation [33–37]. To date, the crystal structures of two trypanosomatid eIF4E homologues were reported. The *L. major* EIF4E1 structure was described in the apo form [38], and, in a recent work, we have determined the 3D structure of *Trypanosoma cruzi* EIF4E5 in complex with the mRNA cap-4 and other cap analogues. Structural comparisons showed significant differences in the cap-binding pocket, providing evidences of distinct cap-binding mechanisms between the trypanosomatid eIF4Es [39].

Both the initial description and the structural characterization of the trypanosomatid EIF4E5 has been carried out from *Trypanosoma* species [24,39]. Although major functional and structural features are predicted to be conserved among trypanosomatids, relevant differences might be expected between the *Trypanosoma* and *Leishmania* species due to the substantial period of divergence of

their corresponding lineages, greater than 150 million years [40]. In this work, immunoprecipitation assays combined with mass-spectrometry allowed us to identify EIF4E5 interacting proteins in *Leishmania* and to investigate the role of specific residues in these interactions. Moreover, we have determined the crystal structure of *Leishmania* EIF4E5 in the apo form. Our structural analysis indicates significant conformational changes in EIF4E5 upon cap binding and provide insights into molecular specificities that may be implicated in the association of trypanosomatid EIF4E5 with known and novel partners.

Results

Leishmania EIF4E5 share conserved residues with their *Trypanosoma* EIF4E5 orthologues but are characterized by unique C-terminal ends

The trypanosomatid EIF4E5 was first described in *T. brucei* as one of the smallest eIF4Es identified so far among all major eukaryotic groups, including yeasts, metazoans and trypanosomatid eIF4Es, with the sole exception being the related EIF4E6. Nevertheless, it contains the full core domain typical of the eIF4E super family, including motifs involved in cap binding and required for the interaction with eIF4G partners [24]. EIF4E5 orthologues from various trypanosomatid species have only four out of the eight tryptophan residues typically conserved in other eukaryotic eIF4Es, but these include the equivalents to W56 and W166 from human eIF4E-1, which in this case are required for the binding of the cap structure, and W73, involved in the interaction with eIF4G [33,41,42]. The W102 residue from human eIF4E-1, also required for cap binding, is replaced by another aromatic amino acid, a tyrosine, in trypanosomatid EIF4E5. Figure 1 shows a comparison of the sequences of EIF4E5 orthologues from selected trypanosomatids, including *T. brucei*, *T. cruzi* and various *Leishmania* species. Noteworthy from this alignment is the identification of an additional conserved tryptophan residue within the *Leishmania* EIF4E5 (W45 in *L. infantum* and W50 in *L. major*), which is replaced by an aromatic residue, a tyrosine, in *Trypanosoma* orthologues and a lysine residue in human eIF4E-1. Further differences between the trypanosomatid EIF4E5 and the human eIF4E homologue will be discussed in the next sessions.

To identify differences of possible functional relevance between *Leishmania* and *Trypanosoma* EIF4E5, we first evaluated overall identities and similarities between selected orthologues. The *L. infantum* EIF4E5 shares 98% amino acid identity with its *L. major* orthologue but less than 50% identity with the *T. brucei* and *T. cruzi* orthologues (44 and 47%, respectively). This conservation is similar to that observed for other trypanosomatid eIF4E homologues, such as EIF4E1 and EIF4E4 (97% identities between *L. major* and *L. infantum*, with ~50% or less identity between *Leishmania* and *Trypanosoma* orthologues). For EIF4E5, a noteworthy difference in length and amino acids composition is observed on

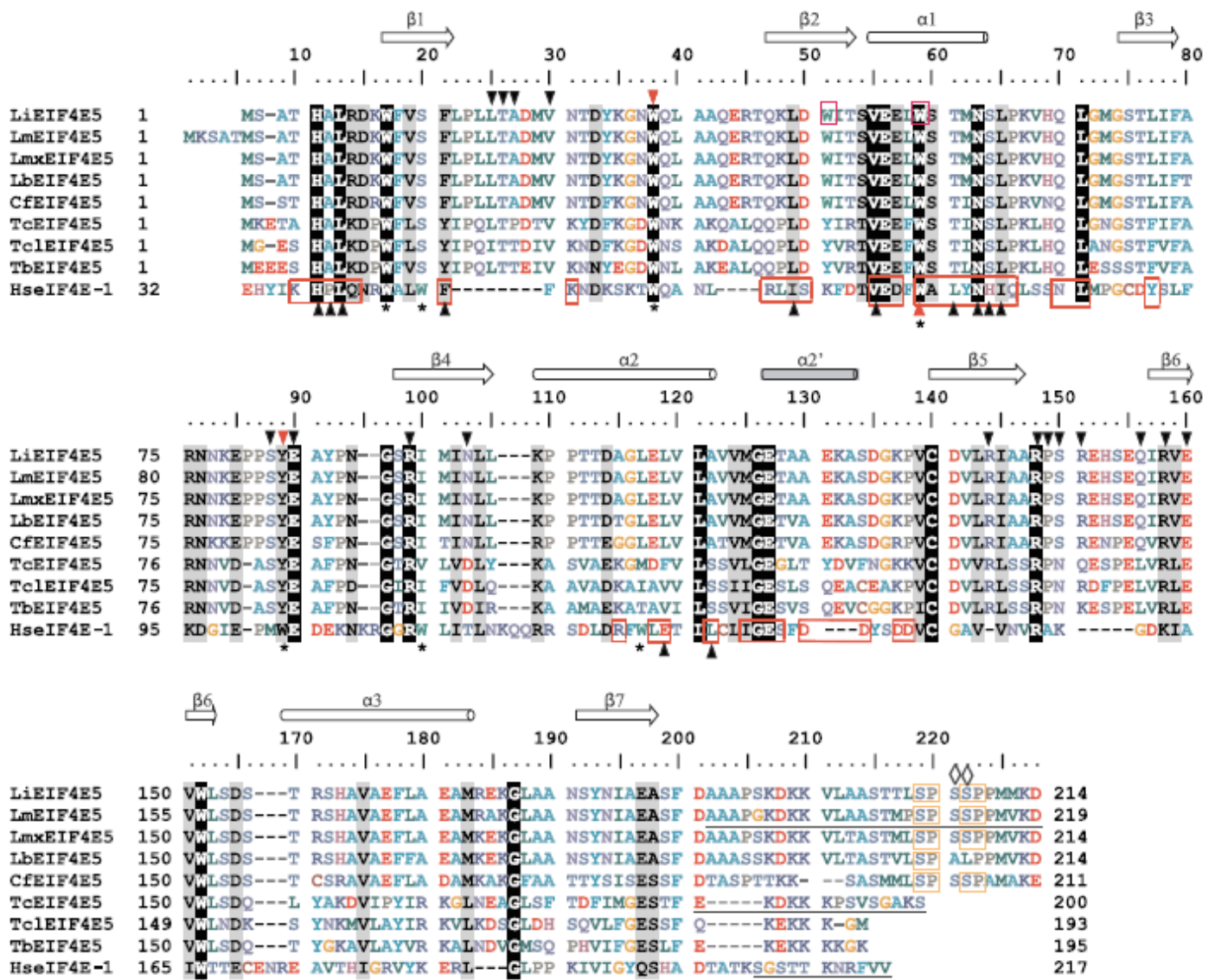


Figure 1. Comparison of EIF4E5 orthologues from different trypanosomatids with *H. sapiens* eIF4E-1. The alignment compares EIF4E5 sequences from *Leishmania infantum* (Li), *L. major* (Lm), *L. mexicana* (Lmx), *L. braziliensis* (Lb), *Crithidia fasciculata* (Cf), *Trypanosoma cruzi* (Tc), *T. congolense* (Tcl) and *T. brucei* (Tb) with the human eIF4E-1 (Hs). Amino acids identical in 100% of the sequences are indicated by the white colour and black background and those similar in 100% of the sequences are highlighted by the grey background. Asterisks represent tryptophan residues generally conserved within the eIF4E protein family. The symbol ▼ indicates the residues involved in cap binding in *T. cruzi* EIF4E5 [39]; in red are highlighted the residues corresponding to human eIF4E-1 W56 and W102 involved in the classical stacking interaction with the m⁷G cap moiety [29]. The ▲ indicates the residues of human eIF4E which participate in the interaction with the eIF4G peptide [36] with the W73 highlighted in red and additional interfacial residues indicated by red boxes. *L. infantum* tryptophans W45 and W53, mutated in this work, are highlighted with pink rectangles. The SP motifs found within the C-terminus of the *Leishmania* and *Crithidia* sequences are indicated by brown boxes, while the validated phosphorylation sites for the *L. infantum* EIF4E5 are marked with the symbol ◊ on top. The secondary structure elements, according to the *L. major* EIF4E5 structure determined in this work, are indicated above the sequences, with the additional α -helix connecting the α 2 helix and the β 5 strand (named α 2') highlighted in grey. The sequences underlined at the C-terminal ends of the *L. major*, *T. cruzi* and human sequences correspond to the regions which are not visible in the 3D structures.

the proteins' C-terminal regions (Figure 1), with the *Leishmania* and *Crithidia* orthologues being roughly 20 amino acids longer than those from the *Trypanosoma* species.

EIF4E5 is constitutively expressed during the growth of *L. infantum* promastigotes

To investigate functional roles for the *Leishmania* EIF4E5 and compare with the data derived from *T. brucei*, we started by generating transgenic *L. infantum* cell lines ectopically expressing a C-terminally HA-tagged form of the protein, as previously done with *L. infantum* EIF4E4 [21].

After stable transfection of *L. infantum* promastigotes, the expression of the tagged protein was evaluated through Western blots with whole protein extracts from different stages of promastigote growth curves. As shown in Figure 2(a), both during exponential growth (up to 48 hours) and stationary phases (~96 hours) the same expression profile is observed for the LiEIF4E5-HA, indicating a constitutive pattern of expression for the parasite's promastigote form. In all samples analysed, the major EIF4E5 band is always accompanied by a second minor isoform of slightly larger molecular weight, whose abundance does not change along the growth culture.

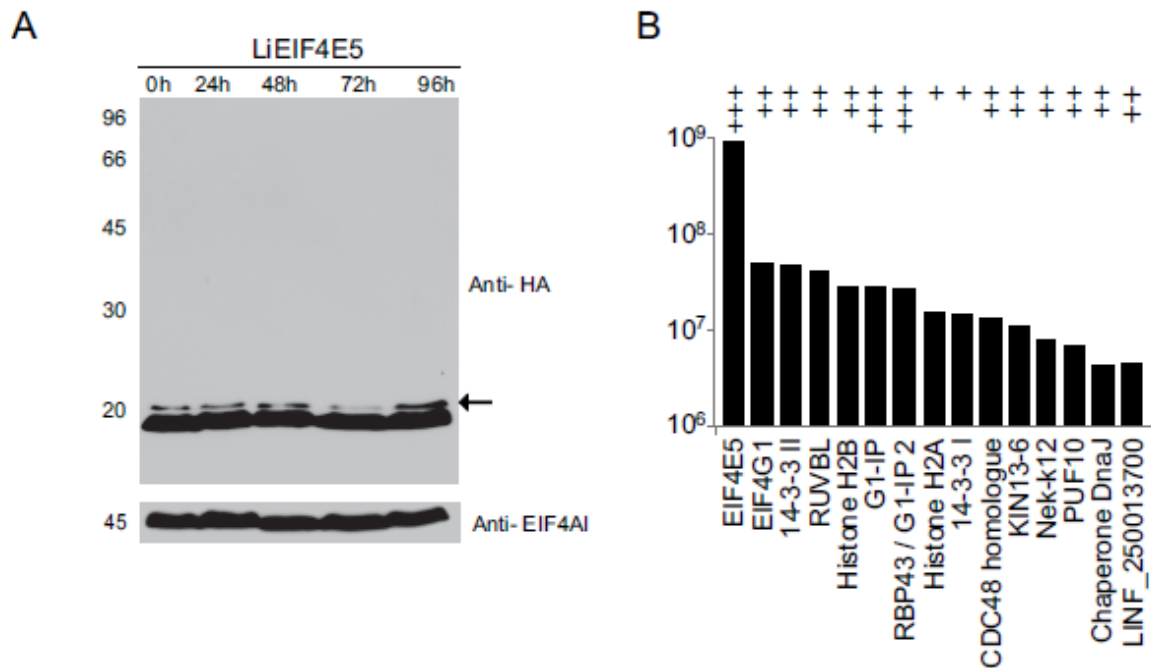


Figure 2. Expression analysis and identification of top-ranking proteins co-precipitating with the *Leishmania infantum* HA-tagged EIF4E5. (A) Western blotting showing the expression of HA-tagged *Leishmania* EIF4E5. After transfection and stable selection, expression of *LiEIF4E5*-HA protein was evaluated at different promastigote growth stages, including exponential (up to 48 hours) and stationary (~96 hours) phases, by western blotting using an anti-HA monoclonal antibody. In all analyses, the anti-EIF4AI antibody was used as a gel loading control. The numbers on the left indicate the sizes in kDa of molecular weight markers and the arrow on the right highlights the minor EIF4E5 isoform. (B) Bar chart with the top-most proteins co-precipitated with *LiEIF4E5*-HA, selected according to the arbitrarily defined 'strength' category, as defined in the Supplementary Table 2 and also shown in Table 1 (all those classified as '+++' and '++' plus the two proteins classified as '+' with intensity values greater than 1×10^7).

EIF4G1, interacting partners and cell-cycle related proteins co-precipitate with the HA-tagged *L. infantum* EIF4E5

The *T. brucei* EIF4E5 has been shown to co-precipitate with EIF4G1 and EIF4G2, two members of the trypanosomatid eIF4G family. Mass-spectrometry data from both EIF4G1 and EIF4G2 associated complexes revealed the presence of *TbEIF4E5* in two independent eIF4F-like subcomplexes. The first subcomplex was formed by EIF4E5, EIF4G1, a protein belonging to the 14-3-3 family (14-3-3 I) and two additional proteins, a mRNA cap guanine-N7 methyltransferase and the RNA Binding Protein 43 (RBP43), also named as *TbG1-IP* (EIF4G1-IP) and *TbG1-IP2* (EIF4G1-IP2), respectively. A second subcomplex was formed by EIF4E5, EIF4G2, two 14-3-3 paralogues (14-3-3 I and II) and another uncharacterized protein that was previously named *Tb-G2-IP* (EIF4G2-IP) [14,24]. Here, a search to identify the *Leishmania* EIF4E5-interacting partners was carried out. First, single knock out (SKO) cell lines were generated by replacing a single *EIF4E5* gene copy with a hygromycin resistance marker, prior to transfection with the plasmid encoding the HA-tagged EIF4E5. Immunoprecipitations (IPs) were then carried out using whole cytoplasmic lysates from *L. infantum* promastigotes expressing *LiEIF4E5*-HA in the context of the SKO background. Whole cytoplasmic lysates of cells lacking the HA-tagged protein were used as negative controls. The IPs from two sets of replicates were then submitted to mass-spectrometry analysis. Proteins specifically co-precipitating with *LiEIF4E5*-HA were identified based on normalized

intensity values and enrichment ratios for the *LiEIF4E5*-HA in comparison with the control, using the averages from the replicates. The whole set of proteins with a minimum enrichment ratio of 1.5 and at least two unique peptides in both replicates with the tagged protein were selected for further analysis and are listed in the Supplementary Table S1, while Table 1 and Supplementary Table S2 list only those proteins which were found with an enrichment ratio greater than 4.0 plus having average intensity values greater than 1/1000 of the value observed for the *LiEIF4E5*-HA bait. These were considered more likely to reflect critical EIF4E5 interactions, with the uncharacterized proteins thus omitted from Table 1. Fifteen proteins having the highest combination of enrichment ratios and intensity values are also shown in Figure 2 (b), for a comparative analysis based only on their intensities, and will be discussed in more detail.

Apart from the HA-tagged *LiEIF4E5* bait, the topmost protein co-precipitated was EIF4G1, with the *L. infantum* EIF4G1-IP1 and RBP43/EIF4G1-IP2 orthologues also found amongst the highest-ranking proteins (Table 1 and Figure 2(b)). Likewise, both 14-3-3 paralogues were also found within the most relevant co-precipitated proteins. These results indicate the conservation in *Leishmania* of one of the EIF4E5 sub-complexes previously identified in *T. brucei*. Histones were also found within the topmost proteins co-precipitated with *LiEIF4E5* (Figure 2(b)), as well as the RuvB-like DNA helicase involved in chromatin remodelling. At this stage it is not clear how these nuclear proteins would be associated with the cytoplasmic EIF4E5 (based on the localization of its *T. brucei* orthologue [24]),

Table 1. Proteins specifically co-precipitating with the HA-tagged *Leishmania* EIF4E5.

Name	Description	Intensity	Ratio	Strength
EIF4E5 – BAIT	Eukaryotic translation Initiation factor 4E type 5	9x10 ⁸	1454.05	+++
Translation initiation factors				
EIF4G1	Eukaryotic translation Initiation factor 4 G type 1	5x10 ⁷	8.26	++
RNA binding and RNA modifying proteins				
G1-IP	mRNA cap guanine-N7 methyltransferase	2.7x10 ⁷	15.09	+++
G1-IP2, RBP43	RNA-binding protein 43	2.6x10 ⁷	17.46	+++
nRBP	Nucleolar RNA-binding protein	8.0x10 ⁶	4.24	-
PUF10	Pumilio 10	6.7x10 ⁶	9.89	++
RBP29	RNA-binding protein 29	1.2x10 ⁶	17.09	-
Protein kinases and related				
14-3-3 II	14-3-3 protein 2	4.7x10 ⁷	4.08	++
14-3-3 I	14-3-3 protein 1	1.4x10 ⁷	8.07	+
NEK-12	Serine/threonine-protein kinase	7.7x10 ⁶	12.12	++
PP2A	Protein phosphatase 2A, regulatory subunit	5.7x10 ⁶	6.06	-
PKA	Protein kinase A, regulatory subunit	2.9x10 ⁶	6.82	-
MKK5	Mitogen-activated protein kinase 5	9x10 ⁵	∞	-
Histones				
H2B	Histone H2B	2.8x10 ⁷	5.85	++
H2a	Histone H2a	1.5x10 ⁷	7.79	+
Enzymes				
RuvB	RuvB-like DNA helicase	3.9x10 ⁷	4.36	++
PPase	Acidocalcisomal pyrophosphatase	4.2x10 ⁶	4.50	-
TK	Thymidine kinase	2.3x10 ⁶	4.02	-
PPME1	Protein phosphatase methyltransferase 1	1.3x10 ⁶	4.95	-
Intracellular transport and cell motility				
CDC48 homologue	Vesicular transport protein, CDC48 homologue	1.3x10 ⁷	15.95	++
KIN-13 6	Kinesin-13 6	1.1x10 ⁷	34.11	++
RabX1	Ras-related protein RabX1	8.4x10 ⁶	4.61	-
Other proteins				
ARL3C	ADP-ribosylation factor protein	9.7x10 ⁶	6.50	+
DnaJ	Chaperone protein DnaJ	4.6x10 ⁶	9.93	++
RFA3	Replication factor A protein 3	3.8x10 ⁶	4.68	-
TWF	Twinfilin	2.8x10 ⁶	10.54	+
WD40 repeat	WD domain-containing protein, G-beta repeat	2.5x10 ⁶	∞	+
NBP35	Nucleotide-binding protein 35	1.6x10 ⁶	7.44	-
PSP1, C-terminal	PSP1 C-terminal domain-containing protein	1.6x10 ⁶	4.65	-

The polypeptides listed here are also detailed on the Supplementary Tables S1 and S2 and includes only identified proteins which fit into the following parameters: two or more peptides found for each EIF4E5 replicate; a minimum of four-fold (4x) enrichment over the negative control ('Ratio'); a minimum average intensity ('Intensity') 1000-fold (1000x) less than that seen for the *Leishmania* EIF4E5-HA bait. The last column arbitrarily defines the 'Strength' of the interaction with the bait considering both the average intensity and the enrichment in comparison with the negative control, according with criteria defined in the Supplementary Table S2.

although some degree of contamination of our cytoplasmic extracts with nuclear proteins is generally seen. EIF4E5 also co-precipitated with the *Leishmania* NEK-12 kinase as well as other protein kinases whose orthologues are linked to cell-cycle regulation in *Trypanosoma* and *L. mexicana* (Table 1). Our IP assays also identified proteins related with intracellular transport and cell motility among the topmost proteins co-precipitating with EIF4E5 (KIN13-6 and CDC48 in Figure 2(b)) with one additional related protein found in Table 1 (RabX1). Additionally, the single uncharacterized protein appearing in the top list of EIF4E5 interacting proteins (LINF_250013700 in Figure 2(b)) has domains indicative of an association with the cell cytoskeleton. Other co-precipitated proteins include PUF10, an RNA binding protein (RBP) member of the Pumilio family (Figure 2(b)), and two other RBPs (Table 1). Finally, the relevance of the association of EIF4E5 with the chaperone DNAJ is unclear. Overall, this analysis highlights a broader panel of proteins co-precipitating with the tagged *Leishmania* EIF4E5, when compared to those previously reported for its *T. brucei* orthologue, and better define several which should be further investigated in order to clarify the EIF4E5 mode of action.

Further analysis of the proteins co-precipitating with EIF4E5

A number of proteins found to co-precipitate with EIF4E5, but with lower enrichment ratios or intensity values, might also reflect functional relevance. To consider this, we evaluated the more comprehensive list of co-precipitated proteins (Supplementary Table S1) and searched for proteins which might be functionally related to EIF4E5 (summarized in Figure 3(a)). We first investigated other eIF4F-related subunits and EIF4G2 was indeed found in the full list of co-precipitated proteins, although associated with enrichment ratio and intensity values more similar to EIF4G4 and its partner EIF4E3, both eIF4F subunits so far found to be unrelated to EIF4E5. More relevant was the identification of a significant number of ribosomal proteins, normally associated with translationally active mRNAs. Likewise, co-precipitation with EIF3B, one of the many subunits of the eIF3 complex, also reinforces an association with the translation machinery (a second eIF3 subunit, EIF3G, was also found with an enrichment ratio greater than two, but it was omitted from the Figure and Supplementary Table since it was represented by only one unique peptide in one of the replicates). Two of the *Leishmania* PABP homologues (PABP1 and

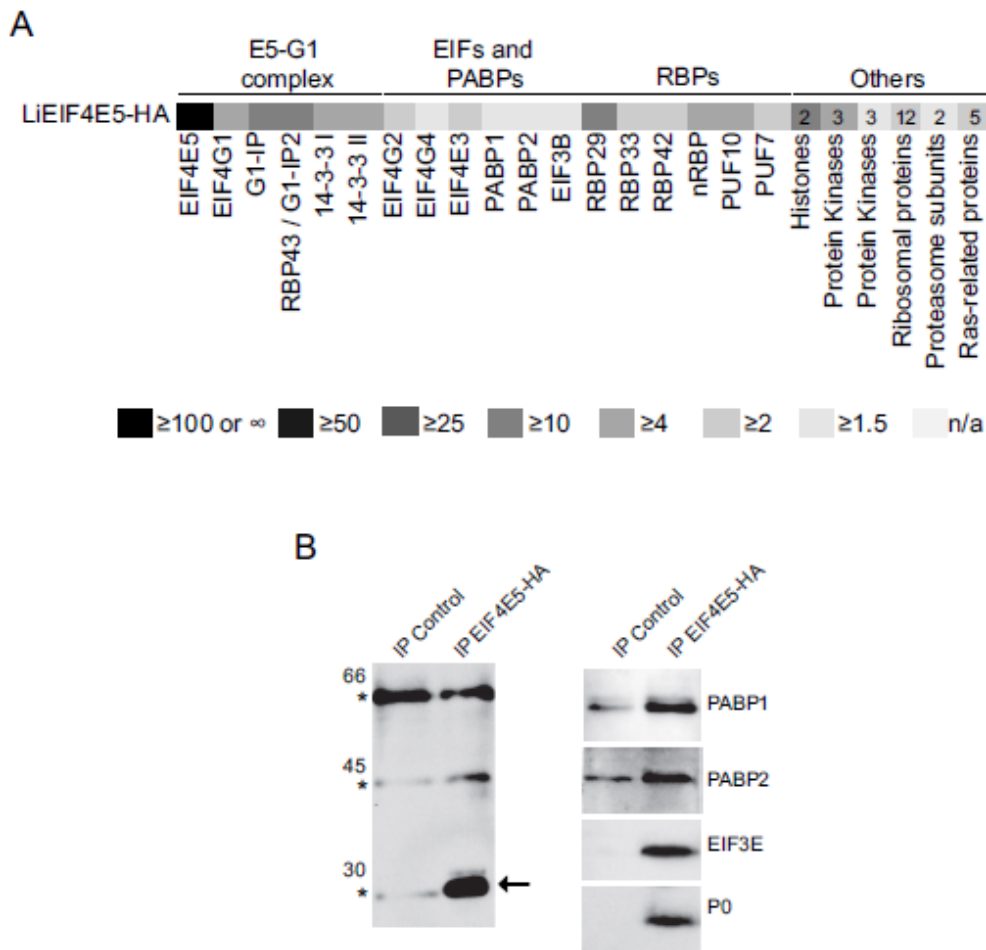


Figure 3. Detailed analysis of selected proteins co-precipitating with *LiEIF4E5-HA*. (A) Average enrichment ratios for selected proteins co-precipitated with *LiEIF4E5-HA* (based on the comprehensive list from Supplementary Table 1) and considered functionally relevant for the understanding of its function. The numbers within the squares indicate several related proteins grouped together. (B) Validation of the co-precipitation of selected proteins with the HA-tagged EIF4E5 by western blotting. The left panel shows the blot with IPs from the *LiEIF4E5-HA* sample and the negative control assayed with the anti-HA antibody. The stars indicate the IgG bands with the sizes of the molecular weight markers also shown and the EIF4E5-HA band indicated by the arrow. The panels on the right show the same samples blotted with selected antisera, as indicated.

precipitated with EIF4E5-HA within the expanded list from Figure 3(a).

With the availability of previously described antisera against *Leishmania* orthologues for the two PABP homologues identified [43], the EIF3E subunit of the eIF3 [44] and the protein P0 [45], constituent of the 60S ribosomal subunit, we decided to use those antisera to independently validate the co-precipitation assays. A new set of IPs were then carried out with cytoplasmic extracts from both *LiEIF4E5-HA* expressing cells and corresponding negative controls, but now with cells having both endogenous copies of the wild-type *LiEIF4E5* gene. The IPs were followed by western blotting using these antisera. As shown in Figure 3 (b), the two PABP homologues were found both with the negative control as well with the tagged *LiEIF4E5*, but with a stronger signal for the latter, consistent with both PABPs binding non-specifically to the beads to some extent but being enriched in the presence of *LiEIF4E5*. In contrast, both EIF4E3 and the ribosomal P0 were only seen in the presence of the tagged protein. HA-tagged versions of two eIF3 subunits, EIF3E and EIF3G, were also expressed in *L. infantum* and used for IPs and mass-spectrometry

identification of the co-precipitated proteins. EIF4E5 was found to co-precipitate with both eIF3 subunits, with EIF3G also found to co-precipitate with EIF4G1 and related protein partners. In contrast, EIF4E3 was the only other *Leishmania* eIF4E found co-precipitating with these two eIF3 subunits (Supplementary Figure). These results confirm the association of EIF4E5 with the translation machinery and validate the identification of the proteins co-precipitated with *LiEIF4E5-HA*.

Assessing the role of selected tryptophan residues in mediating the EIF4E5 interactions

The role of the W73 residue in mediating the mammalian eIF4E-1 interaction with eIF4G homologues and eIF4E-BPs was originally demonstrated both *in vitro* and *in vivo* [33,42]. In yeast eIF4E the equivalent residue was also found to be required for its interaction with yeast eIF4G and the p20 eIF4E binding protein [41]. The equivalent residues have likewise been shown to be required for the binding of the *Leishmania* EIF4E3 and EIF4E4 to their corresponding eIF4G partners, respectively EIF4G4 and

EIF4G3, since their replacement with alanines abolished these interactions [20–22]. The conservation of a tryptophan residue in an equivalent position in the various trypanosomatid EIF4E5 orthologues (Figure 1) indicates a likely role in binding to eIF4G partners and possibly to other co-precipitated proteins. In addition to the conserved W53 (*L. infantum* numbering), the presence of a nearby tryptophan residue conserved in EIF4E5 from *Leishmania* species (W45) raised the question of whether this tryptophan would be also involved in relevant interactions with known or novel protein partners. To investigate these possibilities, we generated two EIF4E5-HA tagged mutants (Trp replaced by Ala) which were transfected into *L. infantum* EIF4E5-SKO promastigotes. Expression of the two mutants was evaluated as for the wild-type *LiEIF4E5*-HA, with both ectopically expressed proteins having identical profiles to the one seen for the wild-type protein, including the presence of a larger molecular weight isoform indicative of post-translational modifications (Figure 4(a)).

The two mutant proteins were immunoprecipitated in two sets of experiments carried out in parallel with the original *LiEIF4E5*-HA, with the co-precipitated proteins also analysed by mass-spectrometry. After two cycles of normalization considering both the overall IP-yield and the intensity of the *LiEIF4E5*-HA bait, the three sets of data were compared aiming to identify proteins whose co-precipitation with EIF4E5 was altered by either of the two mutations (EIF4E5^{W53A} or EIF4E5^{W45A}). We specifically looked for proteins found to be enriched with *LiEIF4E5*-HA in comparison with the negative control, in the

experiments discussed previously (based on the list from Supplementary Table S1), but whose co-precipitation was reduced at least two-fold by either mutation (listed in Table 2 with the data detailed in the Supplementary Tables S3 and S4, respectively, for the EIF4E5^{W53A} and EIF4E5^{W45A} mutants). Remarkably, no noticeable effect of the W53A mutation was observed on the association of the top-ranked proteins identified in the wild-type *LiEIF4E5*-HA IP assay, including EIF4G1, or the other protein partners associated with the EIF4G1 subcomplex. Nonetheless, the W53A mutation has an impact on the association of RBP33; two members of the PUF family (PUF10 and PUF7); several other enzymes and uncharacterized proteins. Overall, most of those affected proteins were already associated with reduced intensity levels ($1-3 \times 10^6$) in the wild-type *LiEIF4E5*-HA and the reduction seen in the mutant is generally less than four-fold. An exception is RBP33, also known as cell-cycle sequence binding phosphoprotein, with intensity values of $\sim 5 \times 10^6$ and more than five-fold reduction in co-precipitation with the mutant protein. As for the EIF4E5^{W45A} mutant, although the mutation did not lead to changes in the co-precipitation of EIF4G1 or other eIF4G homologues, it did lead to a two-fold reduction of RBP33 (G1-IP2). More noteworthy, however, is the major (thirty-fold) drop in intensity seen for the serine/threonine protein phosphatase 2A regulatory subunit (PP2A).

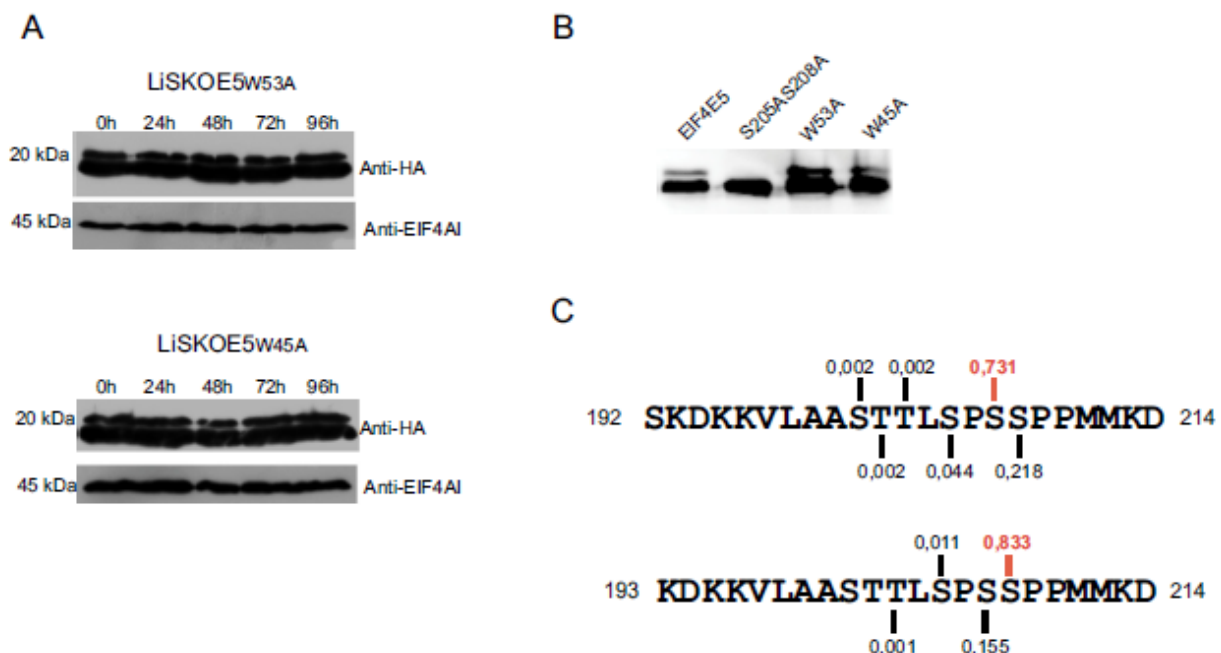


Figure 4. Expression analysis of mutated *LiEIF4E5*-HA and identification of phosphorylated peptides. (A) Western blotting showing the expression of the HA-tagged *LiEIF4E5* mutant proteins. After transfection into *L. infantum* EIF4E5-SKO promastigotes, expression of W53A-HA and W45A-HA EIF4E5 protein was evaluated at different promastigote growth phases as described for the previous figure. (B) Western blotting comparing the expression of the tagged proteins shown in (A) with that of the S205AS208A mutant. (C) Phosphopeptides identified after mass-spectrometry analysis of the HA-tagged EIF4E5. The figure summarizes the data from a representative experiment with the two phosphopeptides identified, as well as the most likely phosphorylated residues (in red), defined according to the p-values shown for each residue. The p-values are indicated by vertical bars.

Table 2. Proteins specifically co-precipitated with the HA-tagged *L*EIF4E5 but impacted by either the W53A or W45A mutations.

Name	Description	Ratio*
RBP33	EIF4E5 _{W53A}	6.75
	Cell cycle sequence binding phosphoprotein	3.63
KHC1	Uncharacterized protein	3.61
	Kinesin – putative	2.94
	PSP1 C-terminal conserved region containing protein – putative	2.75
PUF7	Pumilio/PUF RNA binding protein 7 – putative	2.66
	AP-1 adapter complex gamma subunit – putative	2.63
PUF10	Uncharacterized protein	2.56
	Pumilio protein 10 – putative	2.53
ASAT	Aspartate aminotransferase – putative	2.31
	Protein kinase – putative	2.11
LPG3	Lipophosphoglycan biosynthetic protein/heat shock protein 90 – putative	2.07
	Cytoplasmic translation machinery associated protein – putative	2.03
PP2A	EIF4E5 _{W45A}	30.5
	Protein phosphatase 2A, regulatory subunit	2.5
G1-IP2	RNA-binding protein 43 – putative	
RBP43		

The proteins listed (fully detailed on the Supplementary Tables S3 and S4) are those found to be enriched with the wild-type protein in comparison with the negative control (from Supplementary Table S1) and whose co-precipitation with the EIF4E5_{W53A} or EIF4E5_{W45A} mutants is reduced by a minimum of two-fold. * Ratio calculated between the wild-type *L*EIF4E5 and the mutant proteins based on the averages of the Intensity values for the two replicates for both the wild-type and mutants.

The unique C-terminal end of the *Leishmania infantum* EIF4E5 is a target for post-translational modifications

The trypanosomatid EIF4E3 and EIF4E4 homologues have been shown to be targeted by multiple phosphorylation events in both *T. brucei* and *Leishmania* species. These have so far been shown to be directed at their unique N-terminal regions, although many of the phosphorylation motifs are not conserved across the species [21,46,47]. In contrast, *T. brucei* EIF4E5 and EIF4E6 have been shown to be targets for lysine acetylation [48]. Here, we opted to investigate whether the larger molecular weight isoform seen for the HA-tagged EIF4E5 was also a product of phosphorylation. In *L. infantum*, phosphorylation of both EIF4E4 as well as its PABP1 partner was found to target similar serine-proline or threonine-proline (SP/TP) motifs [21,45]. We first considered the possibility that such motifs might also be targeted in EIF4E5 and searched for them in orthologues from various trypanosomatids. As highlighted in Figure 1, two SP motifs were generally found near the C-terminal ends of the *Leishmania* and *Crithidia* EIF4E5 orthologues. We next generated another EIF4E5 mutant having the two serine residues from these motifs replaced by alanines (S205A/S208A). The new mutant was then assayed with the anti-HA antibody after transfection and expression in *L. infantum* promastigotes. When compared with the wild-type EIF4E5 and the previously described mutants, the S205A/S208A EIF4E5 was the single one lacking the extra isoform (Figure 4(b)). Taking advantage of the mass-spectrometry analyses carried out above, we also searched for phosphorylated peptides derived from the HA-

tagged *L*EIF4E5 proteins in order to map phosphorylated sites. Two phosphorylated peptides were indeed found which mapped to the protein's C-terminus, with the most likely phosphorylated residues identified being S207 and S208 (Figure 4(c)). Overall, these results are consistent with the *Leishmania* EIF4E5 being targeted by phosphorylation directed to the C-terminal SP motifs and/or neighbouring serine residues. The lack of similar motifs in the *Trypanosoma* orthologues highlights that the difference between the *Leishmania* and *Trypanosoma* sequences within their C-terminal ends imply relevant functional implications.

The crystal structure of *L. major* EIF4E5 in the apo form indicates structural rearrangements upon mRNA cap binding

To gain more insight into the molecular mechanisms involved in the interaction between *Leishmania* EIF4E5 and its partners, we were able to crystallize the *L. major* EIF4E5 homologue, which shares 98% identity with *L. infantum* EIF4E5. The crystal structure was determined by molecular replacement using a model generated by the Phyre2 server [49] as the search model. Despite the attempts to obtain the protein structure also in the cap-bound form, suitable diffraction data were obtained only from crystals grown in absence of the ligand. Analysis of *L*mEIF4E5-cap association through microscale thermophoresis (data not shown) indicated low affinity of *L*mEIF4E5 for the cap-4, although we were not able to determine the dissociation constant with sufficient accuracy. These results are in agreement with the results from the *T. cruzi* EIF4E5 [39]. The *L*mEIF4E5 apo structure was refined at 2.0 Å resolution to final R_{work}/R_{free} of 20.1/23.0% (Table 3).

Table 3. Crystallographic data and refinement statistics. Values in parentheses are for the highest resolution shell.

Data Statistics	
Source	SOLEIL-PX1
Wavelength (Å)	0.978565
Resolution (Å)	50–2.0 (2.12–2.00)
Space Group	P3 ₂ 21
Unit Cell (Å)	a = b = 61.89 c = 173.03
Number of Observations	470,983 (61,054)
Number of Unique Reflections	26,851 (4162)
Completeness (%)	99.8 (98.6)
Redundancy	17.5 (14.7)
I/σ	9.1 (0.96)
CC ½	99.8 (44.6)
Refinement Statistics	
R _{work}	0.201
R _{free}	0.230
Bond RMSD Length(Å)/Angle (°)	0.008/0.96
Mean B value (protein overall, Å ²)	41.9
Protein atoms	2807
Number of Water Molecules	204
Ramachandran plot	
Favoured (%)	98.0
Outliers (%)	0

The two asymmetric unit monomers comprise residues 9 to 192 (out of 219 residues per monomer and with a single difference between the *L. major* and *L. infantum* sequences), with the polypeptide chains clearly defined by the electron density. The exception is the region connecting helix $\alpha 2$ and strand $\beta 5$, where residues 123 to 131 could not be modelled in monomer B. Asymmetric unit monomers superposition resulted in a Root Mean Square Deviation (RMSD) of 1.25 Å and we used the monomer A for structural analysis.

Superposition of the *Leishmania* EIF4E5 structure with its *T. cruzi* counterpart in complex with the mRNA cap-4 (PDB 6O7Y) [39] resulted in a RMSD of 2.3 Å for 172 C-alpha

atoms aligned, with 43% of sequence identity. Structural comparison between the holo and apo forms evidenced significant rearrangements (Figure 5(a)). The tryptophan residue that participates in the classical stacking interaction with the m7G moiety in *TcEIF4E5* holo structure (W33) moves away from the m7G pocket in *LmEIF4E5* (W37), as a consequence of the rearrangement of the entire region connecting $\beta 1$ - $\beta 2$ strands (S1-S2) (Figure 5(b)). The side chain of the highly conserved glutamic acid (E84 in *TcEIF4E5*, E89 in *LmEIF4E5*), which forms two hydrogen bonds with the cap-4 guanine base in the *TcEIF4E5* structure, moves towards the solvent in the apo structure. Moreover, together with the S1-S2 region, the loop

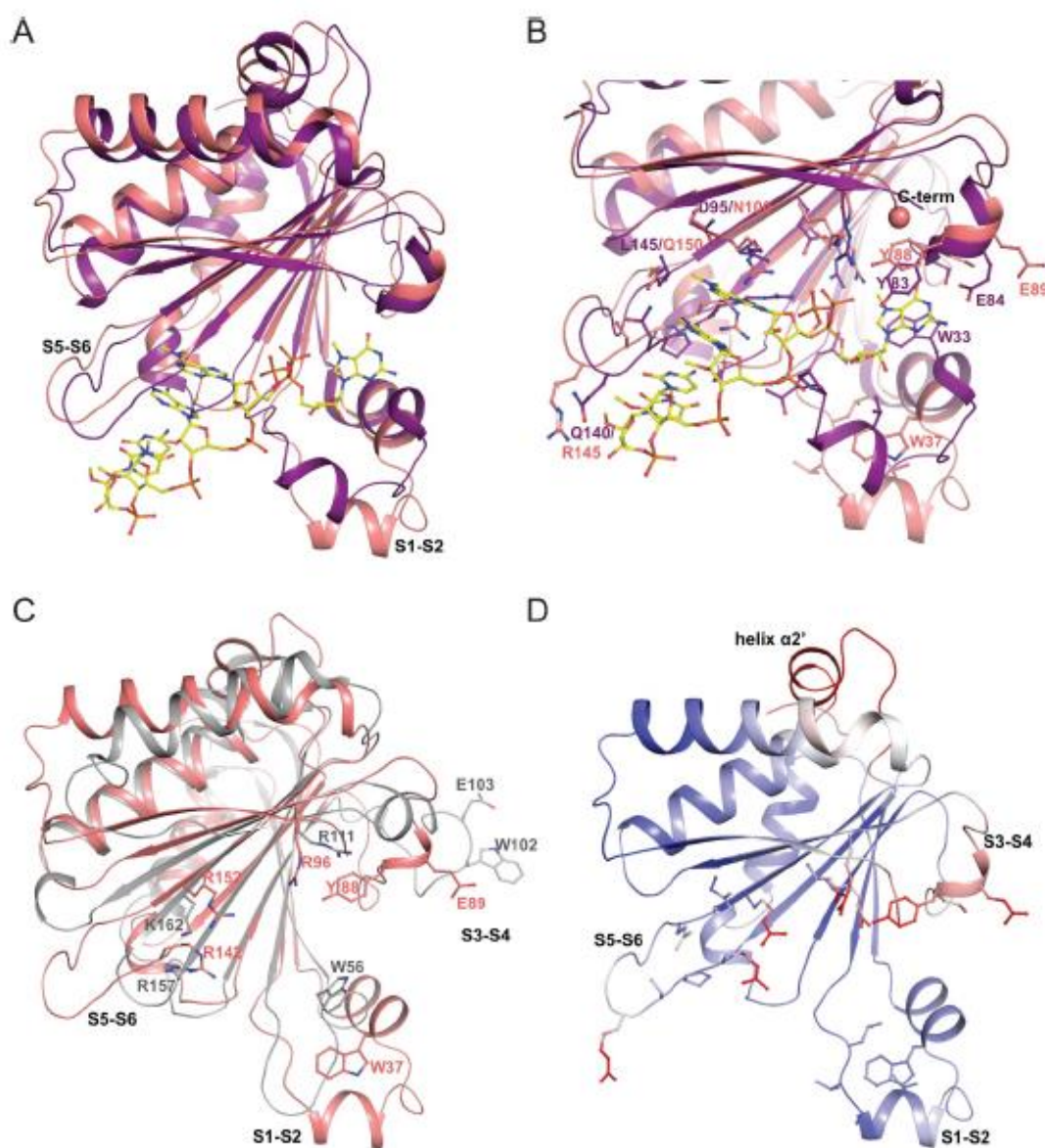


Figure 5. Structural analysis of *L. major* EIF4E5 in the apo form. (A) Overall superposition of *LmEIF4E5* (salmon) and cap-4 bound *TcEIF4E5* (PDB 6O7Y, purple). The superposition evidences structural rearrangements of S1-S2 and S5-S6 regions upon cap-4 binding. The cap-4 is represented in sticks. (B) Detailed analysis of the cap-4 interacting residues in *TcEIF4E5* showing a significant conformational change of W37, Y88 and E89 (*LmEIF4E5* numbering, corresponding to *TcEIF4E5* W33, Y83 and E84) in the absence of the ligand. Non-conservative substitutions of cap-4 interacting residues are also indicated. The last *LmEIF4E5* residue which was modelled in the 3D structure is highlighted by a sphere. (C) Overall superposition of *LmEIF4E5* (salmon) and human eIF4E in the apo form (PDB 2GPQ, grey). The residues of human eIF4E known to participate in the m7G sandwich (W56 and W102) and the basic residues which interacts with cap phosphate groups, together with the corresponding residues in *LmEIF4E5* are represented in sticks and labelled. (D) Analysis of B-factor distribution in the *LmEIF4E* structure. The structure is coloured according to the B-factor values, from blue to red indicating increasing B-factor. The cap interacting residues (by comparison with *TcEIF4E5*) are represented in sticks.

connecting $\beta 5$ - $\beta 6$ strands (S5-S6 loop) adopts a more open conformation in the absence of the cap-4. Although most of the *TcEIF4E5* residues involved in cap-4 binding are conserved in *LmEIF4E5*, three non-conservative substitutions are observed. *TcEIF4E5* D95 and L145 are replaced, respectively by an asparagine (N100) and a glutamine (Q150) in *LmEIF4E5* (Figure 5(b)). However, in *TcEIF4E5* their aliphatic moieties are involved in hydrophobic interactions with the cap-4's first methylated adenine. Therefore, these substitutions may not have an impact in cap-binding. In addition, *TcEIF4E5* Q140 is replaced by an arginine (R145) in *LmEIF4E5*. *LmEIF4E5* R145 could fulfill the same role as *TcEIF4E5* Q140 in cap-binding, by forming a H-bond with the cytosine base and participating in a hydrophobic interaction with the uracil nucleotide. However, we cannot rule out that the R145 basic character may also have an impact on the interaction with the mRNA. As for the *L. major* EIF4E5 C-terminus, the crystal structure determined here lacks the terminal 193 to 219 residues, indicating a flexibility similar to what was observed in human and *T. cruzi* eIF4E homologues (Figure 1, underlined). Interestingly, the last modelled residue (D192) is in the vicinity of the cap-binding site.

A structural comparison of *LmEIF4E5* with the human eIF4E structure in the apo form (Figure 5(c)) revealed differences in the S3-S4 conformation which are expected, considering the proteins' intrinsic flexibilities in the absence of ligands. The $\beta 1$ - $\beta 2$ and $\beta 5$ - $\beta 6$ connections are longer in the eIF4E homologues from trypanosomatids, contributing to additional differences in the cap-binding pocket as previously discussed [39]. However, most of the residues involved in the m7GTP binding in human eIF4E are conserved in *LmEIF4E5* and in other trypanosomatid eIF4E homologues. An analysis of the B-factor distribution along the *LmEIF4E5* structure (Figure 5(d)) also revealed interesting features of the apo form. The highest B-factors are observed in the region connecting the $\alpha 2$ helix and $\beta 5$ strand, which forms an additional α -helix in *LmEIF4E5* (named $\alpha 2'$ in Figures 1 and 5 (d)) when compared with other eIF4E structures. Helix $\alpha 2'$ is not in the vicinity of the cap-binding region and its position is discussed in the next section. The region connecting the $\beta 3$ and $\beta 4$ strands (S3-S4), which includes the tyrosine residue that participates in the m7G sandwich, are also associated with high B-factor values. In contrast, the S1-S2 region, which includes the tryptophan residue that also participates in the stacking interaction, has lower B-factors. Stability of S1-S2 may be partially related to crystal contacts, although its interaction with a neighbour molecule is not extensive.

Structural analysis suggests distinct interaction mechanisms for EIF4E5 binding proteins in trypanosomatids

The structure determination of *L. major* EIF4E5 prompted us to investigate differences in the putative eIF4G interaction surface when compared with other eIF4E homologues. Structural superposition of *LmEIF4E5* with human eIF4E1 [36] and yeast eIF4E [34] in complex with eIF4G peptides (Figure 6) shows that, assuming a similar binding interface,

LmEIF4E5 W58 (equivalent to W53 in *LiEIF4E5*) keeps the conformation of the highly conserved tryptophan (W73 and W75 in human and yeast eIF4E, respectively) and would be in position to interact with the eIF4G canonical helix. In contrast, *LmEIF4E5* W50 (W45 in *LiEIF4E5*) would not participate in the interaction when compared with the human complex (Figure 6(a), right panel). A different scenario is observed in the yeast complex, where F68, structurally aligned with the *LmEIF4E5* W50, is involved in the interaction with eIF4G (Figure 6(b), right panel) through hydrophobic contacts with the C-terminal helix of eIF4G [34].

Analysis of the *LmEIF4E5* electrostatic potential surface and comparison with *T. cruzi* EIF4E5 [39] and human eIF4E1 [36] reveal interesting differences (Figure 7). As well as the distinct charge distribution between EIF4E5 and human eIF4E1 in the cap-binding site, which was already discussed in a previous work [39], superposition of the *T. cruzi* and *L. major* EIF4E5 structures with human eIF4E1 in complex with the eIF4G interacting peptide also evidences a striking difference in charge distribution in the region corresponding to the eIF4G interacting surface (Figure 7, bottom panels). Indeed, the sequence alignment shows that although several of the human eIF4E residues involved in the eIF4G interaction are conserved in the trypanosomatids EIF4E5 (Figure 1), non-conservative substitutions in some positions concur with differences in the electrostatic potential surfaces. Notably, the W73 neighbouring motif in human eIF4E (73-WALY-76) is replaced by 54-WSTI-57 and 58-WSTM-61 in *T. cruzi* and *L. major* EIF4E5, respectively. These substitutions modify the surface surrounding this highly conserved tryptophan residue and most likely have an impact in the EIF4E5 interaction mechanism with protein partners. In addition, the eIF4G non-canonical interacting loop begins after the 'elbow loop' which is anchored to the surface of eIF4E through interactions with conserved eIF4E Asn (Hs N77) and His (Hs H78) [36]. Whereas all trypanosomatid EIF4E5 conserve an asparagine at the same position when compared with *HseIF4E*, the histidine residue is replaced by a serine (see also in Figure 1). This substitution may also contribute to a distinct association mechanism of EIF4E5 with partners when compared with other eIF4E-4 G/4E-BP complexes. Another striking difference is observed in the loop connecting eIF4E helix $\alpha 2$ and strand $\beta 5$. A four-residues insertion is observed in *T. cruzi* and *L. major* EIF4E5 orthologues when compared with *HseIF4E*, and the additional $\alpha 2'$ helix is observed in the *LmEIF4E5* structure (its position is highlighted in Figure 7, bottom-right panel). In human eIF4E, the $\alpha 2$ - $\beta 5$ loop participates in the interface with the N-terminal residues of the eIF4G interacting peptide. Remarkably, the molecular shape and charge distribution of this region is quite distinct among the three proteins, in particular in *LmEIF4E5*. Interestingly, the four-residues insertion within the $\alpha 2$ - $\beta 5$ connection is also observed in EIF4E5 from other trypanosomatids, but its sequence, EKAS in *Leishmania* species, is not conserved in *Trypanosoma* orthologues (Figure 1). In summary, these structural comparisons provide insights into distinct molecular features that may have an impact in the interaction of trypanosomatids EIF4E5 with their binding partners.

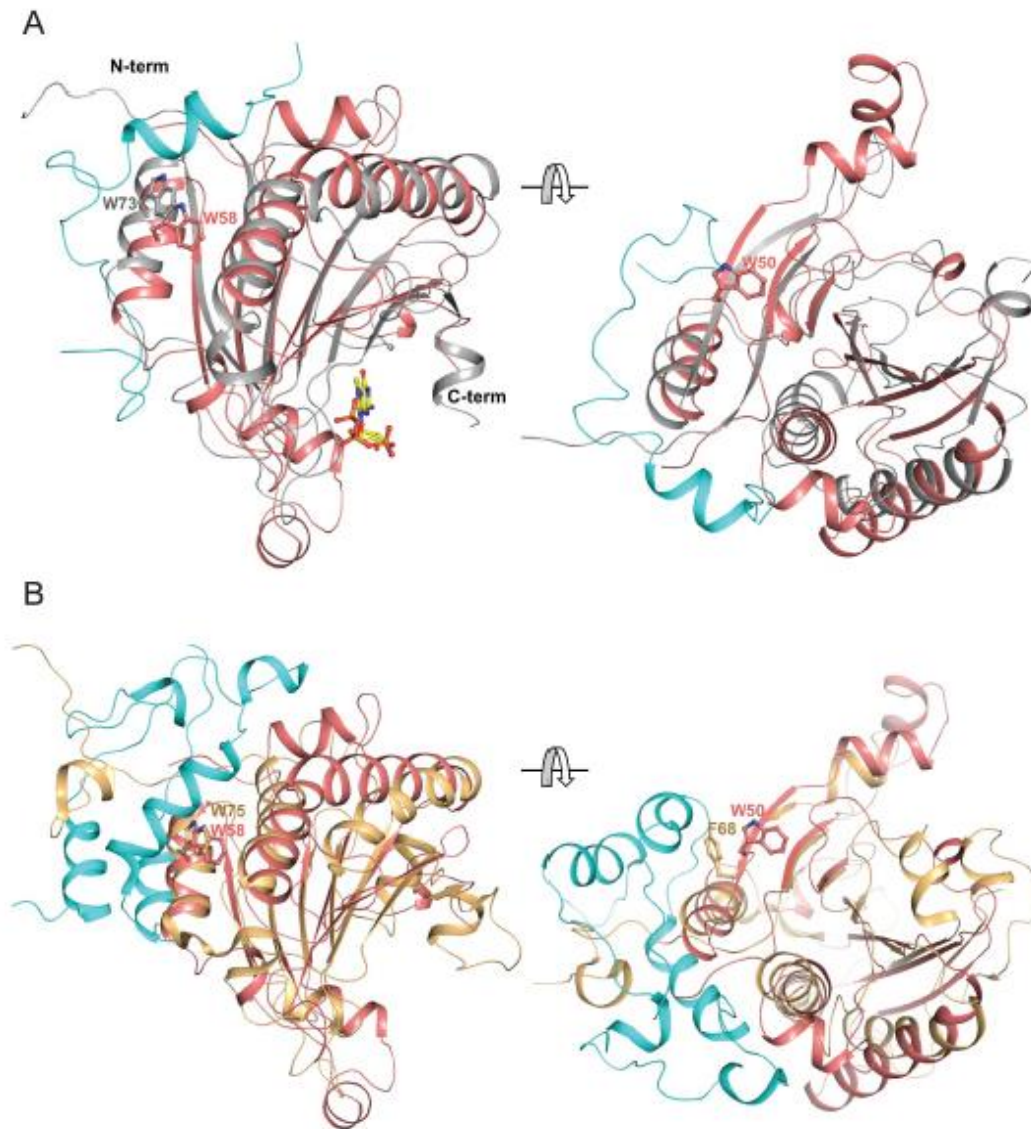


Figure 6. Analysis of the eIF4E-eIF4G interface. (A) Structural superposition of *L. major* EIF4E5 (salmon) and *H. sapiens* eIF4E1 (grey) in complex with an eIF4G peptide (cyan) and m7GTP (sticks) (PDB 5T46). The N- and C-terminal ends, located at the convex and concave sides of eIF4E, respectively, are indicated. On the left, the highly conserved tryptophan (W73 in *HsEIF4E1*) involved in the interaction with the eIF4G canonical helix is shown in sticks. The corresponding W58 in *LmEIF4E5* is also shown. On the right panel, the *LmEIF4E5* W50 is highlighted. The corresponding residue in *HsEIF4E1* (not shown) is not involved in eIF4G interaction. (B) Structural superposition of *LmEIF4E5* (salmon) and *S. cerevisiae* eIF4E (gold) in complex with an eIF4G peptide (cyan) (PDB 1RF8). On the left, *ScEIF4E* W75, which participates in the interaction with the eIF4G canonical helix and the corresponding *LmEIF4E5* W58 are shown in sticks. On the right, *LmEIF4E5* W50 and the corresponding residue in *ScEIF4E* (F68), which participates in eIF4G interaction, are highlighted in sticks. *LmEIF4E5* W58 and W50 corresponds, respectively, to *L. infantum* W53 and W45, mutated in this work.

Discussion

The unique diversity of eIF4E homologues found in trypanosomatids still raises major questions considering their possible roles in these divergent unicellular eukaryotes. At this stage, and based mostly in data from *T. brucei*, EIF4E1 and EIF4E2 are most likely translation regulators, perhaps mediating selective translation repression of target mRNAs with the help of specific RNA binding proteins, such as EIF4E1-IP and SLBP2, respectively [50,51]. In contrast, EIF4E4 and EIF4E3 have been directly or indirectly implicated in translation initiation, a role that seems to require their association with eIF4G partners, respectively EIF4G3 and EIF4G4 (reviewed in [14]). The roles of EIF4E5 and EIF4E6 are so

far less clear, although their association with specific eIF4G homologues does suggest a direct role in translation which needs to be further clarified.

The full-length proteins or fragments of all trypanosomatid eIF4G homologues, as well as EIF4E3 and EIF4E4, have been known for some time to enhance expression of a reporter mRNA when tethered to its 3' UTR [52,53], although only EIF4G3 and EIF4G4, again with their EIF4E4 and EIF4E3 partners, were identified in a search for RNA binding proteins associated with the *T. brucei* polysomes [54]. Rather than reflecting roles other than in translation, the lack of identification of either EIF4E5 or EIF4E6 from both the original tethering assay, as well as the polysomal analyses, might be more related to specific structural features, such as their small

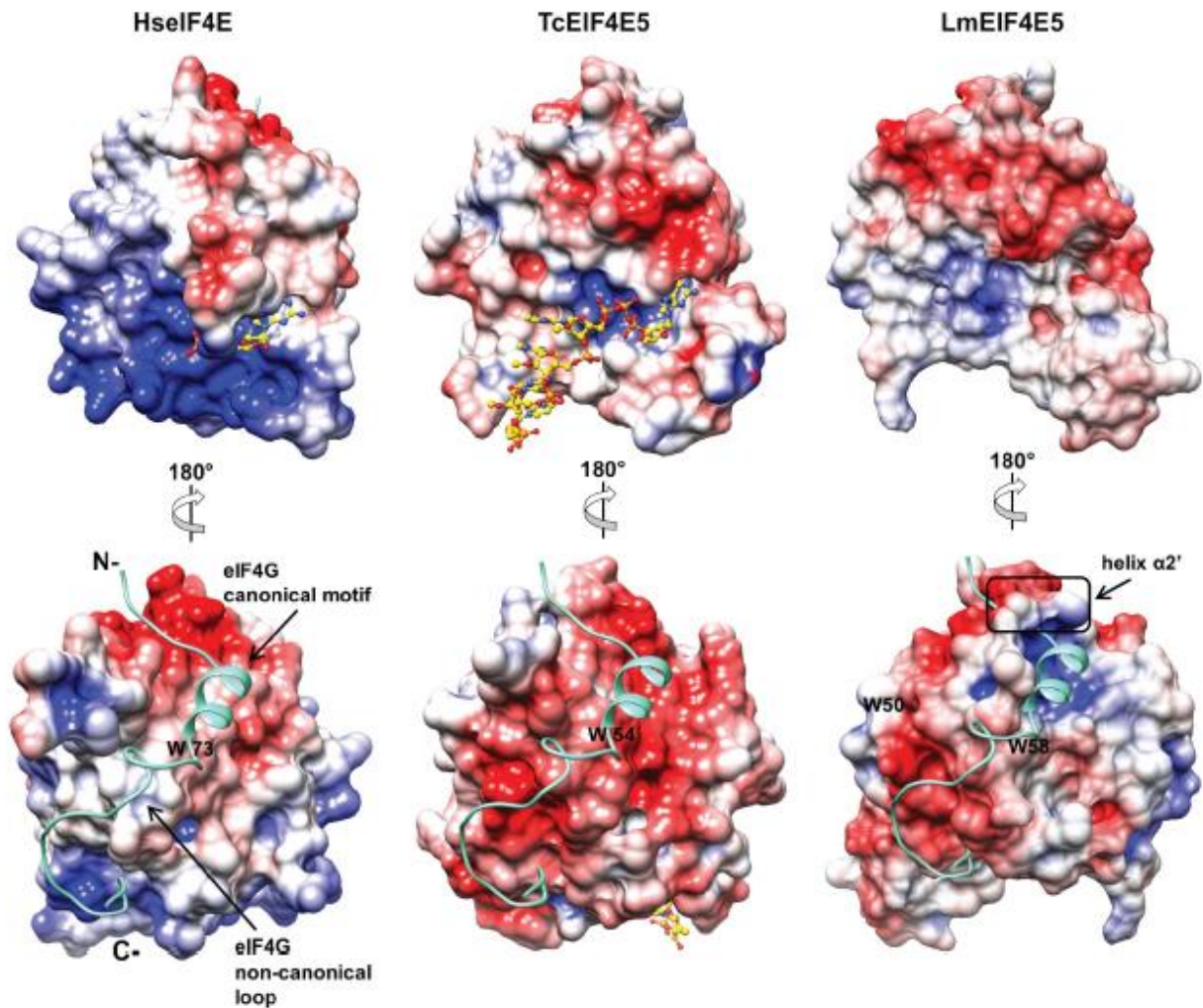


Figure 7. Analysis of trypanosomatids EIF4E5 putative interacting surfaces. Comparison of the electrostatic potential surfaces of *H. sapiens* eIF4E (left, PDB 5T46), *T. cruzi* EIF4E5 (middle, PDB 6O7Y) and *L. major* EIF4E5 (right). The boundaries for potential contour map visualization are ± 5 kT/e. m^7 GTP and cap-4, bound to human eIF4E1 and TcEIF4E5, respectively, are represented in sticks (top panels). The eIF4G peptide bound to the human eIF4E is represented in cyan in the bottom-left panel. The N- and C-terminal ends as well as the canonical helix and non-canonical interacting loop are indicated. Based on structural superposition, the eIF4G peptide from the human complex is positioned in the trypanosomatid EIF4E5 structures, for comparative purposes. The positions of HsEIF4E W73, TcEIF4E5 W54, LmEIF4E5 W58 and W50 are indicated. LmEIF4E5 W58 and W50 corresponds, respectively, to *L. infantum* W53 and W45, mutated in this work. The position of the LmEIF4E5 helix $\alpha 2'$ is indicated in the bottom-right panel.

sizes, or to lower abundances. Indeed, it has been recently reported that EIF4E6 does stimulate expression in a tethering assay and that the EIF4E6/EIF4G5 complex is specifically associated with a number of RNA binding proteins that behave likewise, suggesting all might be directly linked to translation stimulation [55]. Here, the co-precipitation of EIF4E5 with ribosomal subunits as well as with a subunit of the eIF3 complex might reflect a direct presence in polysomes and a role in translation. This would still need to be verified but it would clearly be linked to its association with EIF4G1.

The presence of the EIF4G1-IP, with a putative mRNA cap guanine-N7 methyltransferase activity, as seen in *T. brucei* [24], still raises the possibility of EIF4E5 also being involved in cytoplasmic cap modification, or recapping, an activity that has recently come into focus in mammalian cells [56]. The IP data for EIF4E5 reinforced its association with EIF4G1-IP2 or RBP43 and also revealed an association with other RNA binding proteins. A relevant difference observed for the

Leishmania EIF4E5, in comparison with *T. brucei*, is the lack of substantial EIF4G2 enrichment in the immunoprecipitated fractions. This might reflect interspecific differences which would also be associated with a lack of the *T. brucei* EIF4G2 partner, the uncharacterized Tb-G2-IP (Tb927.11.6010), from *Leishmania* or *Crithidia* species. In this case, it remains to be defined what would be the EIF4G2 role and what are the reasons for its conservation throughout most trypanosomatid and related species [13].

Our structural analysis revealed molecular differences that could account for genera-specific mechanisms involved in the association of trypanosomatids EIF4E5 with their partners. Although the mRNA cap-binding site was found to be conserved between *T. cruzi* and *L. major* EIF4E5 orthologues, major differences were identified in the opposite side, the so-called convex side of the protein, responsible for the interaction with eIF4G partners and other eIF4E-binding proteins. The significant differences observed in the electrostatic

potential surfaces between the two trypanosomatids homologues, as well as between these and the human counterpart, indicate that these interactions involve different binding modes, which may also involve different partners. Our results show that the mutation of a highly conserved tryptophan (W73 in human eIF4E), essential to the association of eIF4E with eIF4G and eIF4E-BP in mammalian and other eukaryotes [20–22,33,41,42], does not abolish the interaction of EIF4E5 with EIF4G1 and the associated proteins in *Leishmania*. This may reflect distinct molecular features being involved in specific eIF4E interactions within trypanosomatids. It is important to notice that the fact that this mutation (*L. infantum* EIF4E5 W53A) does not disrupt the eIF4F-like complex associated with the *Leishmania* EIF4E5 does not necessarily imply that the conserved W53 is not involved in the interaction. However, it indicates that the mutation is not sufficient to abolish the identified EIF4E5 interactions. We hypothesize that substitutions common to both *Leishmania* and *Trypanosoma* species may contribute for a distinct interaction network which could attenuate the role for this tryptophan in mediating binding to specific protein partners. Concerning *LiEIF4E5* W45, the second tryptophan residue mutated in this work, the equivalent residue in human eIF4E does not participate in the interaction with eIF4G, whereas the equivalent residue in yeast eIF4E (F68) was shown to participate in eIF4G interaction *in vitro*, although only under certain experimental conditions [34,37]. Our results showing that this mutation does not abolish the eIF4F-like complex involving *LiEIF4E5* appears to corroborate that W45 is not essential for the eIF4E–eIF4G interaction. In contrast, both W45A and W50A mutations did impact on the association of *LiEIF4E5* with other partners. We hypothesize that both exposed tryptophan residues may have a role in regulating the EIF4E5 affinity for some protein partners, with direct or indirect influence on the stability of transient complexes. Structural data on the complexes of EIF4E5 with specific partners, such as EIF4G1, and/or mutation of two or more residues, will be necessary to better define the details of the interacting surfaces.

Both EIF4E4 and its partner PABP1 are known to be simultaneously phosphorylated by the cell-cycle related kinase CRK1 [57] and their phosphorylated motifs also suggest phosphorylation by MAP-like kinases. In contrast, kinases associated with the phosphorylation of other trypanosomatid eIF4E or even eIF4G homologues are not known. For EIF4E5, our mass-spectrometry analyses have identified at least one relevant kinase that might be involved in its regulation, the NEK-12 kinase homologue. This is a member of a protein kinase family involved in the regulation of microtubule and cytoskeleton synthesis (reviewed in [58]) and whose orthologue, RDK2, is involved in regulating *T. brucei* differentiation [59]. More relevant still is the EIF4E5 association with the 14-3-3 proteins, also identified with the *T. brucei* orthologue [24]. In *T. brucei* these are structurally conserved proteins known to bind phosphorylated motifs and with roles related to regulation of motility, cell-cycle and cytokinesis [60]. They have been shown to function mostly as a heterodimer and to bind specifically to the AKB1 kinase, however their target proteins still have to be properly defined

[61,62]. The co-precipitation of both 14-3-3 proteins strongly reinforces a role in regulating EIF4E5 function. Nevertheless, the lack of conservation of the identified phosphorylated motifs in the *T. brucei* orthologue might be an indication that, instead of EIF4E5, they act targeting a protein partner, such as EIF4G1.

Although the EIF4E4/PABP1 phosphorylation seems to enhance the activity of both proteins during translation [63], the ability of both recombinant, non-phosphorylated, proteins to interact *in vitro* and the position of the phosphorylated residues in EIF4E4 proximal to PAM2 motifs likely mediating PABP1 binding [45] suggests that phosphorylation might inhibit their interaction. For EIF4E3, at least one phosphorylation site mapped to the *L. donovani* orthologue also overlaps with a PAM2 motif [64] suggesting an effect upon the binding to partners such as the PABP homologues. With regards to the impact of the C-terminal phosphorylation of EIF4E5 on its function, structural and functional analysis have demonstrated the relevance of the C-terminal region for the cap binding in human eIF4E [32], *Schistosoma mansoni* eIF4E [30] and *T. cruzi* EIF4E5 [39]. The phosphorylation described here reinforces a relevant role of the EIF4E5 C-terminus, but the effect of this phosphorylation on the protein's function still needs to be better characterized so as to define whether it impacts on the direct binding to the cap, to neighbouring nucleotides or to protein partners. Nevertheless, the lack of sequence conservation for the EIF4E5 C-terminal region between the *Leishmania* and *Trypanosoma* orthologues, mainly concerning the phosphorylation sites, may reflect different regulatory mechanisms acting upon these proteins that also need to be further defined. Overall, our findings constitute additional evidence of the diverse and species-specific mechanisms that may be found in the context of the translation initiation in trypanosomatids.

Materials and methods

Sequence analysis

Protein sequences were aligned using the MAFFT software (default settings) [65]. The sequences of the trypanosomatid EIF4E5 orthologues used in the alignment are all available at the TriTrypDB Webpage (<http://tritrypdb.org/>). The identification of the serine-proline (SP) motifs and the other important sites were done through visual inspection of the sequences. Serine, threonine and tyrosine phosphorylation status of *LiEIF4E5* protein was predicted using the online open access server NetPhos 3.1 [66,67], considering only residues with the threshold value for the prediction higher than 0.5. TriTrypDB accessions: *Leishmania infantum* (Li)-LINF_360011500; *L. major* (Lm)- LmjF.36.0590; *L. mexicana* (Lmx)- LmxM.36.0590; *L. braziliensis* (Lb)- LbrM.35.0710; *Crithidia fasciculata* (Cf)- CFAC1_250013200; *Trypanosoma cruzi* (Tc)- BCY84_16159; *T. congolense* (Tcl)- TcIL3000_10_4200 and *T. brucei* (Tb)- Tb927.10.5020.

DNA manipulations

The *L. infantum* genomic DNA, from the MHOM/MA/67/ITMAP-263 strain, was isolated using DNAzol (Life Technologies) followed by amplification of the full-length wild-type EIF4E5 (Trytrip; LINF_360011500; NCBI: XP_001469544.1), using primers flanked by sites for the restriction enzymes BamHI and HindIII. Cloning procedures, site-directed mutagenesis and generation of the constructs for ectopic expression and gene deletion (single and double knockouts) were essentially carried out as previously described for the EIF4E4 and PABP1 genes [21,45]. Briefly, using the full-length wild-type EIF4E5 gene as template, two sets of PCR reactions were performed for the site-directed mutagenesis required for the EIF4E5^{W53A} and EIF4E5^{W45A} mutants, while just one PCR reaction was necessary for the EIF4E5^{S05S208} mutant, since the mutated sequence was localized within its C-terminus. Wild-type and mutant genes were cloned into pGEM-T Easy vector (Promega), sequenced and subcloned into the BamHI/HindIII sites of the *Leishmania* expression vector pSPBT1YNEOα. To generate the construct for gene deletion, consisting of 500 bp fragments flanking the coding sequence from the *L*EIF4E5 gene joined to the coding sequences of the hygromycin resistance gene, a PCR fusion-based strategy was used. The resulting amplified fragment was cloned into the pGEM-T Easy vector, sequenced, and recovered by digestion with NotI, followed by gel-purification and transfection into *L. infantum* promastigotes. Relevant primers are listed in the Supplementary Table S5.

Parasite growth and transfection

Leishmania infantum MHOM/MA/67/ITMAP-263 promastigotes were cultured in Schneider's insect medium (Sigma) pH 7.2 and 1% penicillin / streptomycin, supplemented with 10% foetal bovine serum and 0.2% haemin, at 26°C. Growth curves were set up using late-stationary phase cells passaged to fresh medium at a pre-established cell density of 10⁶ cells/ml. Samples were taken at selected time points for processing, SDS-PAGE, and blotting. Transfections were carried out using standard electroporation procedures, followed by growth in Schneider's medium in the presence of neomycin (80 µg/ml G418, Sigma), used for circular plasmids (episomal expression), and hygromycin B (80 µg/ml, Sigma) for linear DNA fragments (for integration and gene deletion).

Antibodies and Western blotting

Western blotting was performed using standard procedures. Antisera against the *Leishmania* PABP1, PABP2, EIF3E and ribosomal protein P0, as well as their use, have been previously described [43–45]. The rabbit anti- EIF4AI antibody (1:5000 dilution) from *T. brucei* [68] was used as a control due to its constitutive expression. For the detection of the different HA-tagged proteins, a commercial anti-HA monoclonal antibody (1:3000 dilution) from Sigma was used.

Cytoplasmic extract preparation and immunoprecipitation

For the mass-spectrometry analysis, total cytoplasmic extracts from exponentially grown *L. infantum* promastigotes, either non-transfected cells (negative control) or transgenic cell lineages expressing either the HA-tagged wild-type *L*EIF4E5 or the EIF4E5^{W53A} and EIF4E5^{W45A} mutants, were generated in duplicates after cells lysis using nitrogen cavitation [69]. Briefly, late exponentially grown *L. infantum* promastigotes were harvested and washed once in ice cold PBS, followed by resuspension in HEPES-lysis buffer (20 mM HEPES-KOH pH7.4, 75 mM potassium acetate, 4 mM magnesium acetate, 2 mM DTT, supplemented with EDTA-free protease inhibitors from Roche) to a concentration of 4 × 10⁷ cells/ml. The resuspended cells were transferred into the cavitation chamber of the cell disruption vessel (Parr Instruments) and incubated at 4°C under pressure of 70 bar for 45 minutes, followed by rapid decompression. The lysates were submitted to centrifugation for 10 minutes at 17.000 g, 4°C to remove cellular debris and the supernatants, the cytoplasmic extracts, aliquoted and stored at –80°C.

All cytoplasmic extracts were mixed with Monoclonal Anti-HA Agarose (Sigma) as per manufacturer's protocol. Briefly, 80 µg of the Anti-HA agarose were washed five times with ice cold PBS supplemented with protease inhibitors (Roche) followed by the incubation with the cytoplasmic extracts for 2 h at 4°C. After, centrifugation and removal of the supernatant, the beads were washed four more times with PBS plus protease inhibitors. The resulting immunoprecipitated complexes were then directly eluted from the beads in SDS-PAGE sample buffer. They were analysed through SDS-PAGE and Western-blotting using antibodies against the HA-tag to confirm the efficiency of the precipitation reaction. The immunoprecipitations were performed in duplicates using different cytoplasmic extracts.

Mass-spectrometry analysis

For mass-spectrometry (MS) and protein identification, the immunoprecipitated samples from the HA-tagged EIF4E5 and both W45A and W53A mutants were loaded into 15% SDS-PAGE gels and allowed to enter the resolving gel. Electrophoresis was interrupted before protein fractionation with the gel slices containing the whole protein samples excised. These slices were analysed at the Proteomics facility of the Carlos Chagas Institute – FIOCRUZ, where they were submitted to in-gel tryptic digestion and mass-spectrometry analysis, as previously described [44]. The same dataset was used for the subsequent search for post-translational modifications using the MaxQuant platform (version 1.6.14) [70]. Protein identification was based on the *L. infantum* protein sequence database (*L. infantum* JPCM5, version from 29 March 2016 available at TriTrypDB). For subsequent analysis, a first normalization was performed based on the sum of the intensities of all identified proteins, which was calculated for each replicate of the tagged proteins and the negative control. The highest sum was then used to calculate the ratio to the sums of the other samples and this ratio was

then used to normalize the individual intensities for each identified protein. From then on only proteins with at least two or more peptides in both replicates of the HA-tagged proteins were considered. Enrichment ratios between the average intensities generated from the HA-tagged EIF4E5_{WT} IPs and the average intensities from the control IPs were next calculated for each co-precipitated protein and both these intensities and enrichment ratios used to define the proteins included in the Tables and Figures. For the analyses of the IPs with the *LiEIF4E5* mutants, the intensity values of the co-precipitated proteins were normalized for a second time using the intensity value obtained for the HA-tagged EIF4E5 bait from the wild-type samples as reference. For this, the intensities obtained for HA-tagged EIF4E5_{W50A} and HA-tagged EIF4E5_{W45A} baits were divided by the intensity value obtained for the HA-tagged EIF4E5 wild-type. The resulting ratio values were then multiplied by the intensities of the co-precipitating proteins from each respective mutant sample. Finally, the ratios for the intensities between the proteins in the wild type and mutant samples were calculated and compared.

Expression and purification of *Leishmania major* EIF4E5

The DNA fragment encoding the full-length *LmeIF4E5* was amplified from *L. major* (Friedlin) genomic DNA flanked by sites for the Nco I/Not I restriction enzymes and cloned into the same sites of the pET-21d vector (Novagen). *Escherichia coli* BL21 (DE3) slyD⁻ cells [71] carrying the plasmid pRARE2 (Novagen) were transformed with the recombinant plasmid and grown at 37°C in LB medium plus ampicillin and induced at 30°C (OD₆₀₀ ~ 0.6) for four hours with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells from one-litre cultures were harvested, resuspended in 20 mL of buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 10 mM 2-Mercaptoethanol) and lysed using a microfluidizer processor (Microfluidics™). The soluble fraction was isolated by centrifugation at 25,000 x g for 30 minutes at 4°C and loaded onto a 1 mL His-Trap HP column (GE Healthcare Life Sciences) equilibrated with buffer A. *LmeIF4E5* was eluted with a two-step gradient in a 10 column volumes (CV) linear gradient from 0% to 10% buffer B (buffer A + 500 mM imidazole) followed by a 20 CV linear gradient from 10% to 100% buffer B. Fractions containing the target protein were pooled, dialyzed against 50 mM Tris-HCl pH 8.0 and loaded onto a 1 mL Hi-Trap Q HP column (GE Healthcare Life Sciences). *LmeIF4E5* was eluted with a 20 CV linear gradient from 0 to 1 M NaCl, 50 mM Tris-HCl pH 8.0. The sample was concentrated to a final volume of 2 mL and loaded onto a Superdex 75 16/60 (GE Healthcare Life Sciences) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT. The eluted protein was concentrated to approximately 5 mg/mL for crystallization assays. Expression of *LmeIF4E5* in *E. coli* was achieved prior to the start of the *L. infantum* EIF4E5 work, so no attempts were made to express the *LiEIF4E5* in *E. coli*.

Crystallization, data collection and processing

Protein samples in the presence and absence of the cap-4 analogue were submitted to crystallization trials by sitting drop vapour-diffusion method using commercial screens. Optimization of the initial crystallization conditions were performed by varying precipitant and protein concentrations. Suitable crystals were obtained only in the absence of the ligand. The best crystals were obtained by hanging drop vapour diffusion at 18°C by mixing the *LmeIF4E5* protein at 5 mg/mL in 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT with crystallization buffer containing 1.25 M Sodium Citrate and 0.1 M Sodium Cacodylate pH 7.0. The crystals were cryo-protected with 20% (v/v) glycerol added to the mother liquor prior to flash cooling in liquid nitrogen. X-ray diffraction data were collected at the PROXIMA 1 beam line of the Synchrotron SOLEIL using an EIGER 16 M detector (Dectris). Diffraction data were processed with XDS package [72]. Data statistics are presented in the Table 3.

Structure determination, refinement and analysis

The structure of the *L. major* EIF4E5 was determined by molecular replacement using MOLREP [73]. A homology model generated by the Phyre2 server [49] was used as the search model. Refinement of the structure was performed alternating cycles of BUSTER [74] with visual inspection and manual rebuilding using COOT [75]. Model validation was performed using the MolProbity Web Server (<http://molprobity.biochem.duke.edu/>) [76]. Electrostatic potential calculations were performed using the Adaptive Poisson-Boltzmann Solver through the software Chimera [77]. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB ID 7KCJ).

Acknowledgments

The authors acknowledge the FIOCRUZ Technical Platform program through the Platforms for Protein Purification and Characterization and Proteomics at the Carlos Chagas Institute as well as the DNA sequencing Platform at the Aggeu Magalhães Institute. The authors are grateful to Synchrotron SOLEIL for beam time awarded at PROXIMA1 beam line (Proposal number 20191165) and all the support they benefited as SOLEIL's users. The reported work in Dr. de Melo Neto's laboratory was supported by grants provided by the Brazilian funding agencies FACEPE (APQ-1662-2.02/15) and CNPq (401282/2014-7). The structural work was supported by the Paraná State Science Support Foundation (FA 005/2016), CNPq (442323/2019-0) and by a PTR (Programmes Transversaux de Recherche) grant (PTR 190-19) from Institut Pasteur Paris. Nilson I. T. Zanchin, Beatriz G. Guimarães and Osvaldo P. Melo Neto are CNPq research fellows (grants nr. 304167/2019-3; 304788/2018-0; 310032/2019-9). Studentships for the graduate and undergraduate students (GBL, TYVLC, ANALMB, LAA and RPAV) were provided by CAPES, CNPq, and FACEPE, while ERF was a recipient of a BJT fellowship awarded by CNPq as part of the 401282/2014-7 grant.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the CNPq [442323/2019-0]; CNPq [401282/2014-7]; Paraná State Science Support Foundation [FA 005/2016]; Programmes Transversaux de Recherche - Pasteur Institute [PTR 190-19]; FACEPE [APQ-1662-2.02/15].

ORCID

Nilson Ivo Tonin Zanchin  <http://orcid.org/0000-0002-1153-0694>
 Beatriz Gomes Guimarães  <http://orcid.org/0000-0002-8546-2460>
 Osvaldo Pompilio de-Melo-Neto  <http://orcid.org/0000-0001-5402-7346>

References

- [1] Kaufer A, Ellis J, Stark D, et al. The evolution of trypanosomatid taxonomy. *Parasites Vectors*. 2017;10(1):1–17.
- [2] El-Sayed NM, Myler PJ, Blandin G, et al. Comparative genomics of trypanosomatid parasitic protozoa. *Science*. 2005;309(5733):404–409.
- [3] Günzl A, Günzl A. The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *EukaryotCell*. 2010;9:1159–1170.
- [4] Fernández-Moya SM, Estévez AM. Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosomatid protozoan parasites. *Wiley Interdiscip Rev RNA*. 2010;1(1):34–46.
- [5] Clayton CE. Networks of gene expression regulation in *trypanosoma brucei*. *Mol Biochem Parasitol*. 2014;195(2):96–106.
- [6] Clayton C. Regulation of gene expression in trypanosomatids: Living with polycistronic transcription. *Open Biol*. 2019;9(6):9.
- [7] Hinnebusch AG. The scanning mechanism of eukaryotic translation initiation. *AnnuRevBiochem*. 2014;83:779–812.
- [8] Merrick WC. eIF4F: a retrospective. *JBiolChem*. 2015;290:24091–24099.
- [9] Merrick WC, Pavitt GD. Protein synthesis initiation in eukaryotic cells. *Cold Spring Harb Perspect Biol*. 2018;10(12):a033092.
- [10] Batool A, Aashaq S, Andrabi KI. Eukaryotic initiation factor 4E (eIF4E): a recap of the cap-binding protein. *J Cell Biochem*. 2019;120(9):14201–14212.
- [11] Borden KLB. The eukaryotic translation initiation factor eIF4E wears a “cap” for many occasions. *Translation*. 2016;4(2):e1220899.
- [12] Volpon L, Osborne MJ, Borden KLB. Biochemical and structural insights into the eukaryotic translation initiation factor eIF4E. *Curr Protein Pept Sci*. 2019;20(6):525–535.
- [13] Bannerman BP, Kramer S, Dorrell RG, et al. Multispecies reconstructions uncover widespread conservation, and lineage-specific elaborations in eukaryotic mRNA metabolism. *PLoS One*. 2018;13(3):1–23.
- [14] Freire E, Sturm N, Campbell D, de Melo Neto O. The role of cytoplasmic mRNA cap-binding protein complexes in *trypanosoma brucei* and other trypanosomatids. *Pathogens*. 2017;6(4):55.
- [15] Dhalia R, Reis CRSR, de Melo Neto OPP, et al. Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. *Mol Biochem Parasitol*. 2005;140(1):23–41.
- [16] Freire ER, Dhalia R, Moura DMN, et al. The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. *Mol Biochem Parasitol*. 2011;176(1):25–36.
- [17] Yoffe Y, Zuberek J, Lewdorowicz M, et al. Cap-binding activity of an eIF4E homologue from *Leishmania*. *RNA*. 2004;10(11):1764–1775.
- [18] Yoffe Y, Zuberek J, Lerer A, et al. Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in *Leishmania*. *Eukaryot Cell*. 2006;5(12):1969–1979.
- [19] Yoffe Y, Leger M, Zinoviev A, et al. Evolutionary changes in the *Leishmania* eIF4F complex involve variations in the eIF4E–eIF4G interactions. *Nucleic Acids Res*. 2009;37(10):3243–3253.
- [20] Zinoviev A, Manor S, Shapira M. Nutritional stress affects an atypical cap-binding protein in *Leishmania*. *RNABiol*. 2012;9(12):1450–1460.
- [21] de Melo Neto OP, da Costa Lima TDC, Xavier CC, et al. The unique *Leishmania* EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. *RNA Biol*. 2015;12(11):1–13.
- [22] Zinoviev A, Leger M, Wagner G, et al. A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. *Nucleic Acids Res*. 2011;39(19):8404–8415.
- [23] Freire ER, Malvezzi AM, Vashisht AA, et al. *Trypanosoma brucei* translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog. *Eukaryot Cell*. 2014;13(7):896–908.
- [24] Freire ER, Vashisht AA, Malvezzi AM, et al. eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in *Trypanosoma brucei*. *RNA*. 2014;20(8):1272–1286.
- [25] Niedzwiecka A, Marcotrigiano J, Stepinski J, et al. Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *JMolBiol*. 2002;319:615–635.
- [26] Poulin F, Gingras AC, Olsen H, et al. 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. *J Biol Chem*. 1998;273(22):14002–14007.
- [27] Khaleghpour K, Pyronnet S, Gingras A-C, et al. Translational homeostasis: Eukaryotic translation initiation factor 4E Control of 4E-binding protein 1 and p70 S6 kinase activities. *Mol Cell Biol*. 1999;19(6):4302–4310.
- [28] Piserà A, Campo A, Campo S. Structure and functions of the translation initiation factor eIF4E and its role in cancer development and treatment. *J Genet Genomics*. 2018;45(1):13–24.
- [29] Marcotrigiano J, Gingras AC, Sonenberg N, et al. Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell*. 1997;89(6):951–961.
- [30] Liu W, Zhao R, McFarland C, et al. Structural insights into parasite eIF4E binding specificity for m7G and m2,2,7G mRNA caps. *J Biol Chem*. 2009;284(45):31336–31349.
- [31] Liu W, Jankowska-Anyszka M, Piecyk K, et al. Structural basis for nematode eIF4E binding an m(2,2,7)G-Cap and its implications for translation initiation. *Nucleic Acids Res*. 2011;39(20):8820–8832.
- [32] Tomoo K, Shen X, Okabe K, et al. Crystal structures of 7-methylguanosine 5'-triphosphate (m7GTP)- and P1-7-methylguanosine-P3-adenosine-5', 5'-triphosphate (m7GppA)-bound human full-length eukaryotic initiation factor 4E: Biological importance of the C-terminal flexible region. *Biochem J*. 2002;362(3):539–544.
- [33] Marcotrigiano J, Gingras AC, Sonenberg N, et al. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *MolCell*. 1999;3:707–716.
- [34] Gross JD, Moerke NJ, Von Der Haar T, et al. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell*. 2003;115(6):739–750.
- [35] Umenaga Y, Paku KS, In Y, et al. Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. *Biochem Biophys Res Commun*. 2011;414(3):462–467.
- [36] Grüner S, Peter D, Weber R, et al. The structures of eIF4E-eIF4G complexes reveal an extended interface to regulate translation initiation. *Mol Cell*. 2016;64(3):467–479.
- [37] Gruner S, Weber R, Peter D, et al. Structural motifs in eIF4G and 4E-BPs modulate their binding to eIF4E to regulate translation initiation in yeast. *Nucleic Acids Res*. 2018;46(13):6893–6908.
- [38] Meleppattu S, Arthanari H, Zinoviev A, et al. Structural basis for LeishIF4E-1 modulation by an interacting protein in the human parasite *Leishmania major*. *Nucleic Acids Res*. 2018;46(7):3791–3801.

- [39] Reolon LW, Vichier-Guerre S, De Matos BM, et al. Crystal structure of the *Trypanosoma cruzi* EIF4E5 translation factor homologue in complex with mRNA cap-4. *Nucleic Acids Res.* 2019;47(11):5973–5987.
- [40] Kaufer A, Stark D, Ellis J. Evolutionary insight into the trypanosomatidae using alignment-free phylogenomics of the kinetoplast. *Pathogens.* 2019;8(3):8.
- [41] Ptushkina M, Von Der HT, Vasilescu S, et al. Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *Embo J.* 1998;17(16):4798–4808.
- [42] Pyronnet S, Imataka H, Gingras AC, et al. Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *Embo J.* 1999;18(1):270–279.
- [43] Da Costa Lima TD, de Melo Neto OP, Moura DMN, et al. Functional characterization of three leishmania poly(a) binding protein homologues with distinct binding properties to RNA and protein partners. *Eukaryot Cell.* 2010;9(10):1484–1494.
- [44] Rezende AM, Assis LA, Nunes ECC, et al. The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates—identification of conserved and divergent features based on orthologue analysis. *BMC Genomics.* 2014;15(1):1175.
- [45] De Melo Neto OP, Da Costa Lima TDC, Merlo KC, et al. Phosphorylation and interactions associated with the control of the *Leishmania* poly-A binding protein 1 (PABP1) function during translation initiation. *RNA Biol.* 2018;6286:1–17.
- [46] Pereira MMCMC, Malvezzi AMM, Nascimento LMM, et al. The eIF4E subunits of two distinct trypanosomatid eIF4F complexes are subjected to differential post-translational modifications associated to distinct growth phases in culture. *Mol Biochem Parasitol.* 2013;190(2):82–86.
- [47] Shrivastava R, Drory-Retwitzer M, Shapira M. Nutritional stress targets LeishIF4E-3 to storage granules that contain RNA and ribosome components in *Leishmania*. *PLoS Negl Trop Dis.* 2019;13(3):e0007237.
- [48] Moretti NS, Cestari I, Anupama A, et al. Comparative proteomic analysis of lysine acetylation in trypanosomes. *J Proteome Res [Internet].* 2018;17(1):374–385.
- [49] Kelley L, Mezulis S, Yates C, et al. The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc.* 2015;10(6):845–858.
- [50] Terrao M, Marucha KK, Mugo E, et al. The suppressive cap-binding complex factor 4EIP is required for normal differentiation. *Nucleic Acids Res.* 2018;46(17):8993–9010.
- [51] Freire ER, Moura DMN, Bezerra MJR, et al. *Trypanosoma brucei* EIF4E2 cap-binding protein binds a homolog of the histone-mRNA stem-loop-binding protein. *Curr Genet.* 2018;64(4):821–839.
- [52] Erben ED, Fadda A, Lueong S, et al. A genome-wide tethering screen reveals novel potential post-transcriptional regulators in *Trypanosoma brucei*. *PLOS Pathog.* 2014;10(6):e1004178.
- [53] Lueong S, Merce C, Fischer B, et al. Gene expression regulatory networks in *T trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Mol Microbiol.* 2016;100(3):457–471.
- [54] Klein C, Terrao M, Inchaustegui GD, et al. Polysomes of *Trypanosoma brucei*: Association with initiation factors and RNA-binding proteins. *PLoS One.* 2015;10(8):e0135973.
- [55] Melo Do Nascimento L, Terrao M, Marucha KK, et al. The RNA-associated proteins MKT1 and MKT1L form alternative PBP1-containing complexes in *Trypanosoma brucei*. *J Biol Chem.* 2020;295(32):10940–10955.
- [56] Trotman JB, Schoenberg DR. A recap of RNA recapping. *Wiley Interdiscip Rev RNA.* 2018;10(1):e1504.
- [57] Hu H, Gourguechon S, Wang CC, et al. The G1 Cyclin-dependent kinase CRK1 in *trypanosoma brucei* regulates anterograde protein transport by phosphorylating the COPII subunit sec31. *J Biol Chem.* 2016;291(30):15527–15539.
- [58] Fry AM, Bayliss R, Roig J. Mitotic regulation by NEK kinase networks. *Front Cell Dev Biol.* 2017;5:1–13.
- [59] Jones NG, Thomas EB, Brown E, et al. Regulators of *trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-Wide RNAi screen. *PLoS Pathog.* 2014;10(1):e1003886.
- [60] Inoue M, Nakamura Y, Yasuda K, et al. The 14-3-3 proteins of *Trypanosoma brucei* function in motility, cytokinesis, and cell cycle. *J Biol Chem.* 2005;280(14):14085–14096.
- [61] Inoue M, Yasuda K, Uemura H, et al. *Trypanosoma brucei* 14-3-3I and II proteins predominantly form a heterodimer structure that acts as a potent cell cycle regulator in vivo. *J Biochem.* 2013;153(5):431–439.
- [62] Inoue M, Okamoto K, Uemura H, et al. Identification and characterization of a cell division-regulating kinase AKB1 (associated kinase of *Trypanosoma brucei* 14-3-3) through proteomics study of the Tb14-3-3 binding proteins. *J Biochem.* 2015;158(1):49–60.
- [63] An T, Liu Y, Gourguechon S, et al. CDK phosphorylation of translation initiation factors couples protein translation with cell-cycle transition. *Cell Rep.* 2018;25(3204–3214):e5.
- [64] Tsigankov P, Gherardini PF, Helmer-Citterich M, et al. Regulation dynamics of leishmania differentiation: deconvoluting signals and identifying phosphorylation trends. *Mol Cell Proteomics.* 2014;13(7):1787–1799.
- [65] Katoh K, Misawa K, Kuma K, et al. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 2002;30(14):3059–3066.
- [66] Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol.* 1999;294(5):1351–1362.
- [67] Blom N, Sicheritz-Pontén T, Gupta R, et al. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics.* 2004;4(6):1633–1649.
- [68] Dhalia R, Marinsek N, Reis CRSR, et al. The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. *Nucleic Acids Res.* 2006;34(9):2495–2507.
- [69] Mureev S, Kovtun O, Nguyen UTT, et al. Species-independent translational leaders facilitate cell-free expression. *Nat Biotechnol.* 2009;27(8):747–752.
- [70] Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 2008;26(12):1367–1372.
- [71] Roof WD, Horne SM, Young KD, et al. slyD, a host gene required for phi X174 lysis, is related to the FK506-binding protein family of peptidyl-prolyl cis-trans-isomerases. *J Biol Chem.* 1994;269(4):2902–2910.
- [72] Kabsch W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biol Crystallogr.* 2010;66(2):133–144.
- [73] Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr Sect D Biol Crystallogr.* 2010;66(1):22–25.
- [74] Bricogne G, Blanc E, Brandl M, et al. BUSTER version 2.10.3. Cambridge, United Kingdom:Global Phasing Ltd.; 2017.<http://www.globalphasing.com>.
- [75] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr.* 2004;60(12):2126–2132.
- [76] Chen VB, Arendall WB, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D Biol Crystallogr.* 2010;66(1):12–21.
- [77] Pettersen EF, Goddard TD, Huang CC, et al. UCSF chimera?A visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605–1612.