

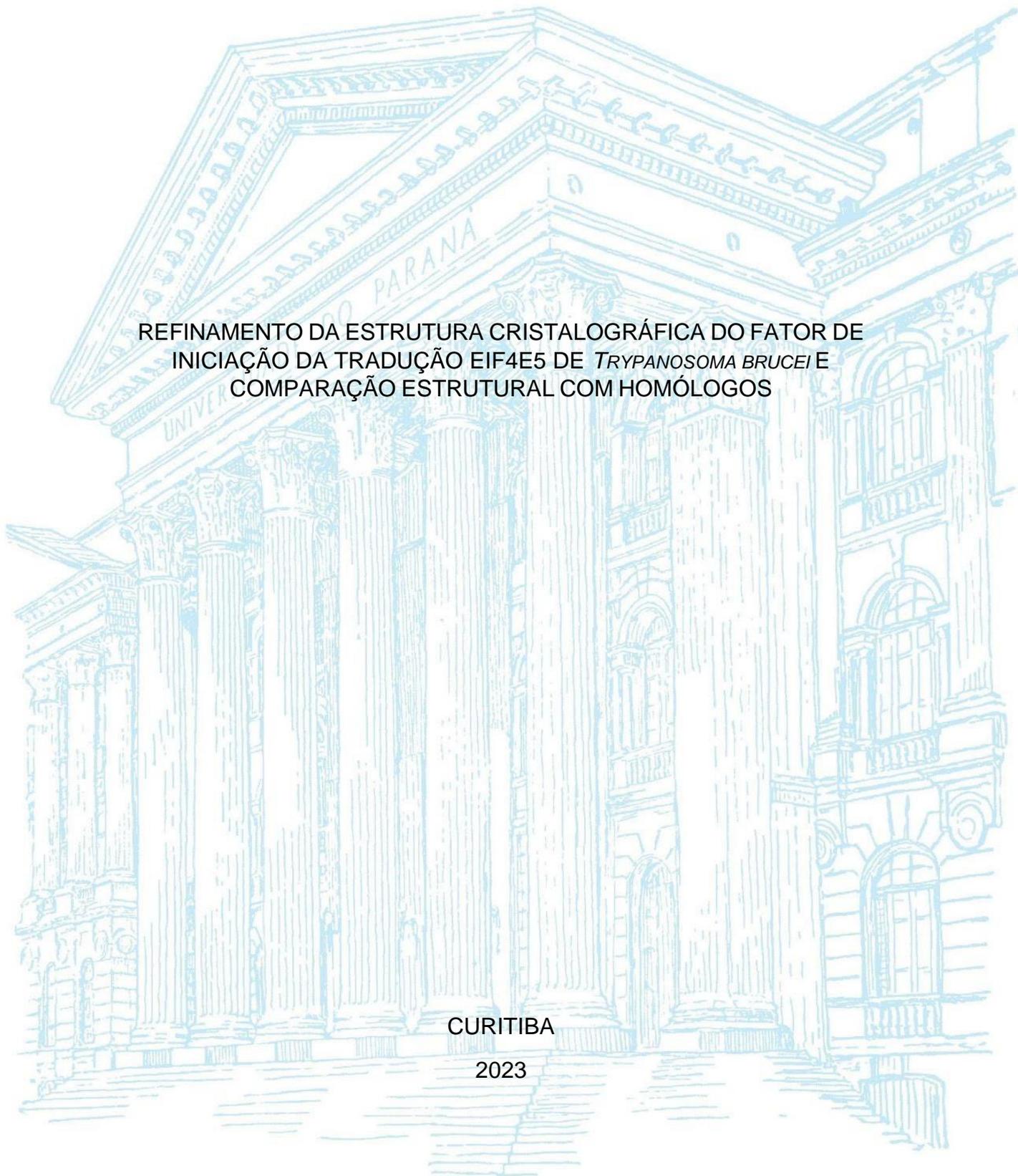
UNIVERSIDADE FEDERAL DO PARANÁ

BEATRIZ MARIA DA SILVA PEREIRA

REFINAMENTO DA ESTRUTURA CRISTALOGRÁFICA DO FATOR DE  
INICIAÇÃO DA TRADUÇÃO EIF4E5 DE *TRYPANOSOMA BRUCEI* E  
COMPARAÇÃO ESTRUTURAL COM HOMÓLOGOS

CURITIBA

2023



BEATRIZ MARIA DA SILVA PEREIRA

REFINAMENTO DA ESTRUTURA CRISTALOGRÁFICA DO FATOR DE  
INICIAÇÃO DA TRADUÇÃO EIF4E5 DE *TRYPANOSOMA BRUCEI* E  
COMPARAÇÃO ESTRUTURAL COM HOMÓLOGOS

Trabalho de conclusão de curso apresentado ao curso de graduação em Biomedicina, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Bacharel em Biomedicina.

Orientador(a): Prof. Dr. Wanderson Duarte da Rocha

Coorientador(a): Profa. Dra. Beatriz Gomes Guimarães

CURITIBA

2023



## TERMO DE APROVAÇÃO

BEATRIZ MARIA DA SILVA PEREIRA

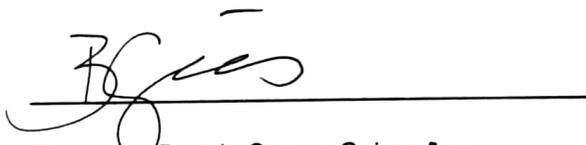
### REFINAMENTO DA ESTRUTURA CRISTALOGRÁFICA DO FATOR DE INICIAÇÃO DA TRADUÇÃO EIF4E5 DE *TRYPANOSOMA BRUCEI* E COMPARAÇÃO ESTRUTURAL COM HOMÓLOGOS

Trabalho de Conclusão de Curso apresentado ao curso de Graduação em Biomedicina, Setor de Ciências Biológicas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Bacharel em Biomedicina.



Prof. Dr. Wanderson Duarte da Rocha

Orientador – Departamento de Bioquímica, UFPR



Profa. Dra Beatriz Gomes Guimarães

Co-orientadora – Departamento de Bioquímica, UFPR



Profa. Dra Elaine Machado Benelli

Departamento de Bioquímica, UFPR

Cidade, 20 de junho de 2023.

Dedico este trabalho aos meus pais, que fizeram inúmeros esforços para que eu sempre tivesse acesso a uma ótima formação e educação.

## AGRADECIMENTOS

A Deus, por me conceder as oportunidades que desejo vivenciar para meu crescimento e minha evolução pessoal, espiritual e profissional.

Aos meus pais, que sempre me ensinaram o quanto a educação é importante e que sempre renderá bons frutos, pelo apoio e incentivo incondicional em todos os meus anos de vida, a encontrar e seguir o caminho do conhecimento com o qual me identifico.

A minha orientadora de Iniciação Científica, Dra. Beatriz Gomes Guimarães, pela oportunidade e confiança a mim concedida, pelos ensinamentos diários sobre como fazer ciência, e sobre como compreender e analisar estruturas de proteínas de maneira a exprimir informações e conhecimento, sendo para mim uma referência e exemplo como pesquisadora a ser seguido, especialmente na área de Biologia Estrutural e Cristalografia.

Aos meus amigos e colegas do Laboratório de Biologia Estrutural e Engenharia de Proteínas da FIOCRUZ-PR, pelo apoio na bancada enquanto aprendia e realizava meus experimentos, pela paciência em responder minhas inúmeras perguntas e questionamentos, tornando mais tranquilo compreender a complexidade da ciência que fazemos.

A minha irmã, pelo apoio e escuta das minhas falas e discursos sobre estruturas de proteínas, mesmo não compreendendo nada sobre o assunto. E, aos meus amigos, que sempre apoiaram minhas decisões e me incentivaram a seguir meus desejos e sonhos, mesmo que parecessem inalcançáveis.

Aos meus professores, que desde a escola, pré-vestibular e universidade, me passaram e ensinaram o conhecimento que hoje utilizo como base para todas as coisas mais complexas que desejo aprender. Em especial, a André Hoelz, pesquisador do CALTECH, que me proporcionou a primeira oportunidade de aprender sobre estrutura de proteínas, tendo sido essencial para minha descoberta pessoal sobre o que eu realmente gosto de fazer.

Você é livre quando gasta tempo de sua vida com as coisas que te motivam.

(JOSÉ MUJICA)

## RESUMO

Protozoários da classe kinetoplastida despertam interesse da ciência pela sua importância como agentes etiológicos de doenças como a tripanosomíase humana africana, doença de Chagas e leishmanioses. Além disso, também são estudados pela sua história evolutiva e por suas características celulares e moleculares únicas, desde a organização de seu material genético até a regulação de vias importantes para manutenção de seu metabolismo. Este trabalho tem como foco a iniciação da tradução, orquestrada por doze fatores cujo papel é realizar a montagem do ribossomo 80S competente para a elongação. Esta etapa da tradução é altamente regulada, consistindo no ponto mais importante de controle do metabolismo de proteínas. Devido às suas peculiaridades em tripanossomatídeos comparados a outros organismos eucariotos, a elucidação e caracterização estrutural e funcional das proteínas envolvidas na iniciação da tradução se torna relevante, entre as quais aquelas que formam o complexo eIF4F (eIF4E, eIF4G, eIF4A), que possui papel relevante neste processo. Os tripanossomatídeos possuem seis homólogos dos fatores eIF4Es, a proteína ligadora da estrutura cap presente na extremidade 5' do RNA mensageiro. A presença destes seis homólogos constitui mais uma característica importante destes organismos. Por conta disso, o laboratório de Biologia Estrutural e Engenharia de Proteínas do Instituto Carlos Chagas (FIOCRUZ – PR) têm trabalhado extensivamente para desvendar e compreender o papel destes fatores na iniciação da tradução. O objetivo é elucidar os mecanismos que podem diferenciar estas proteínas, além de seus papéis na regulação da tradução e mecanismos bioquímicos envolvidos na interação com os parceiros dos complexos de iniciação. Sabendo disso, o objetivo deste projeto foi refinar a estrutura cristalográfica do fator de iniciação da tradução eIF4E5 de *Trypanosoma brucei* (TbEIF4E5), co-cristalizado com o análogo cap-4 do mRNA, e analisado por difração de raios X. Além disso, realizamos a comparação estrutural de TbEIF4E5 com as proteínas homólogas de *T. cruzi* e *L. major*, além de eIF4E de humanos.

**Palavras chaves:** Fatores de Iniciação da Tradução, eIF4E, cap-4, *Trypanosoma brucei*, estrutura cristalográfica eIF4E5.

## ABSTRACT

Protozoa of the kinetoplastids class arouse scientific interest for their importance as etiologic agents of diseases such as human African trypanosomiasis, Chagas disease, and leishmaniasis. In addition, they are also studied for their evolutionary history and for their unique biochemical characteristics, from the organization of their genetic material to the regulation of relevant pathways for the maintenance of their metabolism. This work focuses on translation initiation, orchestrated by twelve factors whose role is to carry out the assembly of the 80S ribosome competent for elongation. This step is highly regulated, and is the most important control point of protein metabolism, overriding the control of DNA metabolism, for example. Due to its peculiarities in trypanosomatids compared to other eukaryotic organisms, the elucidation and structural and functional characterization of proteins involved in the initiation of translation becomes relevant, including those that form the EIF4F complex (EIF4E, EIF4G, EIF4A), which has a relevant role in this process. Trypanosomatids possess six homologs of EIF4E's factors, the cap structure-binding protein present in the 5' end of the messenger RNA. The presence of these six homologs constitutes another important characteristic of these organisms. Because of this, the Laboratory of Structural Biology and Protein Engineering of the Carlos Chagas Institute (FIOCRUZ - PR) has been working extensively to unravel and understand the role of these factors in translation initiation. The goal is to elucidate the mechanisms that may differentiate these proteins, as well as their roles in the regulation of translation and the biochemical mechanisms involved in the interaction with the partners of the initiation complexes. Knowing this, the aim of this project was to refine the structure of the translation initiation factor EIF4E5 of *Trypanosoma brucei* (TbEIF4E5), co-crystallized with the mRNA analog *cap-4*, and analyzed by X-ray diffraction. In addition, we performed a structural comparison of TbEIF4E5 with homologous proteins from *T. cruzi* and *L. major*, and eIF4E from humans.

Keywords: Translation Initiation Factors, eIF4E, *cap-4*, *Trypanosoma brucei*, crystallographic structure eIF4E5

## LISTA DE FIGURAS

Figura 1 - Representação esquemática do mecanismo de iniciação da tradução em eucariotos .....	18
Figura 2 - Representação esquemática da estrutura cap-4 presente em tripanosomatídeos.....	21
Figura 3 - Estrutura global do fator de iniciação eIF4E (rosa) de humanos em complexo com um fragmento de eIF4G (ciano) e o m <sup>7</sup> GTP (Depósito PDB 5T46) .....	25
Figura 4 - Detalhe da interação de <i>Tc</i> eIF4E5 com o <i>cap</i> -4 .....	26
Figura 5 - Representação da definição dos parâmetros de uma célula unitária .....	28
Figura 7 - Mapa de densidade eletrônica visualizado em nível de sigma 0.98 RMSD .....	39
Figura 8 - Representação do modelo cristalográfico do fator de iniciação da tradução eIF4E5 de <i>Trypanosoma brucei</i> .....	41
Figura 9 - Regiões de interação entre <i>Tb</i> eIF4E5 e <i>cap</i> -4 do mRNA.....	43
Figura 10 - Alinhamento estrutural dos fatores de iniciação da tradução .....	44
Figura 11 - Comparação das superfícies eletrostáticas dos fatores EIF4E5 de tripanossomatídeos e eIF4E de <i>H. sapiens</i> .....	47

## LISTA DE TABELAS

Tabela 1 - Os sete sistemas cristalinos, valores de suas arestas e ângulos, classificação na Rede de Bravais e símbolo correspondente.....	29
Tabela 2 - Dados cristalográficos e estatística do refinamento.....	36
Tabela 3 - Fitas beta e alfa hélices presentes na estrutura de <i>Tb4E5</i> , resíduos onde se inicia e termina cada uma das estruturas secundárias presentes .....	40
Tabela 4 - Dados referentes ao alinhamento estrutural de <i>TbEIF4E5</i> com as proteínas homólogas de <i>T. cruzi</i> , <i>L. major</i> e eIF4E de <i>Homo sapiens</i> .....	42
Tabela 5 - Análise estrutural dos modelos experimentais quanto ao número de resíduos no N terminal, C terminal e loops $\beta 1-\beta 2$ , $\beta 3-\beta 4$ , $\beta 5-\beta 6$ dos fatores de iniciação da tradução de <i>T. brucei</i> ( <i>Tb eIF4E5</i> , dados ainda não publicados), <i>T. cruzi</i> ( <i>Tc eIF4E5</i> , PDB 5T46), <i>L. major</i> ( <i>Lm eIF4E5</i> , PDB 7KCJ) e <i>Homo sapiens</i> ( <i>Hs eIF4E</i> , PDB 5T46) .....	44
Tabela 6: Resíduos dos fatores de iniciação da tradução EIF4E5 de <i>T. cruzi</i> ( <i>Tc4E5</i> ), <i>T. brucei</i> ( <i>Tb4E5</i> ) e <i>L. major</i> ( <i>Lm4E5</i> ), envolvidos na interação com a estrutura cap do mRNA.....	45

## LISTA DE ABREVIATURAS OU SIGLAS

APBS – do inglês, Adaptive Poisson-Boltzmann Solver

eIFs – do inglês, Eukaryotic initiation factors

GAP - do inglês, GTPase-activating protein

GDP – do inglês, Guanosine diphosphate

GTP – do inglês, Guanosine-5'-triphosphate

m<sup>7</sup>GTP: 7-metil guanosina trifosfato

mRNA – RNA Mensageiro

PABP - do inglês, poly(A)-binding protein

PDB – do inglês, Protein Data Bank

PEG – Polietilenoglicol

PIC - do inglês, pre-initiation complex

PLIP – do inglês, Protein-Ligand Interaction Profiler

Rfree – do inglês, Free R value

RMSD – do inglês, Root Mean Square Deviation

RNAi – RNA de Interferência

siRNA – Pequeno RNA interferente

SLBP - do inglês, stem-loop-binding protein

TC - do inglês, ternary complex

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>14</b>
1.1 OBJETIVOS .....	15
1.1.1 Objetivo geral .....	15
1.1.2 Objetivos específicos .....	15
<b>2 REVISÃO DE LITERATURA</b> .....	<b>15</b>
2.1.1 INICIAÇÃO DA TRADUÇÃO .....	16
2.1.2 COMPLEXO EIF4F .....	19
2.1.3 PECULIARIDADES DA INICIAÇÃO DA TRADUÇÃO EM TRIPANOSOMATÍDEOS .....	20
2.1.3.1 FATORES DE INICIAÇÃO DA TRADUÇÃO EIF4Es e EIF4Gs .....	20
2.1.3.2 CAP DO mRNA.....	20
2.1.3.3 CARACTERIZAÇÃO DOS FATORES EIF4Es. ....	22
2.2 ESTRUTURA DOS FATORES DE INICIAÇÃO DA TRADUÇÃO EIF4Es DE TRIPANOSOMATÍDEOS .....	24
<b>3 MATERIAL E MÉTODOS</b> .....	<b>26</b>
3.1 EXPRESSÃO, PURIFICAÇÃO E CRISTALIZAÇÃO DE EIF4E5 DE <i>T. brucei</i> .....	26
3.2 COLETA E PROCESSAMENTO DOS DADOS DE DIFRAÇÃO.....	27
3.3 DETERMINAÇÃO E REFINAMENTO DA ESTRUTURA.....	30
3.4 ANÁLISES ESTRUTURAIS.....	32
3.4.1 ANÁLISE ESTRUTURAL INDIVIDUAL .....	33
3.4.2 ANÁLISE ESTRUTURAL COMPARATIVA .....	33
3.4.3 ANÁLISE DE SUPERFÍCIE ELETROSTÁTICA.....	34
<b>4 APRESENTAÇÃO DOS RESULTADOS</b> .....	<b>35</b>
4.1 PROCESSAMENTO DOS DADOS E REFINAMENTO .....	35
4.2 REFINAMENTO DA ESTRUTURA DE EIF4E5 DE <i>Trypanosoma brucei</i> .....	35
4.3 ANÁLISE ESTRUTURAL INDIVIDUAL.....	38
4.4 ANÁLISE ESTRUTURAL COMPARATIVA.....	42
4.5 ANÁLISE DE SUPERFÍCIE ELETROSTÁTICA .....	46
<b>5 CONSIDERAÇÕES FINAIS</b> .....	<b>48</b>
5.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS .....	48
<b>REFERENCIAS</b> .....	<b>50</b>

## 1 INTRODUÇÃO

A regulação gênica pós-transcricional é constituída de três pontos principais: processamento, tradução e degradação do mRNA. No processamento do mRNA são realizadas modificações para remoção de íntrons, adição da cauda poli-A na extremidade 3' e da estrutura cap na extremidade 5', para formar o mRNA maduro. Já a degradação de mRNA é sinalizada a partir da ligação de pequenos RNAs como miRNAs e siRNA. Por fim, o processo de tradução, que comumente é dividido em quatro etapas distintas - iniciação, alongação, finalização e reciclagem -, sendo a primeira a mais extensivamente regulada, envolvendo doze fatores de iniciação (eIFs). Dentre estes fatores está o eIF4E, proteína ligadora do cap do mRNA, que compõe o complexo de iniciação da tradução eIF4F, cujo papel é auxiliar na montagem do ribossomo 80S competente para a alongação. A formação do complexo eIF4F constitui-se um dos pontos relevantes de controle da tradução e é um dos objetivos globais de estudo do projeto. Isso se deve à presença de seis homólogos EIF4E(1-6) em tripanosomatídeos, característica distinta da grande maioria dos organismos eucariotos.

O papel na regulação da iniciação da tradução de cada um destes homólogos ainda é pouco conhecido. Embora haja na literatura dados que visam caracterizar estas proteínas *in vivo* e *in vitro*, informações precisas sobre suas funções específicas ainda são escassas. Além disso, a caracterização estrutural destes fatores, principalmente ligados ao análogo do *cap-4* presente em tripanosomatídeos, contribui para o entendimento dos mecanismos envolvidos e especificidades no reconhecimento do mRNA durante a iniciação da tradução. Diferenças estruturais nestes fatores podem levar a diferentes mecanismos e especificidades de reconhecimento do mRNA e de proteínas parceiras na formação dos complexos de iniciação e, portanto, a diferentes mecanismos de regulação nestes protozoários.

Sabendo disso, se torna essencial a resolução destas estruturas por técnicas como cristalografia de proteínas e criomicroscopia eletrônica, a fim de gerar dados relevantes que possam ser usados para inferir informações sobre sua função biológica e mecanismos moleculares. E, com isso, elucidar o processo da regulação da iniciação da tradução em tripanosomatídeos.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo geral

Determinar a estrutura cristalográfica do fator de iniciação da tradução eIF4E5 de *Trypanosoma brucei* (TbEIF4E5), e realizar sua comparação com as estruturas experimentais já resolvidas dos fatores eIF4E de *Homo sapiens*, eIF4E5 de *Trypanosoma cruzi* e eIF4E5 de *Leishmania major*.

### 1.1.2 Objetivos específicos

i) Realizar o refinamento da estrutura de eIF4E5 de *Trypanosoma brucei* a partir de dados de difração de raio X da proteína cristalizada, previamente coletados.

ii) Comparar globalmente a estrutura refinada com as homólogas de *Homo sapiens* e de dois tripanosomatídeos, *T. cruzi* e *L. major*.

iii) Comparar o sítio de interação de eIF4E5 com o *cap-4* do mRNA, de TbEIF4E5 com as homólogas de *H. sapiens* e de *T. cruzi* e *L. major*.

iv) Comparar a superfície eletrostática de TbEIF4E5 com as homólogas de *H. sapiens*, *T. cruzi* e *L. major*, na região de interação com o *cap-4*.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 TRADUÇÃO DE PROTEÍNAS EM EUCARIOTOS

O processo de tradução consiste na síntese de proteínas por meio da leitura do mRNA transcrito. É subdividido em quatro etapas: iniciação, alongação, finalização e reciclagem, respectivamente. O objetivo da iniciação é a montagem das subunidades maior (60S) e menor (40S) para formar o ribossomo competente para a alongação (80S) (JACKSON RJ; HELLEN CUT; PESTOVA TV, 2010). Mediante o reconhecimento do códon de iniciação (AUG) do mRNA, presente no iniciador aminoacil-tRNA, começa o alongamento da sequência peptídica mediada pela atividade dos fatores eIF1A, eIF2, eIF3, eIF5A (revisões em MARINTCHEV et al., 2004; GOODFELLOW, I. G.;

ROBERTS, L. O, 2008; MARINTCHEV et al., 2009; DEVER TE; DINMAN JD; GREEN R, 2018; HERSHEY, J. W. B.; SONENBERG, N.; MATHEWS, M. B, 2012). Após esta etapa há a finalização, em que ocorre o reconhecimento dos códons de parada (UAA, UAG e UGA) no sítio A do ribossomo, orquestrada pelos fatores eRF1 e eRF3 que promovem a liberação de proteínas recém-sintetizadas. Por fim, ocorre a reciclagem do ribossomo, com a dissociação deste do mRNA, seguido pelas duas subunidades da organela que são desmontadas para que sejam utilizadas por pelo menos mais oitenta novos ciclos de tradução (HELLEN CUT, 2018; YOUNG DJ; GUYDOSH NR, 2022).

### 2.1.1 INICIAÇÃO DA TRADUÇÃO

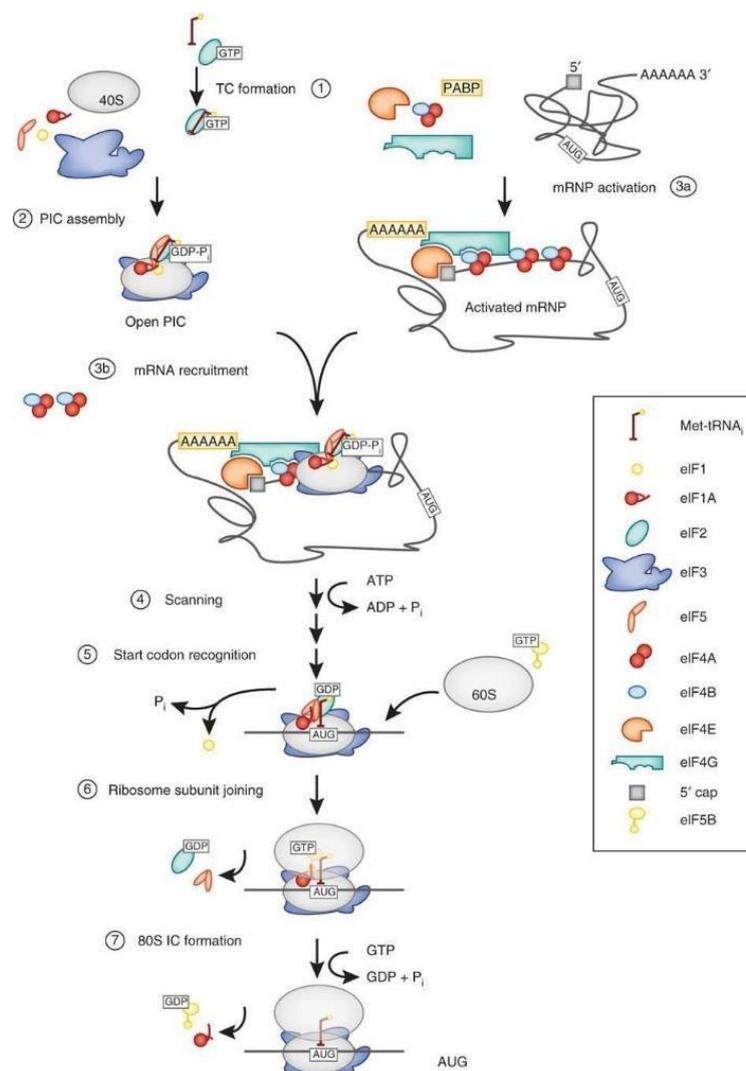
A iniciação da tradução (Figura 1) é caracterizada como a etapa de maior regulação dentre as quatro fases da síntese de proteínas. Isso se deve à presença de doze fatores (eIF4s) responsáveis por coordenarem a montagem do ribossomo 80S (AITKEN CE; LORSCH JR, 2012).

O processo inicia com a formação do complexo ternário (TC, do inglês ternary complex), composto pelo fator eIF2, GTP e met-tRNA. Seu papel é garantir que o iniciador met-tRNA seja entregue ao sítio P do ribossomo para que ocorra adequadamente o reconhecimento do códon de iniciação (HRONOVÁ V; VALÁŠEK LS, 2017). Este é então recrutado pela subunidade menor do ribossomo (40S), que em complexo com os fatores eIF3, eIF1, eIF1A e eIF5, forma o complexo de pré-iniciação (PIC, do inglês pre-initiation complex) da tradução (43S), numa região próxima ao cap do mRNA. Estes fatores possuem papéis distintos nesta etapa de regulação. eIF3 forma uma ponte de conexão entre componentes que se ligam a entrada e saída do mRNA na subunidade 40S do ribossomo, para coordenar a montagem do PIC. Os fatores eIF1 e eIF1A ligam-se cooperativamente na subunidade menor do ribossomo em sua conformação aberta para estabilizar o PIC. Já eIF5, proteína ativadora de GTPase (GAP, do inglês GTPase-activating protein), hidrolisa a molécula de GTP presente no complexo ternário (TC) a GDP, propiciando a dissociação de fatores de iniciação para que a subunidade maior 60S do ribossomo seja acoplada.

O PIC em sua conformação aberta, juntamente com o mRNP ativado, ou seja, o mRNA transcrito maduro, associado com os fatores eIF4G, eIF4B, eIF4A

e eIF4E (se liga ao cap do mRNA), formam a estrutura capaz de iniciar o escaneamento do mRNA. O fator eIF4B é uma proteína ligadora de mRNA que possui o domínio RRM (motivo de reconhecimento de RNA). Ela interage com mRNA para facilitar a ligação do ribossomo nesta molécula, atuando como um facilitador da formação do PIC 43S.

Então, após o reconhecimento do códon AUG pela estrutura formada mencionada, há a liberação de eIF1 e a subunidade 60S do ribossomo é recrutada para formar o ribossomo 80S, competente para a elongação. Posteriormente, há a hidrólise do GTP pelo fator eIF5B que promove umrearranjo conformacional na estrutura montada do ribossomo, juntamente com a dissociação de eIF4A, o que torna o ribossomo 80S preparado para o ciclo de alongamento da cadeia peptídica (revisões em GRAY NK; WICKENS M, 1998; GINGRAS AC; RAUGHT B; SONENBERG N, 1999; JACKSON RJ; HELLEN CU; PESTOVA TV, 2010; AITKEN CE; LORSCH JR, 2012).



**Figura 1 - Representação esquemática do mecanismo de iniciação da tradução em eucariotos (retirado de AITKEN et al., 2012).** (1) A iniciação começa pela formação do complexo ternário contendo eIF2-GTP e o tRNA iniciador. (2) O complexo é associada-se à subunidade 40S com a participação dos fatores eIFs 1, 1A, 3 e 5 para formar o PIC (pre-initiation complex). (3a) Em paralelo o mRNA é ligado aos fatores eIF4E e PABP (poly-A binding protein) formando um mRNP ativado. Uma vez ligado à extremidade 5' do mRNA, o complexo PIC realiza a varredura até a localização do códon de iniciação (AUG) (4). (5) O reconhecimento do códon de iniciação desencadeia a liberação de eIF1 e a conversão de eIF2 ao seu estado ligado à GDP, cessando o processo de varredura. (6) eIF2-GDP e eIF5 se dissociam e a subunidade 60 associa-se para formar o complexo de iniciação 80S. Os diferentes componentes dos complexos são identificados no retângulo à direita.

### 2.1.2 COMPLEXO EIF4F

O complexo eIF4F é formado durante a fase de iniciação e participa da montagem do PIC 43S. Isso ocorre durante a formação da estrutura apta para o escaneamento do mRNA, em que estão presentes as estruturas do mRNP ativado, junto dos fatores eIF4B, eIF4A, eIF4G e eIF4E, mas ainda sem a subunidade maior do ribossomo (revisões em GRAY NK; WICKENS M, 1998; GINGRAS AC; RAUGHT B; SONENBERG N, 1999; JACKSON RJ; HELLEN CU; PESTOVA TV, 2010; AITKEN CE; LORSCH JR, 2012). Estes três últimos fatores são os que compõe o complexo (MERRICK WC, 2015), descoberto em reticulócitos de coelho, em um experimento que visou compreender como as células estudadas iriam maximizar a síntese proteica na ausência do fator eIF4B (GRIFO JA et al., 1983).

O fator eIF4A é uma proteína helicase dependente de ATP. Ela contém o motivo DEAD box (sequência de resíduos Asp-Glu-Ala-Asp), envolvido na catálise do desenovelamento de mRNA dependente de ATP (ANDREOU AZ; KLOSTERMEIER D, 2013), além do motivo de reconhecimento de RNA (RRM, do inglês RNA recognition motif). Sendo assim, a função principal deste fator da tradução é desenovelar as estruturas secundárias do mRNA na região 5' não traduzidas (5'-UTR), de maneira que o PIC realize o escaneamento do mRNA para que o códon de início seja encontrado (ROGERS GW Jr; KOMAR AA; MERRICK WC, 2002).

eIF4G atua como uma plataforma de ancoragem para fatores eIFs. A região de ligação ao fator eIF4E foi mapeada no N terminal da proteína (LAMPHEAR BJ et al., 1995), enquanto o domínio conservado MIF4G (do inglês middle domain of eukaryotic initiation factor 4G) está envolvido na interação com a helicase eIF4A (VIRGILI G et al., 2013). Além deste, na região C terminal existe um sítio adicional de ligação de eIF4A (IMATAKA H, SONENBERG N, 1997) e de ligação do fator eIF3 (LAMPHEAR BJ et al., 1995).

Por fim, o fator eIF4E, que se liga ao cap do mRNA. Esta proteína está envolvida com a regulação da iniciação da tradução e com o controle de crescimento celular (SONENBERG N; GINGRAS AC, 1998). Durante a fase da iniciação realiza interação com o fator eIF4G, além de sofrer mudanças conformacionais no bolsão de interação com o cap, que permitem a estabilização

do complexo eIF4F ativo para a tradução. Outro parceiro de interação é o fator 4E-BP, abundantes em condições de crescimento, cujo estado majoritário é hiperfosforilado. No entanto, quando as células não estão crescendo ativamente e, portanto, com menor necessidade de síntese proteica, os fatores 4E-BP são desfosforilados, o que permite sua ligação ao 4E. Sendo assim, a associação destes fatores produz complexos de tradução inativos (revisão em VON DER HAAR T et al., 2004).

### 2.1.3 PECULIARIDADES DA INICIAÇÃO DA TRADUÇÃO EM TRIPANOSOMATÍDEOS

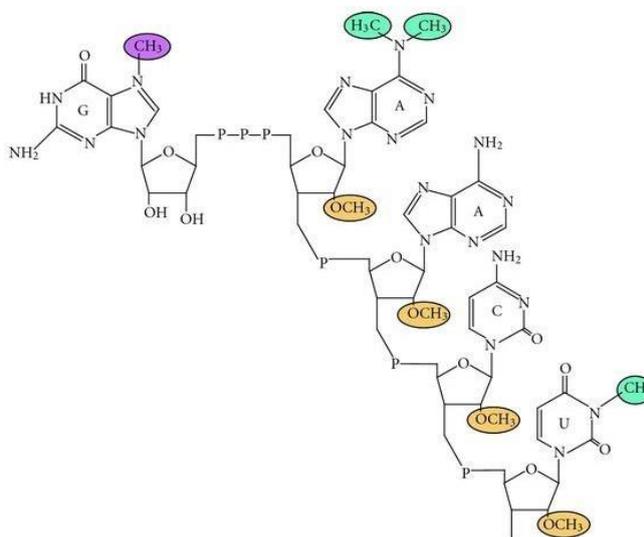
Tripanossomatídeos possuem características únicas na regulação de expressão gênica, grande parte realizada através de mecanismos pós-transcricionais (revisões em CLAYTON, 2002; CLAYTON; SHAPIRA, 2007; CLAYTON, 2019).

#### 2.1.3.1 CAP DO mRNA

O cap é uma estrutura composta de nucleotídeos modificados m<sup>7</sup>Gppp(n)-ligado (ROTTMAN F, SHATKIN AJ, PERRY RP, 1974) que desempenha um papel crucial na estabilização do mRNA junto ao ribossomo (SHIMOTOHNO K et al., 1977) como também impede a degradação do mRNA nascente (FURUICHI YASUHIRO, 1977) durante a etapa de transcrição do DNA realizada pela Polimerase II. Essa modificação na extremidade 5' do transcrito primário, juntamente com a remoção dos íntrons da sequência - que para os tripanosomatídeos ocorre pelo processo de *trans-splicing* (LIANG XUE-HAI et al., 2003; MICHAELI S, 2011) - e a adição da cauda poli A na extremidade 3', são necessários para o processamento desta molécula, de maneira que seja formado o mRNA maduro.

Em especial, a extremidade 5' do mRNA maduro é requerida para a tradução de proteínas (MUTHUKRISHNAN S et al., 1975), sendo dependente da metilação que ocorre na posição sete da guanosina (BOTH GW; BANERJEE AK; SHATKIN AJ, 1975; NIEDZWIECKA A et al., 2002). Os tripanosomatídeos possuem um análogo do cap distinto da grande maioria dos eucariotos, o chamado *cap-4*, ao invés do clássico m<sup>7</sup>GTP (BOTH GW et al., 1975; SHATKIN AJ, 1976). Este possui estrutura bastante particular pois além do m<sup>7</sup>GTP na

extremidade 5', comum aos eucariotos, o *cap-4* de tripanossomatídeos possui uma sequência de nucleotídeos AACU hipermodificada, com metilações 2'-O nas riboses dos quatro nucleotídeos e metilações nas bases dos nucleotídeos 1 e 4 (Figura 2) (BANGS JD et al., 1992).



**Figura 2: Representação esquemática da estrutura *cap-4* presente em tripanossomatídeos (retirado de ZINOVIEV A et al., 2012).**

### 2.1.3.2 FATORES DE INICIAÇÃO DA TRADUÇÃO EIF4Es e EIF4Gs

Diferente da grande maioria dos eucariotos, os tripanossomatídeos possuem seis genes conhecidos que codificam para seis distintos fatores eIF4Es, inclusive em *Trypanosoma brucei* (BERRIMAN M et al., 2005).

O primeiro fator descrito foi eIF4E1 de *Leishmania major* (YOFFE Y et al., 2004), também identificado junto de mais três fatores homólogos de *L. major*, e conservados nos demais tripanossomatídeos (DHALLIA R et al., 2005). Após estes, o quinto (FREIRE ER et al., 2014a) e sexto fator identificado também foram descritos (FREIRE ER et al., 2014b). Estes seis fatores homólogos têm sido separados em categorias distintas. eIF4E1 e eIF4E2 constituem o primeiro

grupo, já eIF4E3 e eIF4E4 o segundo grupo e eIF4E5 e eIF4E6 o terceiro e último grupo. As similaridades apresentadas por estes grupos de fatores de iniciação residem nas características estruturais das proteínas devido a sequência de aminoácidos e composição de N e C terminais das proteínas, mas também características funcionais (FREIRE ER et al., 2011).

Além disso, hoje também já são conhecidos cinco homólogos distintos para os fatores eIF4Gs, que ao se complexarem com os diferentes homólogos de 4Es formam diferentes complexos, tanto estruturalmente quanto funcionalmente. Estes já foram descritos para os tripanosomatídeos *T. brucei* (FREIRE ER et al., 2014a; FREIRE ER et al., 2014b) e *L. major* (DHALLIA R et al., 2005; YOFFE Y et al., 2006; YOFFE Y et al., 2009; DE MELO NETO OP et al., 2015; SHRIVASTAVA R et al., 2021).

### 2.1.3.3 CARACTERIZAÇÃO DOS FATORES EIF4Es

A primeira caracterização dos fatores 4Es de tripanosomatídeos foi realizada para *LmEIF4E1*, *LmEIF4E2* e *LmEIF4E3* na fase de vida promastigota do parasita. Buscou-se determinar como se dava a interação destes fatores com o cap do mRNA através de ensaios de *pull down* utilizando uma coluna de sepharose com 7-metil-GTP imobilizado. Os resultados mostraram que 4E1 se liga ao cap, diferente de 4E2 e 4E3. Além disso, para 4E1 e 4E2 constatou-se que seria improvável que estes fatores sozinhos conseguissem regular a iniciação da tradução na fase de vida do parasita no qual os estudos foram realizados. Em paralelo, a modelagem molecular de *LmEIF4E1* evidenciou características estruturais mantidas nesta proteína que realmente permitiriam sua ligação ao cap (DHALLIA R et al., 2005). E, quando foi gerado mutantes pelo sistema CRISPR-Cas9, verificou-se alterações na morfologia em *Leishmania*. As células não formaram a morfologia comum a fase de vida promastigota, além de formarem o flagelo defeituosamente, o que culminou em uma menor capacidade do parasita em infectar macrófagos cultivados (TUPPERWAR N; SHRIVASTAVA R; SHAPIRA M, 2019).

Posteriormente, foram modelados o fator *LmEIF4E1* e *LmEIF4E(2-4)* que também mostraram a conservação do sítio de interação das proteínas com o cap. Além disso, ensaios de interação com colunas de 7-metil-GTP-Sepharose

mostraram que das quatro homólogas estudadas apenas *LmEIF4E2* não interagiu com a resina. Ensaio quantitativo foram feitos utilizando a técnica de fluorimetria que mostraram que a afinidade pelo análogo do cap m7GTP é  $Lm4E1 > Lm4E3 > Lm4E4 > Lm4E2$ . No entanto, *Lm4E2* apresentou alta afinidade pelo *cap-4* (YOFFE Y et al., 2006), contrariando resultados obtidos anteriormente (DHALIA R et al., 2005), mas em concordância com dados obtidos quanto a ligação deste fator a proteína histona de mRNA (SLBP, do inglês stem-loop-binding protein). Este foi determinado por ensaios de *pull down*, cujo resultado indica uma possível função distinta da participação na regulação da iniciação da tradução pois o papel de eIF4E2 pode residir na seleção diferencial de mRNAs que possuem estruturas SLBP (FREIRE ER et al., 2018). Outro estudo verificou que células knock-out para eIF4E2 se multiplicaram lentamente, além de serem estabilizadas pela proteína PUF9 (TUPPERWAR N et al., 2021), envolvida na estabilização de determinados transcritos durante a fase S do ciclo celular de *T. brucei* (ARCHER SK et al., 2009).

Para o fator *Lm4E3*, foi identificada abundante ligação no análogo do cap m7GTP em contraste com o análogo do *cap-4*, indicando, inicialmente, que possivelmente este homólogo não participaria da iniciação da tradução em tripanosomatídeos (YOFFE Y et al., 2006). Acrescenta-se a este dado, a identificação de interação entre este mesmo fator eIF4E3, mas de *brucei*, com eIF4G3 e eIF4G4, indicando um possível papel diferente desta homóloga na tradução (FREIRE ER et al., 2011).

Para eIF4E4 constatou-se dependência da fase de vida promastigota de *major*, para efetiva formação do complexo eIF4F, já que em amastigotas não há ligação da proteína ao cap (ZINOVIEV A et al., 2011). Outro dado relevante é quanto a importância deste homólogo em fases de crescimento de *brucei*, que se dá pela fosforilação de resíduos da região N-terminal, bem como para eIF4E3 (PEREIRA MM et al., 2013). Posteriormente, os eventos de fosforilação foram melhor caracterizados, que ocorrem, inclusive, com melhor eficiência na presença da proteína de ligação a cauda poli(A) (PABP, do inglês poly(A)-binding protein), que interage em três domínios distintos da região N-terminal (DE MELO NETO OP et al., 2015). Esta proteína possui papel importante na promoção da circularização do mRNA durante a fase de tradução junto de fatores que compõem com o complexo eIF4F (WELLS SE et al., 1998).

Ainda, quanto a dependência da atividade de alguns fatores pelo estágio de vida do tripanosomatídeo, foi mostrado que eIF4E1 de *Leishmania major* não interage, tanto na fase promastigota e amastigota, com as homólogas eIF4Gs, excluindo um possível envolvimento deste fator na iniciação da tradução dependente do cap (ZINOVIEV A et al., 2011).

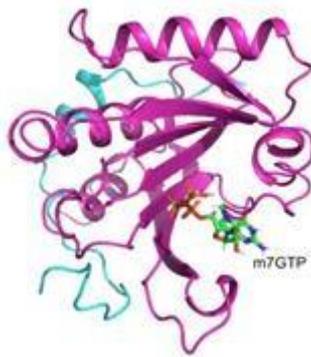
Foram encontrados complexos eIF4E5-eIF4G1 e também determinada a constante de dissociação pelo método de fluorimetria de eIF4E5 com o análogo do *cap-4* ( $K_{as} = 0,55 \pm 0,01 \mu\text{M}^{-1}$ ), curiosamente menor do que a constante de associação com o análogo do cap m7GTP ( $K_{as} = 0,12 \pm 0,01 \mu\text{M}^{-1}$ ) (FREIRE ER et al., 2014a). Também, foi determinado por ensaios de fluorimetria que TbeIF4E6 se liga ao análogo *cap-4*. Os valores das constantes de associação calculados relevaram baixa afinidade da proteína pelo análogo do cap m<sup>7</sup>GpppA mas esta foi maior para os análogos de m7GTP ( $K_{as} 0,16 \pm 0,1 \mu\text{M}^{-1}$ ) e *cap-4* ( $K_{as} 0,16 \pm 0,2 \mu\text{M}^{-1}$ ). Adicionalmente, foi verificado que esta homóloga possui papel na motilidade do parasita (FREIRE ER et al., 2014b). Em *Leishmania* foi identificado que o fator eIF4E6 é importante para a morfologia do parasita na fase promastigota. A interação deste fator com a homóloga eIF4G5, foi determinada por ensaios de *pull down* (TUPPERWAR N et al., 2021). Também, foi determinado em *T. brucei* sua importância para o crescimento do parasita na fase sanguínea tripomastigota, bem como para tradução de proteínas (FALK F et al., 2022).

## 2.2 ESTRUTURA DOS FATORES DE INICIAÇÃO DA TRADUÇÃO EIF4Es

Os fatores eIF4E's possuem seu *core* principal bem descrito, com a conservação de sete a oito estruturas secundárias beta antiparalelas e três a quatro alfa hélices. Além disso, sua estrutura é comparada à uma ferradura, devido à presença da face côncava, em que ocorre o reconhecimento do *cap* do mRNA (Figura 3). Esta foi revelada pela primeira estrutura do fator determinada, cuja região de interação com o *cap* possui característica hidrofílica e com superfície eletrostática carregada positivamente (MARCOTRIGIANO, J. et al, 1997).

Existem resíduos importantes e conservados nas proteínas da família eIF4E's que participam da interação com o *cap*. Dois resíduos aromáticos (W56

e W102 em eIF4E de humanos) fazem interações do tipo  $\pi$ - $\pi$  *stacking*, com o anel m<sup>7</sup>G do *cap*, em uma interação chamada de “sanduíche”. Um resíduo de ácido glutâmico (E103 de humanos) estritamente conservado nos fatores de iniciação 4E's estabiliza esta ligação através de duas pontes salinas através de sua cadeia lateral. Na região dos trifosfatos do *cap* há uma interação atípica 5'-5' entre o m<sup>7</sup>guanossina e o próximo nucleosídeo. Devido à carga negativa dos fosfatos há a presença majoritária de resíduos básicos que realizam interações do tipo ponte salinas. Por fim, na região das bases nitrogenadas o tipo de interação predominante é a ligação de hidrogênio (MARCOTRIGIANO, J. et al, 1997; MATSUO, H. et al, 1997; TOMOO, K. et al, 2002).



**Figura 3: Estrutura global do fator de iniciação eIF4E (rosa) de humanos em complexo com um fragmento de eIF4G (ciano) e o m<sup>7</sup>GTP (Depósito PDB 5T46) (GRÜNER S et al., 2016).**

Já foram determinadas três estruturas de fatores de iniciação eIF4Es de tripanosomatídeos. A primeira foi a estrutura do fator eIF4E1 de *major* em complexo com a proteína 4E-IP1, ainda com função desconhecida, mas que possui ortólogos em diferentes tripanosomatídeos (MELEPPATTU S et al., 2018). Após esta, foi resolvido o fator eIF4E5 de *cruzi* em complexo com o análogo do *cap*-4 do mRNA, revelando detalhes da interação e mecanismo de ligação com o ligante (Figura 4), importante ponto de regulação da iniciação da tradução (REOLON LW et al, 2019). A mais recente descoberta foi a estrutura de eIF4E5 de *major* (DE LIMA GB et al., 2021).

Então, a partir da estrutura de *T. cruzi* foram identificados resíduos característicos das proteínas da família 4E's, como os dois aromáticos que



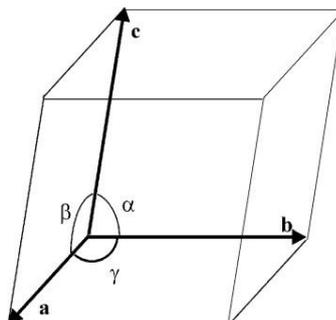
terminal e purificada em três etapas, primeiro por cromatografia de afinidade, seguida de cromatografia de troca iônica e, por fim, cromatografia de exclusão por tamanho. Os ensaios de cristalização foram realizados pelo método *sitting drop* a 18 °C utilizando kits comerciais que incluem uma grande variedade de condições. As condições mais promissoras foram otimizadas e os melhores cristais foram obtidos na condição 0,04 M Imidazol/ 0,06 M MES pH 6,5, 0,06 M divalentes (MgCl<sub>2</sub>/CaCl<sub>2</sub>), 18% PEG mix (40% PEG 500 MME/ 20%PEG20000).

### 3.2 COLETA E PROCESSAMENTO DOS DADOS DE DIFRAÇÃO

A difração de raio X em cristais de proteínas é uma técnica importante na determinação da estrutura tridimensional destas moléculas. Ela se baseia em raios X, emitidos por uma fonte em um comprimento de onda específico, que interagem com a nuvem de elétrons presente nos átomos das proteínas cristalizadas, o que leva à difração dos raios X em padrões característicos. Estes são usados para a determinação da sua estrutura tridimensional (SMYTH MS; MARTIN JH, 2000).

Cristais são estruturas tridimensionais periódicas formadas por arranjos ordenados de íons, moléculas ou átomos constituintes (AMEH ES, 2019). As propriedades de simetria dos cristais são governadas por leis estabelecidas por Christian Weiss (1780-1856), que identificou sete possíveis sistemas de simetria cristalina, levando em consideração as quatro operações fundamentais: rotação, translação, reflexão e inversão (WASEDA Y; MATSUBARA E; KOZO S, 2011).

A menor unidade estrutural dos cristais é conhecida como célula unitária. Esta é caracterizada por seis parâmetros distintos: três comprimentos de aresta representados pelos vetores *a*, *b*, *c* e três ângulos  $\alpha$ ,  $\beta$ ,  $\gamma$  formados entre esses vetores, conforme ilustrado na Figura 5 (PARKER M., 2003). Os valores dos vetores da célula unitária do cristal obtido para a proteína eIF4E5 de *brucei* estão descritos na Tabela 2.



**Figura 5: Representação da definição dos parâmetros de uma célula unitária.** *a*, *b* e *c* correspondem ao tamanho das arestas da célula e os ângulos  $\alpha$ ,  $\beta$ , e  $\gamma$  presentes entre elas. Fonte: Parker M., 2003.

A partir da cristalização de uma proteína, várias das mesmas moléculas estarão dispostas de maneira padronizada em células unitárias, seguindo um eixo de simetria que advém do empacotamento estruturado dessas moléculas no cristal (CHAYEN NE; SARIDAKIS E, 2008; McPHERSON A; GAVIRA, JA, 2014). Essa organização padronizada está diretamente relacionada com a resolução obtida dos dados, pois uma menor mosaicidade - pequenas imperfeições ou desordens na estrutura cristalina, que podem resultar em ligeiras variações nas orientações dos planos cristalinos devido a defeitos estruturais, tensões internas ou desalinhamentos nas camadas atômicas - aumenta a capacidade de difração de raios X (DURBIN SD; FEHER G, 1996). O limite de resolução dos dados de difração, obtidos do cristal de TbE14E5 conta na Tabela 2.

Os cristais são classificados nos sete sistemas cristalinos, uma classificação atribuída a René Just Haüy (1743-1822). Pertencentes aos sete sistemas cristalinos, existem quatorze células unitárias diferentes descritas por Auguste Bravais (1811-1863) e são conhecidas como Redes de Bravais (HELLIWELL J, 1992). Cada célula unitária possui características específicas de simetria, incluindo o número e a direção dos eixos de simetria presentes, o que caracteriza o grupo espacial ao qual o cristal pertence, que descrevem completamente a simetria do cristal. A Tabela 1 descreve as características dos sistemas cristalinos e as redes de Bravais associadas.

Sistema	Comprimentos e ângulos axiais	Redes de Bravais	Símbolo da rede
Cúbico	$a = b = c, \alpha = \beta = \gamma = 90^\circ$	Simplex	<i>P</i>
		Corpo-centrado	<i>I</i>
		Face-centrada	<i>F</i>
Tetragonal	$a \neq b \neq c, \alpha = \beta = \gamma = 90^\circ$	Simplex	<i>P</i>
		Corpo-centrado	<i>I</i>
Ortorrômbico	$a \neq b \neq c, \alpha = \beta = \gamma = 90^\circ$	Simplex	<i>P</i>
		Corpo-centrado	<i>I</i>
		Base-centrada	<i>C</i>
		Face-centrada	<i>F</i>
Trigonal	$a = b \neq c, \alpha = \beta = 90^\circ, \gamma = 120^\circ$	Simplex	<i>R</i>
	$a = b = c, \alpha = \beta = \gamma \neq 90^\circ$ (romboédrico)		
Hexagonal	$a = b \neq c, \alpha = \beta = 90^\circ, \gamma = 120^\circ$	Simplex	<i>P</i>
Monoclínico	Três eixos desiguais, um par não em ângulos retos $a \neq b \neq c, \alpha = \beta = 90^\circ, \gamma \neq 90^\circ$	Simplex	<i>P</i>
		Base centrada	<i>C</i>
Triclínico	$a \neq b \neq c, \alpha \neq \beta \neq \gamma \neq 90^\circ$	Simplex	<i>P</i>

**Tabela 1: Os sete sistemas cristalinos, valores de suas arestas e ângulos, classificação na Rede de Bravais e símbolo correspondente.** Fonte: WASEDA Y; MATSUBARA E; KOZO S, 2011.

A coleta de dados do cristal da proteína de interesse foi realizada na linha de luz PROXIMA2A do Síncrotron SOLEIL (França) pela Dra. Beatriz Gomes Guimarães. A difração de raios X por cristais resulta em um padrão de difração decorrente da interferência das ondas eletromagnéticas espalhadas por todos os átomos do cristal. A distribuição e intensidade dos pontos de difração depende do conteúdo e simetria da célula unitária do cristal. De uma maneira simplificada, a coleta de dados consiste na medida do padrão de difração, ou seja, na detecção das intensidades dos feixes difratados nas várias direções, quando o cristal é exposto aos raios X em diferentes orientações.

Os dados de difração foram processados utilizando-se o programa XDS (KABSCH W., 2010). O processamento envolve três etapas principais: indexação, integração e escalonamento. A indexação consiste em atribuir “coordenadas” ou índices  $h k l$  aos pontos de difração. Nesta etapa, através de uma análise de distâncias entre os pontos de difração podem ser determinados os parâmetros de rede e o sistema cristalino ao qual o cristal pertence. A integração dos dados consiste na estimativa de intensidade para cada um dos pontos de difração ou “reflexões” medidas. A atribuição da intensidade a cada ponto de difração  $h k l$  envolve a aplicação de fatores de correção aos valores medidos pelo detector de fótons utilizado no experimento. Uma vez atribuídas as intensidades a todas as reflexões, estas são comparadas, a fim de se detectar padrões de simetria. A simetria do cristal é refletida no padrão de difração, dando origem a conjuntos de reflexões de “mesma” intensidade, as chamadas reflexões equivalentes. Por fim, na etapa de escalonamento são realizadas correções dos dados, através da comparação das intensidades de reflexões equivalentes. Nesta etapa é também calculado o conjunto de reflexões únicas, que reúne as reflexões equivalentes e calcula as médias de intensidades e erros associados. Através da análise das reflexões equivalentes e suas relações de simetria é determinado o grupo espacial correspondente àquele conjunto de dados.

### 3.3 DETERMINAÇÃO E REFINAMENTO DA ESTRUTURA

O refinamento de estruturas tridimensionais de proteínas refere-se ao processo de modificar o modelo através do ajuste do backbone da proteína e das cadeias laterais dos resíduos, visando obter modelos que se aproximem do estado “nativo” da molécula (HEO L; PARK H; SEOK C, 2013; ADIYAMAN R; MCGUFFIN LJ, 2019). A construção e ajuste do modelo são realizados com base no mapa de densidade eletrônica, calculado a partir dos fatores de estrutura, que por sua vez são obtidos dos dados de difração de raios X do cristal de interesse e um conjunto inicial de fases provenientes, por exemplo, do método de substituição molecular.

O fator de estrutura,  $F_{hkl}$ , representa a descrição matemática da somatória das ondas difratadas pelo cristal após o fenômeno físico de interferência construtiva. A equação do fator de estrutura é descrita como:

$$F(\mathbf{h}) = \sum f(hkl) \exp(2\pi i(hx + ky + lz))$$

onde  $f(hkl)$  é o fator de espalhamento do átomo.  $h$   $k$   $l$  são os índices de Miller, obtidos para cada ponto do mapa de difração, na etapa de indexação do processamento dos dados de difração de raios X.

A função densidade eletrônica é a transformada de Fourier do Fator de Estrutura, descrita pela equação:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum F(\mathbf{h}) \exp(i\alpha) \exp(-2\pi i \mathbf{h} * \mathbf{r})$$

onde  $V$  corresponde ao volume da célula unitária do cristal analisado.  $F(\mathbf{h})$  é o módulo de  $F_{hkl}$ .  $\alpha$  sua fase estimada.

O mapa de densidade eletrônica é a representação gráfica tridimensional da função densidade eletrônica. (LAMB, A. L.; KAPPOCK, T. J.; SILVAGGI, N. R, 2015). Esta descreve a distribuição de elétrons presente em cada átomo na célula unitária do cristal que contribuem para a intensidade no padrão de difração, onde o modelo de interesse será ajustado.

Sendo assim, inicialmente, foi gerado no Alphafold Collab um modelo do fator de iniciação da tradução eIF4E5 de *brucei* a partir da sequência da proteína depositada no banco de dados TriTrypDB (código de identificação Tb927.10.5020) Este modelo foi utilizado como modelo inicial para a aplicação do método de substituição molecular, que fornece fases iniciais para o cálculo do primeiro mapa de densidade eletrônica. Um primeiro ciclo de ajustado modelo é feito utilizando-se o programa COOT (EMSLEY; COWTAN, 2004). Assim, foram realizados ajustes e otimizações para melhorar a sobreposição entre as coordenadas atômicas do modelo e o mapa de densidade, respeitando os parâmetros estereoquímicos da estrutura. Posteriormente, iniciaram-se os ciclos de refinamento no espaço recíproco, utilizando as novas coordenadas atômicas presentes no arquivo pdb, juntamente com o arquivo mtz contendo os dados experimentais. Esses ciclos de refinamento foram executados nos programas BUSTER ou Phenix. Como resultado desse refinamento, as coordenadas atômicas atualizadas foram importadas novamente para o COOT,

junto do arquivo mtz vindo do novo ciclo de refinamento, sucessivas vezes até que o modelo de estrutura tridimensional da proteína eIF4E5 de *T. brucei* apresentasse bons índices de refinamento.

O critério adotado para decidir se um ciclo de refinamento seria continuado foi baseado no valor do "Free R value" (Rfree), que é um indicador de validação da estrutura que reflete a concordância dos dados experimentais ( $F_{obs}$ ) com o modelo ( $F_{calc}$ ) (BRÜNGER, A. T, 1992). É calculado pela diferença do F observado, calculado com os dados experimentais obtidos da difração de raios X e posteriormente processados, em relação ao F calculado, calculado com os parâmetros advindos do modelo.

Dessa forma, se o valor do Rfree diminuísse em relação ao ciclo anterior, o refinamento prosseguia. Por outro lado, se o valor do Rfree aumentasse, retornava-se ao ciclo anterior para testar novos ajustes no mapa de densidade. O objetivo era sempre obter um valor de Rfree mais baixo do que o apresentado no ciclo anterior. Essa abordagem foi adotada para garantir a melhoria contínua da qualidade da estrutura, levando em consideração a concordância com os dados experimentais e os critérios geométricos dos resíduos da proteína.

### 3.4 ANÁLISES ESTRUTURAIS

A análise estrutural foi realizada através da sobreposição dos modelos experimentais das homólogas eIF4E5 dos organismos *T. brucei* (este trabalho), *T. cruzi* (PDB 6O7Y), *L. major* (PDB 7KCJ) e eIF4E *H. sapiens* (PDB 5T46). Uma primeira análise global da estrutura buscou detalhar o tamanho das regiões N terminais, C terminais e loops de conexão entre elementos de estrutura secundária conservados, com fim de comparação entre os fatores de iniciação da tradução. Além disso, a análise local da região de interação destas proteínas com o cap do mRNA também foi realizada, tanto pela descrição dos resíduos de aminoácidos envolvidos, quanto pela superfície eletrostática de cada um dos fatores.

### 3.4.1 ANÁLISE ESTRUTURAL INDIVIDUAL

A partir do arquivo com as coordenadas atômicas, gerado como output do último ciclo de refinamento no programa Phenix (D. LIEBSCHNER, et al., 2019), o modelo experimental de TbEIF4E5, em complexo com o análogo *cap-4* do mRNA foi analisado nos programas Chimera (PETTERSEN et al., 2004), Pymol (THE PYMOL MOLECULAR GRAPHICS SYSTEM, SCHRÖDINGER, LLC) e *Protein-Ligand Interaction Profiler* (PLIP) (ADASME F MELISSA et al., 2021), com o objetivo realizar uma análise global da proteína e uma análise local do sítio de interação com o ligante.

Sendo assim, para a análise global foi realizado o detalhamento de quais são e quais as extensões das estruturas secundárias fitas beta e alfa hélices presentes na proteína através do programa Chimera. A ferramenta utilizada foi a de “Sequência”, em que são reconhecidas e coloridas tais estruturas secundárias na sequência da proteína alvo, com os respectivos números dos resíduos de aminoácidos.

Para a análise local, foi detalhado como se dá a interação do fator *TbEIF4E5* com o ligante, determinando quais são os resíduos participantes e ligações mediadas por eles, através de suas cadeias laterais e/ou cadeias principais. Esta etapa foi realizada no programa PLIP, em que foi utilizado como input o arquivo de coordenadas atômicas do modelo experimental, obtido do processo de refinamento, e como output foi gerado um arquivo com as coordenadas dos resíduos da proteína participantes da interação com o *cap-4*, bem como tipos de ligações associadas entre estes aminoácidos eo ligante, para posterior visualização e interpretação no programa Pymol.

### 3.4.2 ANÁLISE ESTRUTURAL COMPARATIVA

A partir das sobreposições estruturais, buscou-se identificar os resíduos, nas diferentes homólogas, que possivelmente estariam envolvidos na ligação com o *cap-4*. Para isso, a estrutura cristalográfica de eIF4E5 de *T. cruzi* (PDB 6O7Y), determinada em complexo com o *cap-4* (REOLON LW et al., 2019), foi utilizada como base comparativa através de alinhamento estrutural com os

demais modelos experimentais de interesse. Assim, foram feitas análises globais quanto à conservação do core principal da proteína, tamanho do N e C terminal e regiões de loop (Tabela 3) e, também, análises locais de resíduos de aminoácidos que possivelmente participariam da ligação no sítio de interação dos fatores eIF4E com o *cap-4* (Tabela 4). Foram consideradas as características físico-químicas dos resíduos e posições equivalentes da estrutura utilizada com referência.

### 3.4.3 ANÁLISE DE SUPERFÍCIE ELETROSTÁTICA

Com o objetivo de realizar uma análise comparativa da distribuição de cargas na região de interação com o *cap-4*, os modelos tridimensionais foram utilizados para os cálculos dos mapas de potencial eletrostático. Para isso, a preparação dos destes envolveu a remoção do *cap-4* nos quais está presente, *T. cruzi* (PDB: 6O7Y) e *T. brucei* (dados ainda não publicados).

Assim, para a geração das superfícies eletrostáticas, os arquivos de coordenadas atômicas correspondentes a cada estrutura foram utilizados como input para o programa PDB2PQR (DOLINSKY TJ et al., 2004 e DOLINSKY TJ et al., 2007), que prepara as estruturas para os cálculos posteriores das propriedades eletrostáticas, reconstruindo átomos, adicionando hidrogênios, atribuindo cargas atômicas e raios de campos de força especificados, gerando como output um arquivo de extensão pqr utilizado pelo programa Adaptive Poisson-Boltzmann Solver (APBS) (JURRUS E. et al., 2018). Este resolve a equação de Poisson-Boltzmann para o cálculo do potencial eletrostático em torno da estrutura de interesse. O arquivo de saída gerado é utilizado para colorir as superfícies moleculares de acordo com seu “mapa” eletrostático. Tal representação é gerada utilizando-se o programa Chimera (PETTERSEN et al., 2004). Também, considerando que o modelo experimental de *Leishmania major* não possui ligante, e como o objetivo da análise é avaliar a superfície de interação com o *cap-4*, foi realizada a sobreposição da estrutura com o ligante, advindo tanto da estrutura resolvida de *T. cruzi* (PDB 6O7Y) quando da estrutura resolvida de *T. brucei*, utilizando a ferramenta de sobreposição estrutural MatchMaker também no programa Chimera.

## 4 APRESENTAÇÃO DOS RESULTADOS

### 4.1 PROCESSAMENTO DOS DADOS E REFINAMENTO

As estatísticas dos dados de difração e refinamento da estrutura são apresentadas na Tabela 2. O cristal de *TbEIF4E5* possui célula unitária pertencente ao sistema cristalino trigonal com grupo espacial  $P3_121$ . Este difratou em alta resolução (1.35 Å). Foram obtidos bons valores de  $R_{work}$  e de  $R_{free}$  no refinamento, mostrando boa fidelidade do modelo. A completeza dos dados processados, que é definida pelo número de reflexões cristalográficas coletadas em comparação com o número de reflexões teoricamente possíveis únicas para a simetria do cristal analisado, foi de 100%. O número médio de medições por reflexão individual simetricamente única, chamado de multiplicidade, também foi bastante elevado, garantindo boa acurácia na estimativa das intensidades, o que é refletido em outro indicador de qualidade dos dados, calculada pelo coeficiente de correlação de Pearson  $CC\ \frac{1}{2}$ . Valores de  $CC_{1/2}$  superiores a 40% na última camada de resolução são considerados bastante satisfatórios.

**Tabela 2.** Dados cristalográficos e estatística do refinamento. Os valores entre parêntesis correspondem à última camada de resolução.

<b><i>Estatística dos dados</i></b>	
Fonte	SOLEIL-PX2A
Comprimento de onda (Å)	0.978565
Resolução (Å)	47.04 – 1.35 (1.43- 1.35)
Grupo de espaço	P3 <sub>1</sub> 21
Célula unitária (Å)	a=b= 72.44 c=94.07
Número de observações	1260557 (202892)
Número de reflexões únicas	63204 (10040)
Completeza (%)	99.9 (99.5)
Multiplicidade	19.9 (20.2)
I / $\sigma$ (I)	25.5 (1.2)
CC ½	100.0 (51.6)
<b><i>Estatísticas do refinamento</i></b>	
R <sub>work</sub>	0.1773
R <sub>free</sub>	0.2020
Comprimento de ligação RMSD (Å)/ Ângulo (°)	0.014/1.490
Valor B médio (global) (Å <sup>2</sup> )	22.24
Átomos da proteína	1544
Átomos do solvente	211
<b><i>Ramachandran</i></b>	
Favoráveis (%)	98.21%
Outliers (%)	0%

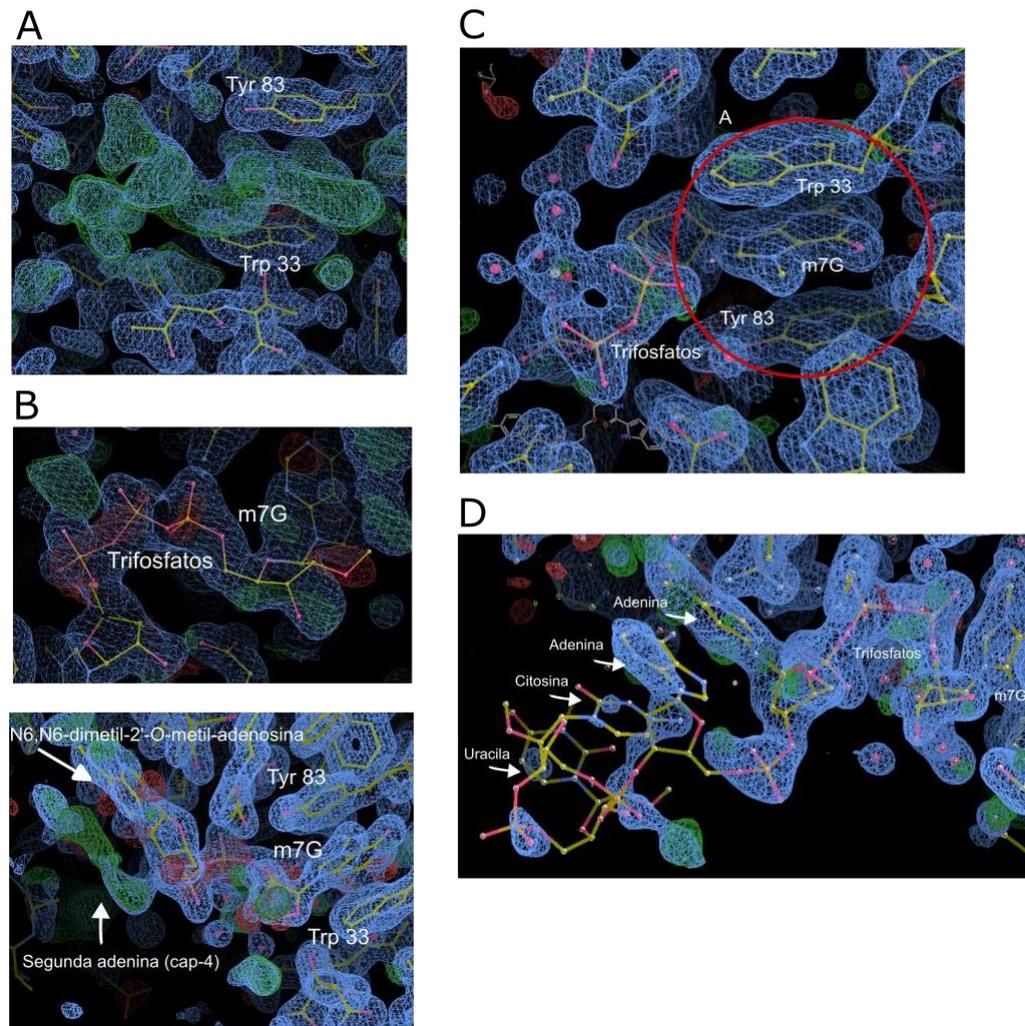
#### 4.2 REFINAMENTO DA ESTRUTURA DE EIF4E5 DE *Trypanosoma brucei*

Durante a construção iterativa do modelo através do programa COOT, o mapa de densidade eletrônica indicou claramente a presença do *cap-4*, co-cristalizado com a proteína *TbEIF4E5* (Figura 6A). A conformação do ligante foi

otimizada através de vários ciclos de refinamento, que evidenciou claramente o posicionamento do fragmento m<sup>7</sup>GTP (Figura 6B). Por outro lado, observa-se flexibilidade conformacional a partir do fosfato que conecta a primeira e segunda adenosinas, indicada pela ausência de densidade eletrônica definida correspondendo aos três últimos nucleotídeos. A flexibilidade é provavelmente consequência da possibilidade de diferentes rotâmeros para o grupo fosfato e pelo fato que os três últimos nucleotídeos do ligante não são estabilizados por interações com a proteína. O modelo do ligante (Figura 6C-D) foi refinado com valor de ocupação de 0.5 para os átomos dos dois últimos nucleotídeos.

Todos os resíduos da proteína puderam ser modelados, inclusive àqueles localizados em loops mais longos que conectam estruturas secundárias. Também, devido à alta resolução dos dados, puderam ser modeladas duplas e até mesmo triplas conformações para uma mesma cadeia lateral de alguns resíduos, em especial cadeias laterais de serinas.

Na modelagem do solvente foram adicionados íons de magnésio, um deles estabilizando o trifosfato do *cap-4*, além de íons cálcio, cloro e moléculas de *polietilenoglicol* (PEG), advindos das condições de cristalização da proteína. O modelo final inclui 174 moléculas de água.

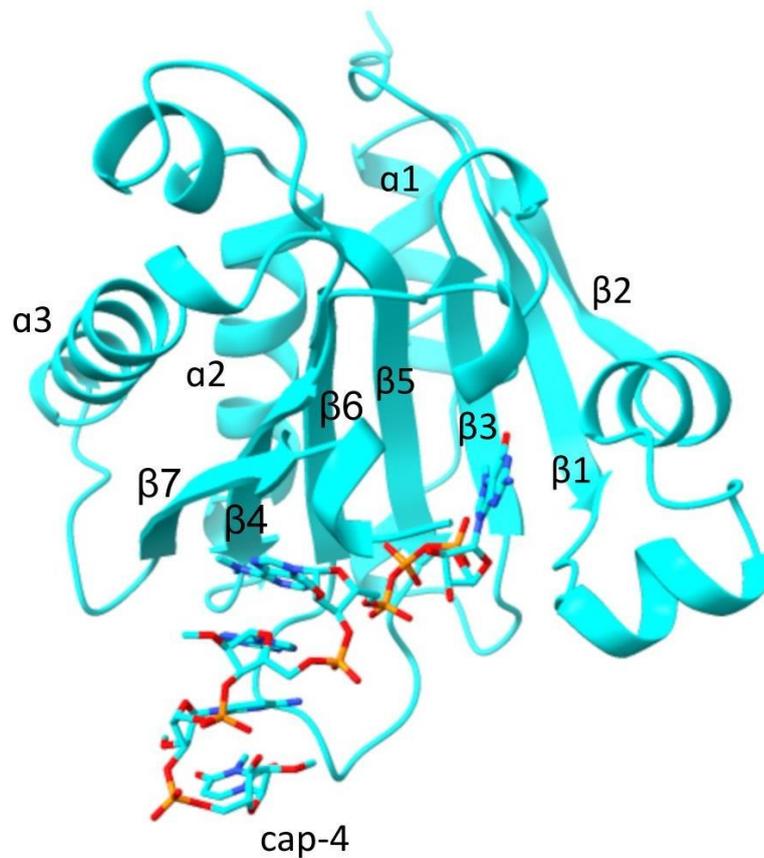


**Figura 6: Mapa de densidade eletrônica ilustrando a modelagem do ligante cap4.** O mapa 2Fo-Fc é mostrado em azul e contornado a 0.98 sigma enquanto o mapa Fo-Fc é mostrado em verde e contornado a 2.5 sigma. (A) Densidade eletrônica positiva do ligante evidenciada em verde. (B) m<sup>7</sup>GTP posicionado na densidade eletrônica, com aparente densidade da segunda base adenina do análogo do cap-4 (C) Mapa de densidade eletrônica após refinamento do ligante, destacando os resíduos que foram a ligação tipo “sanduíche” com o m7G e o grupo trifosfato. (D) Detalhe dos quatro nucleotídeos (AACU) que compõe o cap-4. Observa-se que não há densidade eletrônica correspondente aos dois últimos.

#### 4.3 ANÁLISE ESTRUTURAL INDIVIDUAL

A estrutura *TbeIF4E5* mantém as características da família de proteínas 4E's, com a presença de sete fitas betas e três extensas

alfa-hélices, como apresentado na Figura 7. Além disso, fitas e hélices curtas adicionais são observadas (Tabela 3).



**Figura 7:** Representação do modelo cristalográfico do fator de iniciação da tradução eIF4E5 de *Trypanosoma brucei*. O ligante *cap-4* é mostrado em bastões.

Fitas $\beta$	RESÍDUOS	Alfa hélices	RESÍDUOS
$\beta 1$	W 12 - P 18	$\alpha 1'$	T 22 - N 28
$\beta 2$	Q 42 - V 47	$\alpha 2'$	W 33 - L 40
$\beta 3$	S 70 - A 75	$\alpha 1$	V 50 - S 59
$\beta 4$	T 90 - I 96	$\alpha 3'$	L 63 - Q 65
$\beta 5$	C 129 - S 136	$\alpha 4'$	R 76 - N 78
$\beta 6$	V 146 - L 152	$\alpha 5'$	Y 83 - A 85
$\beta 7$	V179 - L 185	$\alpha 2$	A 99 - I 114
-	-	$\alpha 6'$	S 117 - V122
-	-	$\alpha 3$	Q 155 - V 172
	-	$\alpha 7'$	F 186 - E 189

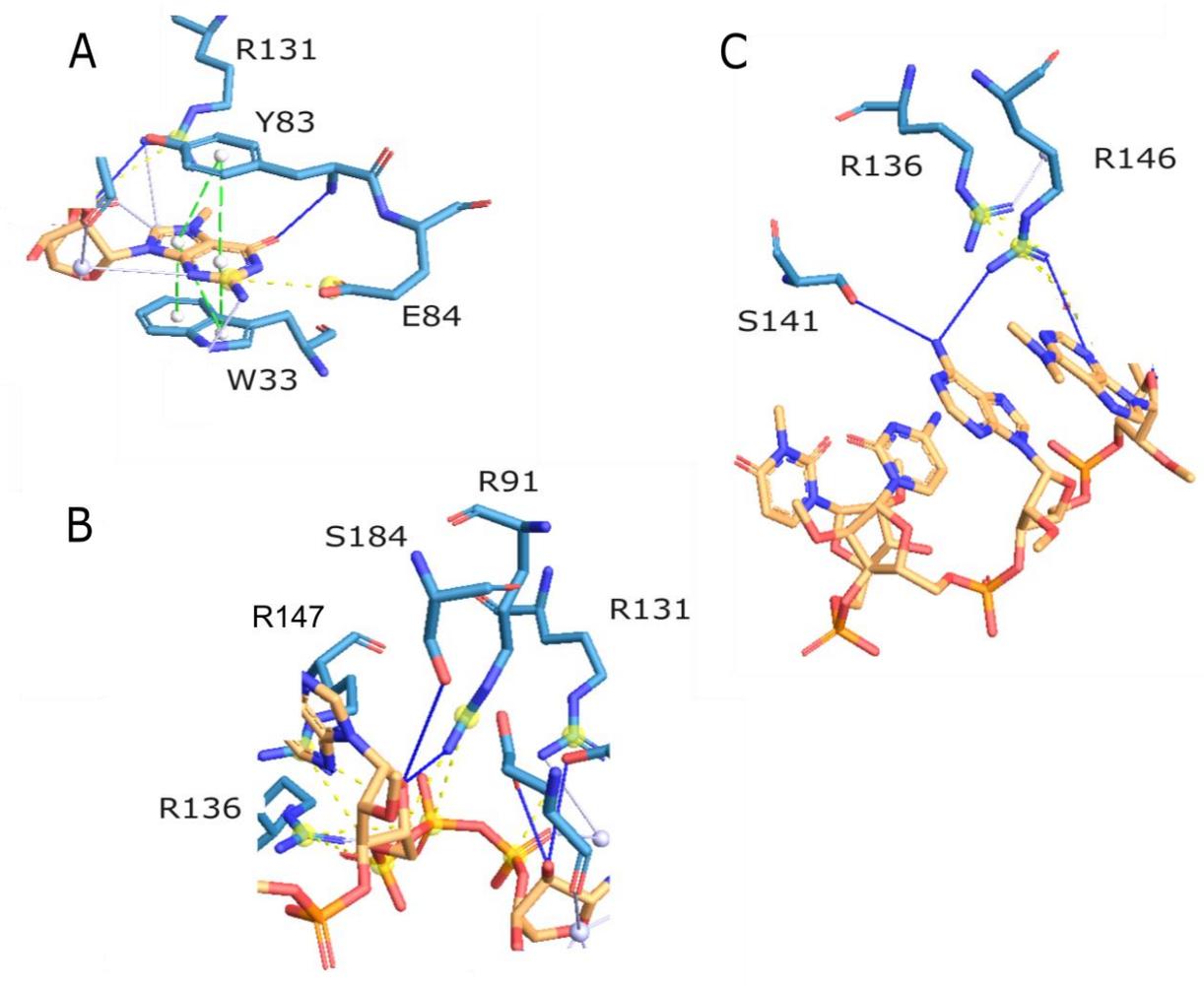
**Tabela 3: Folhas beta e alfa hélices presentes na estrutura de *TbEIF4E5*, resíduos onde se inicia e termina cada uma das estruturas secundárias presentes.** As estruturas secundárias denotadas com apóstrofo possuem um número significativamente menor de resíduos quando comparado às demais presentes na proteína.

A análise detalhada da região de interação com o *cap-4* mostra a formação do *stacking* (sanduíche) com o m<sup>7</sup>G, através dos resíduos aromáticos W33 e Y83, além da interação com o resíduo de ácido glutâmico (E84), estritamente conservado nos fatores eIF4E de tripanosomatídeos e de outros eucariotos. Tal resíduo realiza duas pontes salinas através de sua cadeia lateral, com os dois átomos de nitrogênio carregados presentes na guanina que compõem o *cap-4* do mRNA (Figura 8A).

Além disso, há ligações que ocorrem na região do grupo de trifosfatos, presente na ligação atípica 5'-5' entre o m<sup>7</sup>GTP e a primeira adenina do *cap-4*. Estas ocorrem através dos resíduos básicos de argininas, cuja interação se dá

majoritariamente por pontes salinas e também por ligações de hidrogênio (Figura 8B).

Por fim, curiosamente, apenas as duas primeiras bases nitrogenadas presentes na estrutura do *cap-4* participam de interação com a proteína, mantendo a citosina e uracila livres (Figura 8C).



**Figura 8: Resíduos de *TbEIF4E5* envolvidos na interação com o análogo do *cap-4*.** Os carbonos pertencentes ao ligante estão coloridos em amarelo, já os em azul são os resíduos de aminoácidos da proteína que interagem diretamente com o *cap-4*. Interações mostrada em azul correspondem a ligações de hidrogênio, em verde pi-stacking, em amarelo pontes salinas e em cinza claro ligações de hidrogênio intermediadas por moléculas de água. Em A, detalha-se as interações envolvendo o grupo m7G, em B aquelas envolvendo o trifosfato enquanto em C são representadas as interações envolvendo os demais nucleotídeos.

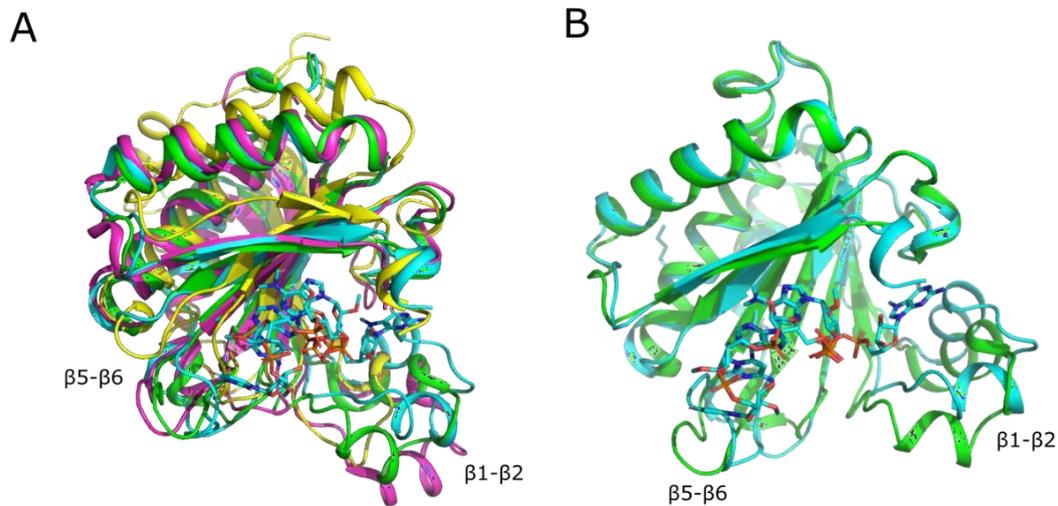
#### 4.4 ANÁLISE ESTRUTURAL COMPARATIVA

A sobreposição do modelo tridimensional de EIF4E5 de *T. brucei*, com as homólogas de *T. cruzi* e *L. major* e eIF4E de *Homo sapiens*, mostra a conservação do *core* característico desta família de proteínas (Figura 9). Os valores de desvio quadrático médio (RMSD, do inglês Root Mean Square Deviation) obtidos a partir dos alinhamentos estruturais refletem a conservação estrutural global das proteínas (Tabela 4).

	Nº CARBONOS REFERÊNCIA	Nº CARBONOS MOVIDOS	RMSD	Nº CARBONOS $\alpha$ ALINHADOS	IDENTIDADE DE SEQUÊNCIA
<b>Tc4E5</b>	183	187	1.6 Å	173	65.9%
<b>Lm 4E5</b>	176	187	2.5 Å	157	44.6 %
<b>Hs 4E</b>	213	187	1.97 Å	150	14.7 %

**Tabela 4: Dados referentes ao alinhamento estrutural de TbEIF4E5 com as proteínas homólogas de *T. cruzi*, *L. major* e eIF4E de *H. sapiens*.**

Apesar da conservação estrutural global, significativas diferenças locais podem ser observadas. Uma delas é no loop que conecta as fitas  $\beta 1$ - $\beta 2$ , onde está ancorado um dos resíduos aromáticos participantes da estrutura do stacking, o que justifica sua relevância (Figura 9). Para a estrutura de *major*, este mesmo loop está mais aberto quando comparado com as de *T. cruzi* e *brucei*, muito provavelmente devido ao fato de a estrutura não ter sido co-cristalizada com o *cap-4* do mRNA. Além disso, ainda que ambas as estruturas de EIF4E5 de *T. cruzi* e *T. brucei* tenham sido resolvidas em complexo com o ligante, a homóloga de *T. brucei* possui um loop mais fechado comparado com as demais. O mesmo ocorre para o loop que conecta as fitas  $\beta 5$ - $\beta 6$ , que possui menor extensão, sendo uma região importante devido ao ancoramento de resíduos que participam das interações com as bases nitrogenadas do *cap* (Figura 9).



**Figura 9 Alinhamento estrutural dos fatores de iniciação da tradução. A) Tb eIF4E5 (ciano), (dados ainda não publicados e Tc eIF4E5 (verde, PDB 5T46). B) Mesmos fatores que na Figura9A, junto de Lm eIF4E5 (rosa, PDB 7KCJ e de Hs eIF4E (amarelo, PDB 5T46).**

A partir da análise destas estruturas foi descrito o tamanho das regiões N e C terminal e conexões entre as fitas betas na Tabela 5. Os loops de conexão entre as fitas  $\beta 1-\beta 2$  possuem a mesma extensão para os três tripanosomatídeos, sendo mais longa quando comparada com a de *sapiens*, além de outro destaque para a região C terminal da homóloga de *major* que também possui maior extensão comparado com as demais estruturas.

ESPÉCIE / PROTEÍNA	N TERMINAL	C TERMINAL	Loop $\beta$ 1- $\beta$ 2	Loop $\beta$ 3- $\beta$ 4	Loop $\beta$ 5- $\beta$ 6
<i>Tb</i> 4E5	1-4 (4)	187-195 (9)	18-41 (24)*	76-89 (14)	137-145 (9)
<i>Tc</i> 4E5	1-11 (11)	178-186 (9)	18-41 (24)*	77-89 (13)	137-145 (9)
<i>Lm</i> 4E5	1-10 (10)	192-219 (28)*	22-45 (24)*	81-94 (14)	142-150 (9)
<i>Hs</i> 4E	1-36 (37)	199-217 (19)	49-59 (11)	96-110 (15)	156-161 (6)

**Tabela 5: Análise estrutural dos modelos experimentais quanto ao número de resíduos no N terminal, C terminal e loops  $\beta$ 1- $\beta$ 2,  $\beta$ 3- $\beta$ 4,  $\beta$ 5- $\beta$ 6 dos fatores de iniciação da tradução de *T. brucei* (Tb eIF4E5, dados ainda não publicados), *T. cruzi* (Tc eIF4E5, PDB 5T46), *L. major* (Lm eIF4E5, PDB 7KCJ) e *sapiens* (Hs eIF4E, PDB 5T46). Foram considerados N e C terminais as regiões que, ao alinhar o modelo da proteína com a estrutura experimental de *T. cruzi* eIF4E5 não possuíam alinhamento com o core principale conservado característico dos fatores de iniciação da tradução. Os valores entre parênteses correspondem ao número de resíduos, quando com asterisco evidencia diferença significativa de tamanho comparado às demais proteínas.**

A análise local foi realizada a fim de compreender e descrever, comparativamente, como ocorre a interação do análogo *cap-4* com as homólogas eIF4E5 dos três organismos, *T. brucei*, *T. cruzi* e *L. major*. Dessa forma, é possível revelar diferenças quanto à interação dos fatores com o ligante que podem implicar na afinidade destes pelo *cap-4* do mRNA para cada organismo.

Sendo assim, existem resíduos importantes e conservados na região de interação do *cap*-4 do mRNA com os fatores 4E's, descritos na Tabela 6, que mantêm a característica da estrutura da família dos fatores eIF4E's.

	<i>Tb4E5</i>	<i>Tc4E5</i>	<i>Lm4E5</i>	<i>Hs4E</i>
	TRP 33	TRP 33	TRP 37	TRP 56
STACKING	TYR 83	TYR 83	TYR 88	TRP 102
	GLU 84	GLU 84	GLU 89	GLU 103
GRUPO TRIFOSFATO	ARG 131	ARG 137	ARG 142	ARG 157
	ARG 91	ARG 91	ARG 96	ARG 112
	ARG 147	ARG 133	ARG 138	ASN 155
	ARG 136	ARG 147	ARG 142	LYS 159
	ARG 91	GLN 140	TRP 156	TRP 166
	ARG 147	ASN139	LEU 23	LYS 206
BASES DE NUCLEOTÍDEOS	SER 184	PRO 138	ALA 26	-
	-	ASN 68	-	-
	-	LEU 20	-	-
	-	ARG 137	-	-
	-	ARG147	-	-

**Tabela 6: Resíduos dos fatores de iniciação da tradução EIF4E5 de *T. cruzi* (*Tc4E5*), *T. brucei* (*Tb4E5*) e *L. major* (*Lm4E5*), envolvidos na interação com a estrutura cap do mRNA.**

Os dois resíduos aromáticos que reconhecem o m<sup>7</sup>GTP, presente no cap-4, e que formam a estrutura do *stacking* são mantidos nas homólogas dos quatro organismos. Embora o resíduo de triptofano 102, presente em eIF4E de *sapiens*, seja substituído pelo resíduo de tirosina nos três fatores homólogos eIF4E5 dos

tripanosomatídeos *T. brucei*, *T. cruzi* e *L. major*, a característica da interação do tipo *pi-stacking* com o anel da guanina permanece.

Além disso, a estrita conservação do resíduo de ácido glutâmico é confirmada ao analisar as interações do fator eIF4E5 de *T. brucei*, *T. cruzi* e *L. major* e *sapiens*, bem como a presença de resíduos básicos que interagem na região do grupo de trifosfatos. Por fim, as interações com as bases nitrogenadas do cap estão presentes em maior quantidade até as duas primeiras adeninas, pois a partir da citosina e uracila são quase inexistentes, como ocorre para a proteína de *T. brucei*, em que as estas ficam livres de qualquer ligação.

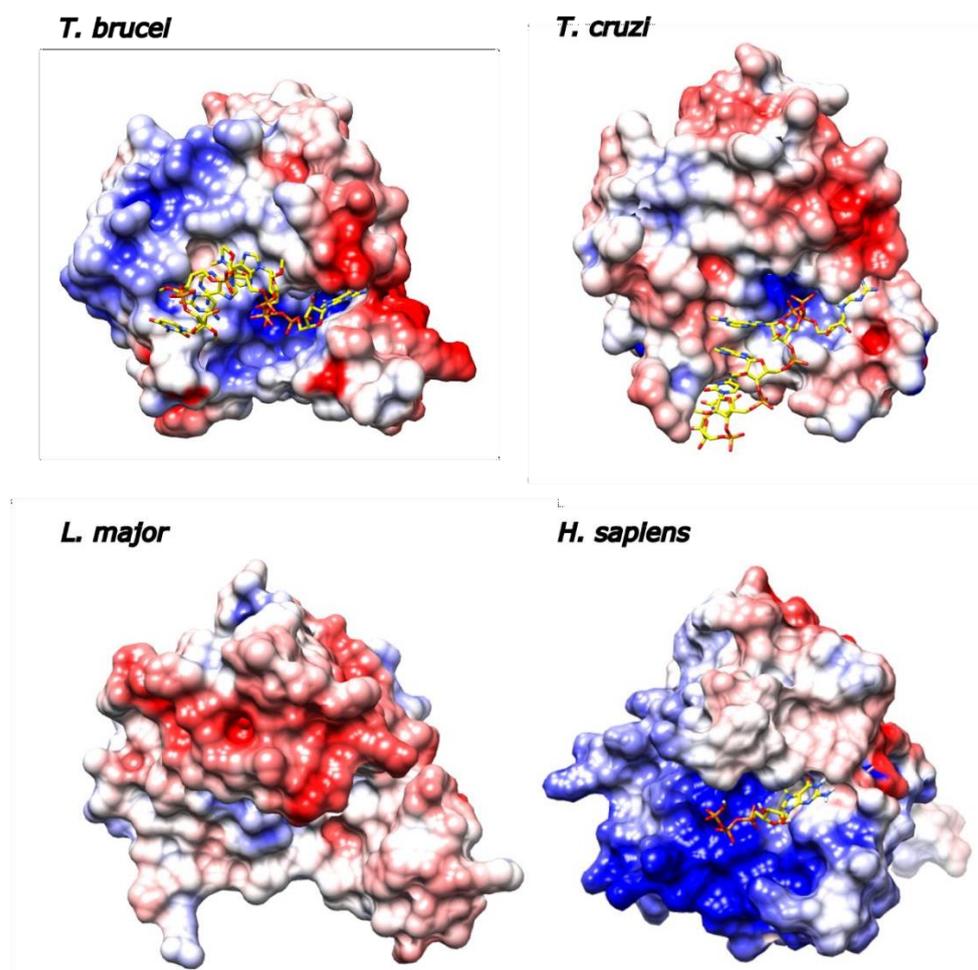
#### 4.5 ANÁLISE DE SUPERFÍCIE ELETROSTÁTICA

A análise das superfícies eletrostáticas (Figura 10) na face de interação com o *cap-4* do RNA evidencia diferenças significativas entre as três homólogas de tripanossomatídeos e destas com a eIF4E humana. As superfícies eletrostáticas dos quatro fatores de iniciação da tradução possuem, na região de interação com o análogo do cap, superfície carregada positivamente, especialmente na região de ancoramento do grupo trifosfatos que compõe o ligante. No entanto, em *LmEIF4E5* esta característica é distinta, com superfície predominante neutra. Além deste fator, *HsEIF4E* possui maior extensão da superfície positiva quando comparado aos demais três fatores de tripanossomatídeos.

Para *T. brucei*, há um deslocamento no posicionamento do *cap-4* comparado com o presente na estrutura de *TcEIF4E5*. Este ocorre, principalmente, na região em que estão localizadas as bases nitrogenadas, que compõe o ligante. Em *Tb4E5* estas bases estão voltadas para a proteína em uma região de superfície neutra e em *Tc4E5* estão voltadas para o solvente. Esta característica distinta entre estes fatores pode implicar na afinidade de ligação entre estas proteínas e o análogo do *cap-4*. Isso se deve ao papel do solvente em promover um ordenamento do posicionamento do cap, o que não ocorre para *Tb4E5*. Dessa maneira, é possível que existam outras conformações para o ligante, como foi observado durante o refinamento da estrutura, pela ausência de mapa de densidade eletrônica para as três últimas bases do cap.

Sendo assim, sugere-se que a interação com o *cap-4* é mediada principalmente pelo m7G e pelo grupo trifosfato, devido a carga negativa que apresenta. Contudo, existem diferenças significativas entre os fatores ortólogos analisados dos três tripanossomatídeos e do fator de iniciação de humanos, evidenciando possíveis diferenças na afinidade destas proteínas pela ligação ao

cap do mRNA. Como resultado, pode haver um possível mecanismo espécie-específico na interação das proteínas ortólogas EIF4E5 com análogos do cap. Além disso, análises dos demais fatores EIF4E1, 4E2, 4E3, 4E4 e 4E6, a partir de modelos gerados por predição de estrutura, indicam diferenças estruturais adicionais entre os diferentes homólogos de tripanossomatídeos (resultados não mostrados).



**Figura 10: Comparação das superfícies eletrostáticas dos fatores EIF4E5 de tripanossomatídeos e eIF4E de *H. sapiens*.** O análogo do *cap-4*, presente nos modelos de TbEIF4E5 e TcEIF4E5 (PDB 6O7Y), é representado em amarelo, em bastões, tal como o m7GTP, presente no modelo de HsEIF4E (PDB 5T46). Os limites para a visualização do potencial eletrostático são -5 (vermelho) a + 5 (azul) kT/e.

## 5 CONSIDERAÇÕES FINAIS

Os resultados apresentados neste trabalho descrevem como se dá o mecanismo de interação do fator de iniciação de tradução EIF4E5 de *Trypanosoma brucei* com o análogo *cap-4* de tripanossomatídeos.

Resíduos conhecidos e descritos na literatura pela sua conservação nas proteínas da família 4E's também estão presentes na estrutura resolvida, como os resíduos aromáticos que formam o stacking e o resíduo de ácido glutâmico que estabiliza esta estrutura. Além disso, a presença de resíduos básicos cujo papel é interagir com a ponte trifosfato que compõe o *cap*, em um sítio de interação carregado positivamente, sugere um possível mecanismo de ancoramento do mRNA. Ainda, curiosamente, com o processo do refinamento verificou-se grande flexibilidade do ligante pela ausência de mapa de densidade nas três últimas bases nitrogenadas, além da ausência de interações nas duas últimas com resíduos da proteína.

Então, embora haja características estruturais mantidas já descritas para a família de proteínas 4E's, inclusive no *core* da proteína, como a manutenção do número de fitas beta e alfa hélices, provavelmente devem haver mecanismos e especificidades distintas entre os seis fatores homólogos de *Trypanosoma brucei*, quanto para ortólogos de *Trypanosoma cruzi* e *Leishmania major* pelo análogo do *cap-4* do mRNA, constatado inclusive quando comparado o sítio de interação da proteína EIF4E de *Homo sapiens* em complexo com o m7GTP, como também para a estrutura resolvida do tripanossomatídeo *T. cruzi*.

### 5.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Embora tenha sido resolvida a estrutura do fator de iniciação EIF4E5 de *Trypanosoma brucei*, elucidando e gerando informações importantes quanto ao mecanismo bioquímico de reconhecimento do *cap-4* do mRNA, ainda são necessárias pesquisas com abordagem e métodos envolvendo Biologia Estrutural, para que seja compreendido e diferenciados os mecanismos das demais proteínas homólogas de *T. brucei*, já que apenas dois dos seis fatores foram resolvidos, e apenas EIF4E5 em complexo com o *cap-4*. Além disso, são

relevantes estudos que visem resolver a estrutura dos fatores ortólogos, como para os tripanosomatídeos mencionados neste trabalho, *T. cruzi* e *L. major*, como também para demais organismos eucariotos. Isto se deve a importância dos fatores de iniciação da tradução na regulação desta etapa da síntese proteica, essencial para a viabilidade celular de diversos organismos, inclusive vetores e agentes etiológicos de doenças importantes que ainda necessitam de drogas eficientes para tratamento e cura, cujos alvos podem residir nestes fatores. No entanto, não só a resolução das estruturas por cristalografia de proteínas, Ressonância Magnética Nuclear (RMN) e criomicroscopia eletrônica, por exemplo, especialmente complexadas com análogos do cap do mRNA, são recomendadas, mas também ensaios biofísicos que visam caracterizar estas proteínas *in vivo* e *in vitro*, funcionalmente e estruturalmente, inclusive em um contexto global da formação e mecanismos do complexo EIF4F, cujas subunidades EIF4G e EIF4A também possuem papéis relevantes, esta primeira, inclusive, com interação direta com os fatores EIF4E's.

Além disso, abordagens computacionais que envolvam dinâmica, *docking* e modelagem molecular são relevantes para gerar *insights* quanto ao mecanismo e especificidade dos fatores de iniciação pelos análogos do cap, bem como do complexo EIF4F, proteínas que o compõe e demais proteínas que participam da montagem do ribossomo, competente para a elongação, dos organismos de interesse.

## REFERÊNCIAS

- ADASME, M. F. et al. **PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA**. *Nucleic acids research*, v. 49, n. W1, p. W530–W534, 2 jul. 2021
- ADIYAMAN, R.; MCGUFFIN, L. J. **Methods for the Refinement of Protein Structure 3D Models**. *International Journal of Molecular Sciences*, v. 20, n. 9, 1 maio 2019a.
- AMEH, E. S. **A review of basic crystallography and x-ray diffraction applications**. *International Journal of Advanced Manufacturing Technology*, v. 105, n. 7–8, p. 3289–3302, 1 dez. 2019a.
- Andreou AZ, Klostermeier D. **The DEAD-box helicase eIF4A: paradigm or the odd one out?** *RNA Biol.*; v10(1). jan 2013.
- ARCHER, S. K. et al. ***Trypanosoma brucei* PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle**. *PLOS pathogens*, v. 5, n. 8, ago. 2009.
- AITKEN, C. E.; LORSCH, J. R. **A mechanistic overview of translation initiation in eukaryotes**. *Nature structural & molecular biology*, v. 19, n. 6, p. 568–576, jun. 2012.
- BANGS, J. D. et al. **Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides**. *Journal of Biological Chemistry*, v. 267, n. 14, p. 9805–9815, 15 maio 1992.
- BERRIMAN, M. et al. **The genome of the African trypanosome *Trypanosoma brucei***. *Science*, v. 309, n. 5733, p. 416–422, 15 jul. 2005.
- BOTH, G. W. et al. **Ribosome binding to reovirus mRNA in protein synthesis requires 5' terminal 7-methylguanosine**. *Cell*, v. 6, n. 2, p. 185–195, 1 out. 1975.
- BOTH, G. W.; BANERJEE, A. K.; SHATKIN, A. J. **Methylation-dependent translation of viral messenger RNAs in vitro**. *Proceedings of the National Academy of Sciences*, v. 72, n. 3, p. 1189–1193, 1 mar. 1975.
- BRÜNGER, A. T. **Free R value: a novel statistical quantity for assessing the accuracy of crystal structures**. *Nature* 1992 355:6359, v. 355, n. 6359, p. 472–475, 1992.
- CHAYEN, N. E.; SARIDAKIS, E. **Protein crystallization: from purified protein to diffraction-quality crystal**. *Nature Methods* 2008 5:2, v. 5, n. 2, p. 147–153, 30 jan. 2008.
- CLAYTON, C. **Regulation of gene expression in trypanosomatids: living with polycistronic transcription**. *Open biology*, v. 9, n. 6, 1 jun. 2019.

CLAYTON, C. E. NEW EMBO MEMBER'S REVIEW: **Life without transcriptional control? From fly to man and back again.** The EMBO Journal, v. 21, n. 8, p. 1881, 4 abr. 2002.

CLAYTON, C.; SHAPIRA, M. **Post-transcriptional regulation of gene expression in trypanosomes and leishmanias.** Molecular and Biochemical Parasitology, v. 156, n. 2, p. 93–101, 1 dez. 2007.

DE LIMA, G. B. et al. **The translation initiation factor EIF4E5 from Leishmania: crystal structure and interacting partners.** RNA biology, v. 18, n. 12, p. 2433–2449, 2021

DE MELO NETO, O. P. 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 biology, v. 12, n. 11, p. 1209–1221, 1 jan. 2015.

DHALIA, R. et al. **Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues.** Molecular & Biochemical Parasitology, v. 140, p. 23–41, 2005.

DURBIN, S. D.; FEHER, G. **PROTEIN CRYSTALLIZATION.** <https://doi.org/10.1146/annurev.physchem.47.1.171>, v. 47, p. 171–204, 1996.

DOLINSKY, T. J. et al. **PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations.** Nucleic acids research, v. 32, n. Web Server issue, 1 jul. 2004.

DOLINSKY, T. J. et al. **PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations.** Nucleic acids research, v. 35, n. Web Server issue, jul. 2007.

EMSLEY, P.; COWTAN, K. **Coot: model-building tools for molecular graphics.** urn:issn:0907-4449, v. 60, n. 12, p. 2126–2132, 26 nov. 2004.

EVANS, P.; MCCOY, A. **An introduction to molecular replacement.** Acta Crystallographica Section D: Biological Crystallography, v. 64, n. Pt 1, p. 1, 1 jan. 2008.

FALK, F. et al. **Roles and interactions of the specialized initiation factors EIF4E2, EIF4E5, and EIF4E6 in *Trypanosoma brucei*: EIF4E2 maintains the abundances of S-phase mRNAs.** Molecular microbiology, v. 118, n. 4, p. 457–476, 1 out. 2022.

FREIRE, E. R. et al. **The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties.** Molecular and biochemical parasitology, v. 176, n. 1, p. 25–36, 2011.

FREIRE, E. R. et al. ***Trypanosoma brucei* translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog.** Eukaryotic cell, v. 13, n. 7, p. 896–908, 2014a.

FREIRE, E. R. et al. ***Trypanosoma brucei* translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog.** Eukaryotic cell, v. 13, n. 7, p. 896–908, 2014b. Free.

FREIRE, E. R. 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, v. 20, n. 8, p. 1272–1286, 1 ago. 2014d.

FREIRE, E. R. et al. **The Role of Cytoplasmic mRNA Cap-Binding Protein Complexes in *Trypanosoma brucei* and Other Trypanosomatids.** Pathogens 2017, Vol. 6, Page 55, v. 6, n. 4, p. 55, 27 out. 2017.

FREIRE, E. R. et al. ***Trypanosoma brucei* EIF4E2 cap-binding protein binds a homolog of the histone-mRNA stem-loop-binding protein.** Current genetics, v. 64, n. 4, p. 821–839, 1 ago. 2018.

FURUICHI, Y.; LAFIANDRA, A.; SHATKIN, A. J. 5'-Terminal structure and mRNA stability. **Nature 1977 266:5599**, v. 266, n. 5599, p. 235–239, 1 mar. 1977.

GINGRAS, A. C.; RAUGHT, B.; SONENBERG, N. **eIF4 Initiation Factors: Effectors of mRNA Recruitment to Ribosomes and Regulators of Translation.** <https://doi.org/10.1146/annurev.biochem.68.1.913>, v. 68, p. 913–963, 1999.

GOODFELLOW, I. G.; ROBERTS, L. O. **Molecules in Focus Eukaryotic initiation factor 4E.** The international journal of biochemistry & cell biology, v. 40, n. 12, p. 2675, 2008.

GRAY, N. K.; WICKENS, M. **CONTROL OF TRANSLATION INITIATION IN ANIMALS.** <https://doi.org/10.1146/annurev.cellbio.14.1.399>, v. 14, p. 399–458, 1998.

GRIFO, J. A. et al. **New initiation factor activity required for globin mRNA translation.** Journal of Biological Chemistry, v. 258, n. 9, p. 5804–5810, 10 maio 1983.

H, I.; N, S. **Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A.** Molecular and cellular biology, v. 17, n. 12, p. 6940–6947, 1 dez. 1997.

HELLIWELL, J. R. **Macromolecular Crystallography with Synchrotron Radiation.** Macromolecular Crystallography with Synchrotron Radiation, 21 maio 1992.

HEO, L.; PARK, H.; SEOK, C. GalaxyRefine: **Protein structure refinement driven by side-chain repacking.** Nucleic acids research, v. 41, n. Web Server issue, 2013.

HERSHEY, J. W. B.; SONENBERG, N.; MATHEWS, M. B. **Principles of Translational Control: An Overview.** Cold Spring Harbor Perspectives in Biology, v. 4, n. 12, p. a011528–a011528, 1 dez. 2012.

HRONOVÁ, V.; VALÁŠEK, L. S. **An emergency brake for protein synthesis.** eLife, v. 6, 25 abr. 2017.

JACKSON, R. J.; HELLEN, C. U. T.; PESTOVA, T. V. **The mechanism of eukaryotic translation initiation and principles of its regulation.** Nature reviews. Molecular cell biology, v. 11, n. 2, p. 113–127, fev. 2010a.

JURRUS, E. et al. **Improvements to the APBS biomolecular solvation software suite.** Protein science: a publication of the Protein Society, v. 27, n. 1, p. 112–128, 1 jan. 2018.

KABSCH, W. **XDS.** Acta Crystallogr D Biol Crystallogr. urn:issn:0907-4449, v. 66, n. 2, p. 125–132, 22 jan. 2010

LAMB, A. L.; KAPPOCK, T. J.; SILVAGGI, N. R. **You are lost without a map: Navigating the sea of protein structures.** Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, v. 1854, n. 4, p. 258–268, 1 abr. 2015.

LAMPHEAR, B. J. et al. **Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation.** The Journal of biological chemistry, v. 270, n. 37, p. 21975–21983, 1995.

LIANG, X. H. et al. **trans and cis splicing in trypanosomatids: Mechanism, factors, and regulation.** Eukaryotic Cell, v. 2, n. 5, p. 830–840, out. 2003.

LIEBSCHNER, D. et al. **Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.** Acta crystallographica. Section D, Structural biology, v. 75, n. Pt 10, p. 861–877, 1 out. 2019.

MARCOTRIGIANO, J. et al. **Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP.** Cell, v. 89, n. 6, p. 951–961, 13 jun. 1997.

MARINTCHEV, A. et al. **Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation.** Cell, v. 136, n. 3, p. 447–460, 6 fev. 2009.

MARINTCHEV, A.; WAGNER, G. **Translation initiation: structures, mechanisms and evolution.** Quarterly Reviews of Biophysics, v. 37, n. 3–4, p. 197–284, nov. 2004.

MATSUO, H. et al. **Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein.** Nature structural biology, v. 4, n. 9, p. 717–724, 1997.

MCPHERSON, A.; GAVIRA, J. A. **Introduction to protein crystallization.** urn:issn:2053-230X, v. 70, n. 1, p. 2–20, 24 dez. 2014.

MELEPPATTU, S. et al. **Structural basis for LeishIF4E-1 modulation by an interacting protein in the human parasite *Leishmania major***. *Nucleic Acids Research*, v. 46, n. 7, p. 3791, 4 abr. 2018.

MERRICK, W. C. **eIF4F: A retrospective**. *Journal of Biological Chemistry*, v. 290, n. 40, p. 24091–24099, 2 out. 2015.

MICHAELI, S. **Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome**. *Future microbiology*, v. 6, n. 4, p. 459–474, abr. 2011.

MUTHUKRISHNAN, S. et al. **5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation**. *Nature* 1975 255:5503, v. 255, n. 5503, p. 33–37, 1975.

NIEDZWIECKA, A. et al. **Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins**. *Journal of molecular biology*, v. 319, n. 3, p. 615–635, 2002.

PARKER, M. W. **Protein Structure from X-Ray Diffraction**. *Journal of Biological Physics*, v. 29, n. 4, p. 341–362, 2003.

PEREIRA, M. M. C. 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**. *Molecular and biochemical parasitology*, v. 190, n. 2, p. 82–86, 2013.

PETTERSEN, E. F. et al. **UCSF Chimera--a visualization system for exploratory research and analysis**. *Journal of computational chemistry*, v. 25, n. 13, p. 1605–1612, out. 2004.

REOLON, L. W. et al. **Crystal structure of the *Trypanosoma cruzi* EIF4E5 translation factor homologue in complex with mRNA cap-4**. *Nucleic Acids Research*, v. 47, n. 11, p. 5973–5987, 20 jun. 2019.

ROGERS, G. W.; KOMAR, A. A.; MERRICK, W. C. **eIF4A: The godfather of the DEAD box helicases**. *Progress in Nucleic Acid Research and Molecular Biology*, v. 72, p. 307–331, 1 jan. 2002.

ROTTMAN, F.; SHATKIN, A. J.; PERRY, R. P. **Sequences containing methylated nucleotides at the 5' termini of messenger RNAs: Possible implications for processing**. *Cell*, v. 3, n. 3, p. 197–199, 1974a.

SHATKIN, A. J. **Capping of eucaryotic mRNAs**. *Cell*, v. 9, n. 4, p. 645–653, 1 dez. 1976.

SHIMOTOHNO, K. et al. **Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis**. *Proceedings of the National Academy of Sciences*, v. 74, n. 7, p. 2734–2738, 1 jul. 1977.

SMYTH, M. S.; MARTIN, J. H. J. **x Ray crystallography**. *Molecular Pathology*, v. 53, n. 1, p. 8, 2000.

SONENBERG, N.; GINGRAS, A. C. **The mRNA 5' cap-binding protein eIF4E and control of cell growth.** *Current Opinion in Cell Biology*, v. 10, n. 2, p. 268–275, 1 abr. 1998.

TOMOO, K. et al. **Crystal structures of 7-methylguanosine 5'-triphosphate (m(7)GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5',5'-triphosphate (m(7)GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region.** *The Biochemical journal*, v. 362, n. Pt 3, p. 539–544, 15 mar. 2002.

TUPPERWAR, N.; SHRIVASTAVA, R.; SHAPIRA, M. **LeishIF4E1 Deletion Affects the Promastigote Proteome, Morphology, and Infectivity.** *mSphere*, v. 4, n. 6, 18 dez. 2019b.

TUPPERWAR, N. et al. **Characterization of an Atypical eIF4E Ortholog in Leishmania, LeishIF4E-6.** *International journal of molecular sciences*, v. 22, n. 23, 1 dez. 2021.

VIRGILI, G. et al. **Structural Analysis of the DAP5 MIF4G Domain and Its Interaction with eIF4A.** *Structure*, v. 21, n. 4, p. 517–527, abr. 2013.

VON DER HAAR, T. et al. **The mRNA cap-binding protein eIF4E in post-transcriptional gene expression.** *Nature Structural & Molecular Biology* **2004 11:6**, v. 11, n. 6, p. 503–511, 25 maio 2004.

WASEDA, Y.; MATSUBARA, E.; SHINODA, K. **X-Ray Diffraction Crystallography.** *X-Ray Diffraction Crystallography*, 2011.

WELLS, S. E. et al. **Circularization of mRNA by eukaryotic translation initiation factors.** *Molecular cell*, v. 2, n. 1, p. 135–140, 1998.

YOFFE, Y. et al. **Cap-binding activity of an eIF4E homolog from Leishmania.** *RNA*, v. 10, n. 11, p. 1764–1775, 1 nov. 2004.

YOFFE, Y. et al. **Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania.** *Eukaryotic cell*, v. 5, n. 12, p. 1969–1979, dez. 2006.

YOFFE, Y. et al. **Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions.** *Nucleic Acids Research*, v. 37, n. 10, p. 3243–3253, 1 jun. 2009.

ZINOVIEV, A. et al. **A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways.** *Nucleic acids research*, v. 39, n. 19, p. 8404–8415, out. 2011.

ZINOVIEV, A.; SHAPIRA, M. **Evolutionary conservation and diversification of the translation initiation apparatus in trypanosomatids.** *Comparative and functional genomics*, v. 2012.