



Universidade Federal de Pernambuco
Centro de Biociências
Bacharelado em Biomedicina

SAULO RAFAEL MENDES PENNA

**Caracterização estrutural *in silico* das isoformas do fator
eIF4E em *Vigna unguiculata***

Recife
2024

SAULO RAFAEL MENDES PENNA

**Caracterização estrutural *in silico* das isoformas do fator
eIF4E em *Vigna unguiculata***

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Biomedicina da Universidade Federal de Pernambuco, como pré-requisito à obtenção do título de Bacharel em Biomedicina.

Orientadora: Ana Maria Benko Iseppon

Coorientador: Madson A. de Luna Aragão

Recife

2024

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Penna, Saulo Rafael Mendes.

Caracterização estrutural in silico das isoformas do fator eIF4E em *Vigna unguiculata* / Saulo Rafael Mendes Penna. - Recife, 2024.

101 : il., tab.

Orientador(a): Ana Maria Benko Iseppon

Coorientador(a): Madson Allan de Luna Aragão

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Biociências, Biomedicina, 2024.

Inclui referências, apêndices.

1. Feijão-caupi. 2. Fator de Tradução Eucariótico. 3. Fitopatógenos. 4. Modelagem Molecular. 5. Dinâmica Molecular. I. Iseppon, Ana Maria Benko. (Orientação). II. Aragão, Madson Allan de Luna. (Coorientação). IV. Título.

570 CDD (22.ed.)

SAULO RAFAEL MENDES PENNA

**Caracterização estrutural *in silico* das isoformas do fator
eIF4E em *Vigna unguiculata***

Trabalho de Conclusão de Curso
apresentado ao Curso de Graduação em
Biomedicina da Universidade Federal de
Pernambuco, como pré-requisito à
obtenção do título de Bacharel em
Biomedicina.

Aprovada em: 08/10/2024

BANCA EXAMINADORA

Orientador: Profa. Dra. Ana Maria Benko Iseppon
Departamento de Genética, Universidade Federal de Pernambuco

MSc. Ayug Bezerra Lemos
Departamento de Genética, Universidade Federal de Pernambuco

MSc. Wilson Dias de Oliveira
Departamento de Genética, Universidade Federal de Pernambuco

AGRADECIMENTOS

A realização desta monografia só foi possível graças ao apoio e incentivo de diversas pessoas e instituições, às quais manifesto minha profunda gratidão.

Primeiramente, agradeço a Deus, por me conceder força, sabedoria e discernimento ao longo desse percurso.

Agradeço de forma especial à minha mãe, por seu amor incondicional, apoio constante e por acreditar em mim em todos os momentos, mesmo nos mais desafiadores. Sua presença foi um pilar fundamental para que eu pudesse concluir essa etapa.

À Universidade Federal de Pernambuco (UFPE), por me proporcionar a formação acadêmica e as condições necessárias para o desenvolvimento deste trabalho.

Ao Laboratório de Genética e Biotecnologia Vegetal, pela estrutura e todo o suporte técnico cedido e, ao Laboratório Nacional de Computação Científica (LNCC), pelo acesso a recursos computacionais fundamentais para as análises realizadas.

Agradeço especialmente à minha orientadora, Professora Ana Maria Benko Iseppon, pela orientação criteriosa, pelo incentivo constante e pela confiança depositada ao longo de todo o processo.

Ao meu coorientador, Madson A. de Luna Aragão, pela valiosa contribuição científica, apoio técnico e conselhos que enriqueceram imensamente este estudo. Sua sabedoria e paciência foram fundamentais para a concretização deste trabalho.

Aos meus colegas de laboratório, Ayug, Wilson e Rafael, pelo companheirismo, pelas discussões científicas e por tornarem o ambiente de trabalho mais leve e colaborativo.

Gostaria de expressar minha gratidão aos meus amigos de curso, Larissa, Kamyille, Brenda, Marina e João Victor, por todo o apoio, incentivo e companheirismo ao longo dessa jornada acadêmica. Agradeço por tornarem essa caminhada mais leve e significativa.

Dito isso, esse trabalho é fruto da amizade e colaboração que construí ao longo dos anos, obrigado a todos que fizeram parte dessa jornada comigo.

Penna, Saulo Rafael Mendes. Caracterização estrutural in silico das isoformas do fator eIF4E em *Vigna unguiculata*. 2024. 101 páginas. Trabalho de Conclusão de Curso (Graduação em Biomedicina) – Universidade Federal de Pernambuco, Recife, 2024

RESUMO

O Brasil é um dos maiores exportadores de *commodities* do mundo, destacando-se nas leguminosas como um dos seus principais produtos. Entre essas leguminosas, o feijão-caupi (*Vigna unguiculata* (L.) Walp), é conhecido pelo elevado teor de proteínas, vitaminas e sais minerais em sua composição, além de possuir grande importância socioeconômica nas Américas, Ásia e África. Relatórios anuais, revelam perdas nas safras dessa cultura, devido à infecção por fitopatógenos. A interação entre biomoléculas da planta e dos fitopatógenos se tornaram um dos principais tópicos para compreensão das infecções por esses patógenos, bem como para busca por soluções biotecnológicas para combatê-las. Logo, estudos que visam compreender as vias utilizadas por fitopatógenos, como o *Cowpea aphidborne mosaic virus* (CABMV), são importantes para a mitigação de perdas relacionadas às infecções causadas em culturas de *V. unguiculata*. Estudos sugerem que uma das principais vias de ação dos patógenos, ocorre pelo sequestro da maquinaria do complexo ribossomal, sendo o principal alvo, o fator de iniciação da tradução eucariótico 4E (eIF4E), fundamental para manutenção da infecção pelo vírus. Dessa forma, o objetivo deste estudo foi investigar e caracterizar, em escala atomística, a estrutura de diferentes proteínas codificadas por homólogos de eIF4E, incluindo aquelas associadas ao mecanismo de tolerância e/ou resistência em cultivares de feijão-caupi. Para tal, minerações foram realizadas de acordo com os dados de anotação dos genomas disponíveis pelo Laboratório de Genética e Biotecnologia Vegetal (LGBV) da Universidade Federal de Pernambuco (UFPE) e, pelos genomas de *P. vulgaris* (ID: 442) e *V. unguiculata* (ID: 540) disponíveis no Phytozome. As modelagens para o fator eIF4E, foram realizadas por seguinte, os modelos foram validados quanto à qualidade de dobramento e estereoquímica e foram submetidos a simulações em dinâmica molecular. Os dados de dinâmica molecular revelaram diferenças entre os sistemas, com destaque para as isoformas do cromossomo 6, que apresentaram diferenciação de cargas superficiais que não favorecem interações com as proteínas dos fitopatógenos. Assim, sendo alvo de estudos promissores para a defesa das safras de *V. unguiculata*, contra esses agentes infecciosos.

Palavras-chave: Feijão-caupi; Fator de Tradução Eucariótico; Fitopatógenos; Modelagem Molecular; Dinâmica Molecular;

Penna, Saulo Rafael Mendes. **In silico Structural Characterization of eIF4E Isoforms in *Vigna unguiculata***. 2024. 101 pages. Undergraduate Thesis (Bachelor's Degree in Biomedicine) – Federal University of Pernambuco, Recife, 2024.

ABSTRACT

Brazil is one of the world's largest commodity exporters, with legumes standing out as one of its main products. Among these legumes, cowpea (*Vigna unguiculata* (L.) Walp) is renowned for its high protein, vitamin, and mineral content, as well as its significant socioeconomic importance in the Americas, Asia, and Africa. Annual reports reveal crop losses of this species due to infection by phytopathogens. The interaction between plant biomolecules and phytopathogens has become one of the main topics for understanding infections by these pathogens, as well as for seeking biotechnological solutions to combat them. Therefore, studies aiming to understand the pathways used by phytopathogens, such as the Cowpea aphidborne mosaic virus (CABMV), are important for mitigating losses related to infections in *V. unguiculata* crops. Studies suggest that one of the main pathogen action pathways occurs through the hijacking of the ribosomal complex machinery, with the main target being the eukaryotic translation initiation factor 4E (eIF4E), which becomes crucial for maintaining viral infection. Thus, the objective of this study was to investigate and characterize, at an atomistic scale, the structure of different proteins encoded by homologs of eIF4E, including those associated with tolerance and/or resistance mechanisms in cowpea cultivars. To achieve this, mining was conducted according to the genome annotation data from the Genetics Department, available from the Plant Genetics and Biotechnology Laboratory (LGBV) at the Federal University of Pernambuco (UFPE), as well as from the genomes of *P. vulgaris* (ID: 442) and *V. unguiculata* (ID: 540) available on Phytozome. The eIF4E factor models were then validated for folding quality and stereochemistry and subsequently subjected to molecular dynamics simulations. The molecular dynamics data revealed structural differences between the systems, with emphasis on chromosome 6 isoforms, which exhibited surface charge differentiation that does not favor interactions with phytopathogen proteins. Thus, these are promising targets for studies aimed at defending *V. unguiculata* crops against these infectious agents.

Key words: Cowpea; Eukaryotic Translation Factor; Phytopathogens; Molecular Modeling; Molecular Dynamics;

LISTA DE ILUSTRAÇÕES

| | |
|--|----|
| Figura 1 – Grãos de <i>V. unguiculata</i> | 17 |
| Figura 2 – Fitopatógenos..... | 19 |
| Figura 3 – Complexo Proteico Eucariótico eIF4F..... | 21 |
| Figura 4 – Estrutura do Fator de Tradução Eucariótico eIF4E..... | 24 |
| Figura 5 – Imagem Ilustrativa de métodos computacionais voltados para predição de modelos tridimensionais de proteínas a partir da sequência primária..... | 25 |
| Figura 6 – Imagem conceitual sobre simulações de sistemas proteicos por simulação em dinâmica molecular..... | 27 |
| Figura 7 – Alinhamento completo das sequências do fator eIF4E das espécies eucarióticas: <i>P. vulgaris</i> , <i>V. unguiculata</i> , <i>M. musculus</i> , <i>A. thaliana</i> , <i>H. sapiens</i> e <i>D. rerio</i> | 35 |
| Figura 8 – Alinhamento das sequências de aminoácidos completa do cromossomo 4 da proteína eIF4E de <i>Phaseolus vulgaris</i> e das cultivares BR14, DEN, IT85F, SI e PO de <i>Vigna unguiculata</i> | 36 |
| Figura 9 – Alinhamento dos domínios das sequências de aminoácidos do cromossomo 4 da proteína eIF4E de <i>Phaseolus vulgaris</i> e das cultivares BR14, DEN, IT85F, SI e PO de <i>Vigna unguiculata</i> | 37 |
| Figura 10 – Alinhamento das sequências de aminoácidos completa do cromossomo 7 da proteína eIF4E de <i>Phaseolus vulgaris</i> e das cultivares BR14, DEN, IT85F, SI e PO de <i>Vigna unguiculata</i> | 38 |
| Figura 11 – Alinhamento dos domínios das sequências de aminoácidos do cromossomo 7 da proteína eIF4E de <i>Phaseolus vulgaris</i> e das cultivares BR14, DEN, IT85F, SI e PO de <i>Vigna unguiculata</i> | 38 |
| Figura 12 – Alinhamento das sequências de aminoácidos completa do cromossomo 6 da proteína eIF4E de <i>Phaseolus vulgaris</i> e das cultivares BR14, IT85F e SI de <i>Vigna unguiculata</i> | 39 |
| Figura 13 – Alinhamento dos domínios das sequências de aminoácidos do cromossomo 6 da proteína eIF4E de <i>Phaseolus vulgaris</i> e das cultivares BR14, IT85F e SI de <i>Vigna unguiculata</i> | 40 |
| Figura 14 – Representação em cartoon do modelo estrutural da eIF4E do cromossomo 4 da cultivar BR14. | 41 |

| | |
|--|----|
| Figura 15 – Representação em cartoon dos modelos estruturais da eIF4E do cromossomo 4 das cultivares DEN, IT85F, PO e SI..... | 42 |
| Figura 16 – Representação em cartoon do modelo estrutural da eIF4E do cromossomo 7 da cultivar BR14..... | 43 |
| Figura 17 – Representação em cartoon dos modelos estruturais da eIF4E do cromossomo 7 das cultivares DEN, IT85F, PO e SI..... | 44 |
| Figura 18 – Representação em cartoon dos modelos estruturais da eIF4E do cromossomo 6 das cultivares BR14, IT85F e SI..... | 45 |
| Figura 19 – Representação em cartoon do alinhamento das modelagens para as sequências de eIF4E do cromossomo 4..... | 46 |
| Figura 20 – Representação em cartoon do alinhamento das modelagens para as sequências de eIF4E do cromossomo 7. | 47 |
| Figura 21 – Representação em cartoon do alinhamento das modelagens para as sequências de eIF4E do cromossomo 6. | 47 |
| Figura 22 – Gráfico de RMSD do cromossomo 4 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 50 |
| Figura 23 – Gráfico de RMSF do cromossomo 4 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 52 |
| Figura 24 – Gráfico das Pontes de Hidrogênio do cromossomo 4 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 53 |
| Figura 25 – Gráfico do Raio de Giro do cromossomo 4 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 54 |
| Figura 26 – Gráfico de RMSD do cromossomo 7 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 55 |
| Figura 27 – Gráfico de RMSF do cromossomo 7 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 56 |
| Figura 28 – Gráfico das Pontes de Hidrogênio do cromossomo 7 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 57 |
| Figura 29 – Gráfico do Raio de Giro do cromossomo 7 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 58 |
| Figura 30 – Gráfico de RMSD do cromossomo 6 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 59 |
| Figura 31 – Gráfico de RMSF do cromossomo 6 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 60 |

| | |
|--|----|
| Figura 32 – Gráfico das Pontes de Hidrogênio do cromossomo 6 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 61 |
| Figura 33 – Gráfico do Raio de Giro do cromossomo 6 das cultivares de <i>Vigna</i> e da <i>P. vulgaris</i> | 62 |
| Figura 34 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 4 da cultivar BR14..... | 64 |
| Figura 35 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 4 da cultivar Denkitiliana..... | 65 |
| Figura 36 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 4 da cultivar IT85F..... | 66 |
| Figura 37 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 4 da cultivar Pingo de Ouro..... | 67 |
| Figura 38 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 4 da cultivar Santo Inácio..... | 68 |
| Figura 39 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 7 da cultivar BR14..... | 69 |
| Figura 40 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 7 da cultivar Denkitiliana..... | 70 |
| Figura 41 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 7 da cultivar IT85F..... | 71 |
| Figura 42 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 7 da cultivar Pingo de Ouro..... | 72 |
| Figura 43 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 7 da cultivar Santo Inácio..... | 73 |
| Figura 44 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 6 da cultivar BR14..... | 74 |
| Figura 45 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 6 da cultivar IT85F..... | 74 |
| Figura 46 – Resíduos de destaque para a diferenciação e perfis de superfície eletrostática do cromossomo 6 da cultivar Santo Inácio..... | 75 |

LISTA DE TABELAS

Tabela 1 – Descrição de scores de validação do QMEAN-DisCo, Z-Score e PROCHECK das estruturas modeladas das cultivares de *Vigna unguiculata*.....49

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|---|
| ABRAFE | Instituto Brasileiro de Feijão e Pulses |
| PABP | Proteína de Ligação a Poli-A |
| CASP | Do inglês, <i>Critical Assessment of Protein Structure Prediction</i> |
| DEN | Cultivar Denkitiliana. |
| PO | Cultivar Pingo de Ouro. |
| SI | Cultivar Santo Inácio. |
| MD | Modelagem Molecular. |
| RMSD | Do inglês, <i>Root-Mean-Square Deviation</i> , Raiz Quadrada do Desvio Quadrático Médio. |
| RMSF | Do inglês, <i>Root-Mean-Square Fluctuation</i> , Raiz Quadrada da Flutuação Quadrática Média. |
| HB | Do inglês, <i>Hydrogen Bond</i> , Pontes de Hidrogênio. |
| RG | Do inglês, <i>Radius of Gyration</i> , Raio de Giro. |
| R18 | Resíduo de Arginina na posição 18 do alinhamento da sequência de domínio. |
| K18 | Resíduo de Lisina na posição 18 do alinhamento da sequência de domínio. |
| P55 | Resíduo de Prolina na posição 55 do alinhamento da sequência de domínio. |
| W69 | Resíduo de Triptofano na posição 69 do alinhamento da sequência de domínio. |
| F120 | Resíduo de Fenilalanina na posição 120 do alinhamento da sequência de domínio. |

SUMÁRIO

| | | |
|----------|---|----|
| 1 | REVISÃO BIBLIOGRÁFICA | |
| 1.1 | <i>VIGNA UNGUICULATA</i> | 17 |
| 1.2 | Fitopatógenos..... | 18 |
| 1.3 | Complexo eucarioto eIF4F..... | 21 |
| 1.4 | <i>Fator de tradução eucarioto eIF4E</i> | 23 |
| 1.5 | Modelagem molecular..... | 24 |
| 1.6 | Simulação em dinâmica molecular..... | 27 |
| 2 | OBJETIVOS | |
| 2.1 | Objetivo Geral..... | 31 |
| 2.2 | Objetivos Específicos..... | 31 |
| 3 | METODOLOGIA | |
| 3.1 | MINERAÇÃO DOS GENES..... | 32 |
| 3.2 | Alinhamento de Sequências e Definição de Domínio Estrutural..... | 32 |
| 3.3 | Modelagem molecular..... | 32 |
| 3.4 | <i>Simulação em dinâmica molecular</i> | 33 |
| 4 | RESULTADOS E DISCUSSÃO | |
| 4.1 | Alinhamentos e Definição de Domínios Estruturais..... | 35 |
| 4.3 | Modelagem Molecular..... | 41 |
| 4.2 | Alinhamento das Estruturas Comparadas e Validação dos Modelos Teóricos..... | 45 |
| 4.4 | Simulação em Dinâmica molecular..... | 49 |
| 4.5 | Perfil de Superfície Eletrostática..... | 63 |
| 5 | CONCLUSÃO | 76 |
| 6 | REFERÊNCIAS | 77 |
| 7 | APÊNDICES | 89 |

INTRODUÇÃO

O feijão-caupi, popularmente conhecido como feijão-frade, feijão-macáçar e feijão-de-corda - espécie botânica *Vigna unguiculata* (L.) Walp. - é um alimento de elevado valor nutricional (FROTA; SOARES, R. A. M.; ARÊAS, 2008). Constitui um importante parcela da fonte proteica da alimentação de diferentes populações do mundo (FREITAS; TEIXEIRA; FERREIRA, R. B., 2004; GONÇALVES *et al.*, 2016; THANGADURAI, 2005). O Brasil se destaca como um dos maiores produtores de feijão-caupi, no entanto, fatores bióticos e abióticos comprometem as lavouras, ocasionando perdas anualmente (LEITE, Rubson Da Costa; LEITE, Robson Da Costa; CRAVO, 2020). Tais perdas estão diretamente relacionadas a infecções por fitopatógenos, que se figuram entre os principais responsáveis pelos prejuízos em lavouras de diversas culturas de leguminosas todos os anos (BASHIR; AHMAD; GHAFOR, 2002; GONÇALVES *et al.*, 2016).

Durante o processo de infecção, alguns fitopatógenos aproveitam o sistema de tradução do hospedeiro vegetal, especificamente fatores de iniciação da tradução eucariótica (eIFs, do inglês *Eukaryotic Translation Initiation Factors*), para direcionar a síntese de proteínas virais, comprometendo as funções da planta. Entre esses fatores, o eIF4E e suas isoformas se destacam, pois são essenciais não apenas para a tradução de proteínas no hospedeiro, mas também para a infecção de vírus de determinadas famílias, como a *Potyviridae* (WALTER *et al.*, 2020). Por essa razão, as isoformas do eIF4E são alvos estratégicos tanto para programas de melhoramento genético, devido ao seu papel na síntese proteica, quanto para o desenvolvimento de cultivares resistentes a fitopatógenos que dependem desses fatores (SINDHU *et al.*, 2019).

Visto que, o fator eIF4E figura como um dos principais componentes do complexo multiprotéico que contribui para o início da tradução, pois, é a partir do seu pareamento no complexo traducional, que ocorre o reconhecimento e a interação entre a região 5-cap dos ácidos ribonucleicos mensageiros (mRNA) e o recrutamento de subunidades dos ribossomos e outros fatores que compõem o complexo da qual a eIF4E faz parte (ALA-POIKELA; RAJAMÄKI; VALKONEN, 2019; COUTINHO DE OLIVEIRA *et al.*, 2019; XU *et al.*, 2017). Esse sistema multimérico é denominado de complexo proteico eIF4F, sendo composto por 4 diferentes proteínas que interagem

entre si e compõem o sítio de ligação entre o complexo, o mRNA e as subunidades do ribossomo (MAYBERRY *et al.*, 2009; SAHA; MÄKINEN, 2020).

Diante do que foi exposto, estudos que caracterizem o eIF4E e suas isoformas no feijão-caupi são de importância para o melhoramento genético, visando melhorias na produção desse grão. Estas pesquisas podem levar à identificação de variações naturais ou modificações que conferem resistência a fitoviroses dependentes desse fator, contribuindo para o desenvolvimento de cultivares mais robustas e resistentes a doenças virais, além de aprofundar o entendimento dos mecanismos moleculares da síntese proteica e infecção viral em plantas.

1. REVISÃO BIBLIOGRÁFICA

1.1 *Vigna unguiculata*

A *Vigna unguiculata* (L.) Walp., conhecida popularmente no Brasil como feijão-caupi (Figura 1), é uma leguminosa cultivada em diversas partes do mundo, com maior distribuição na África, América Latina e em algumas regiões da Ásia (FREITAS; TEIXEIRA; FERREIRA, R. B., 2004; GONÇALVES *et al.*, 2016; THANGADURAI, 2005).

O feijão-caupi apresenta diferentes variações dentro da mesma espécie, resultantes de processos de seleção ou melhoramento genético. Essas variações são denominadas como cultivares, e podem influenciar características como resistência a doenças, produtividade e adaptação a diferentes condições ambientais (BOUKAR *et al.*, 2019; GOMES *et al.*, 2020; SILVA, J. A. *et al.*, 2021). As cultivares da *V. unguiculata* são conhecidas pelas variações genéticas e alterações fenotípicas de coloração dos grãos, folhagem e tamanho do grão (BOUKAR *et al.*, 2019; GOMES *et al.*, 2020).

Figura 1. Grãos de *V. unguiculata*.



Estima-se que a extensão de plantio do feijão-caupi globalmente, tenha uma área de cultivo de cerca de 14 a 15 milhões de hectares, resultando anualmente em mais de 8 milhões de toneladas de grãos (DHANASEKAR; SOUFRAMANIEN; SUPRASANNA, 2021). Informações do Instituto Brasileiro de Feijão e Pulses (ABRAFE), indicaram uma projeção de 630 mil toneladas de feijão-caupi produzidos no Brasil, entre os anos de 2022 e 2023, abrangendo uma média de produtividade de 1410kg por hectare no Distrito Federal (CHEMIM, 2022). Na América Latina, o Brasil continua em crescimento constante em termos de produção e área de cultivo, e, atualmente, ocupa uma das maiores posições de produção desse legume, mundialmente (MINISTÉRIO DA AGRICULTURA (MAPA), 2024).

O feijão-caupi se destaca como uma fonte significativa de macro e micronutrientes, sendo uma das principais fontes de proteína para populações das Américas e da África (JAYATHILAKE *et al.*, 2018). Nutricionalmente, o grão apresenta baixo teor de gordura e elevado valor proteico, com cerca de 23-32% de proteína, 50-60% de carboidratos e, em média, 1% de gordura (CARVALHO *et al.*, 2017; JAYATHILAKE *et al.*, 2018). Apesar de ser um alimento rico em nutrientes, possui baixa densidade calórica quando comparado a outros grãos (JAYATHILAKE *et al.*, 2018; Safety Assessment of Foods and Feeds Derived from Transgenic Crops, Volume 3, 2019). Assim, em diversas regiões do mundo, a produção e o consumo do feijão-caupi contribuem de maneira substancial para a sociedade, especialmente nos aspectos econômicos e nutricionais. No entanto, alguns tipos de estresses bióticos e abióticos, representam ameaças significativas à sua produção, de exemplo, a desidratação radicular e a infecção por alguns tipos de vírus, como, o Cowpea aphid-born mosaic vírus (CABMV) da família *Potyviridae* (BORGES, 2021).

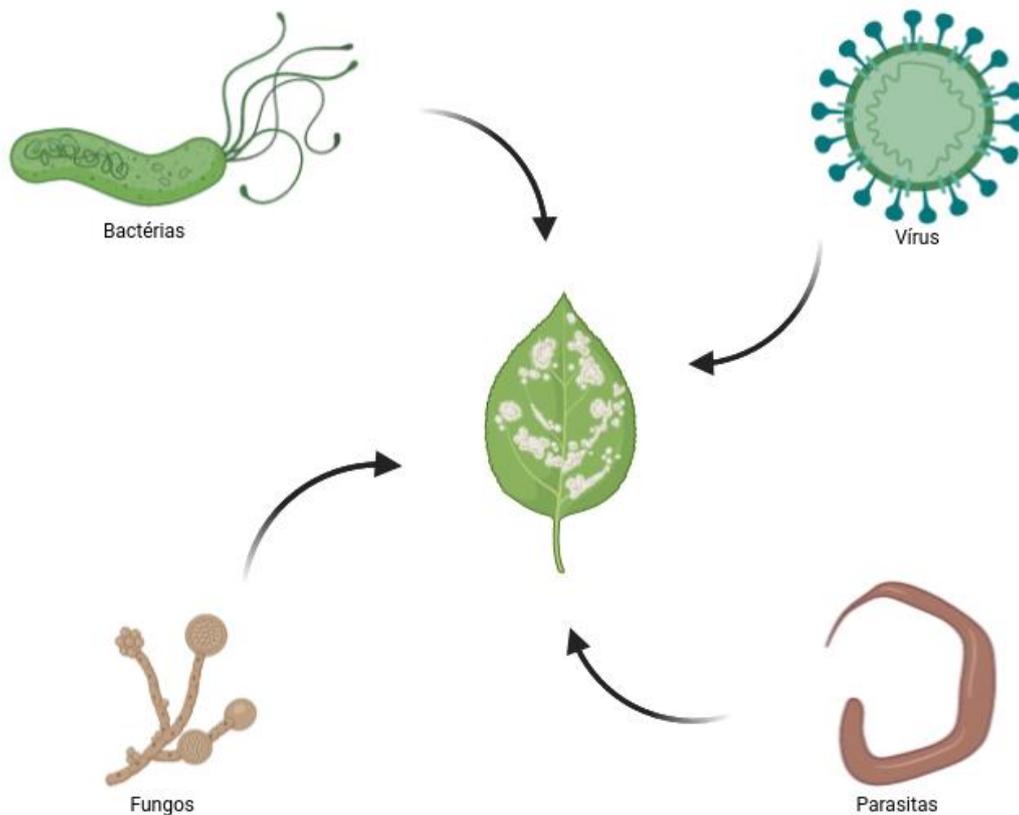
1.2 Fitopatógenos

Na agricultura, os fitopatógenos impactam diretamente o crescimento e a sobrevivência das plantações, além de diminuir a produção e a qualidade da colheita, o que geralmente resulta em perdas financeiras. As plantas podem ser infectadas por parasitas, fungos, bactérias e vírus (Figura 2) (ANDERSEN *et al.*, 2018; NAZAROV *et al.*, 2020).

Esses patógenos podem ser disseminados de forma horizontal, ou seja, por

meio de vetores como solo, água de irrigação e insetos, e, também, de maneira vertical, por transmissão parental (JAIN, A. *et al.*, 2019). Além disso, esses fitopatógenos podem ser espécie-específicos, afetando apenas determinadas espécies de plantas, ou generalistas, capazes de infectar uma variedade de fito-hospedeiros (ANDERSEN *et al.*, 2018).

Figura 2. Fitopatógenos. Esquema representativo de patógenos capazes de causar doenças em plantas.



Fonte: Criado pelo autor em BioRender.com.

Dentre os microrganismos patogênicos, os parasitas *Meloidogyne spp.*, *Pratylenchus spp.*, *Ditylenchus dipsaci* e *Aphelenchoides spp.* são reconhecidos por afetar diretamente a variedade de plantas agrícolas, por meio de danos constantes nas raízes e folhagens (JONES, J. T. *et al.*, 2013; POVEDA; ABRIL-URIAS; ESCOBAR, 2020; SIDDIQUE; GRUNDLER, 2018).

Outro patógeno de interesse ao controle de pragas, são os fungos com mais de 19 mil espécies reconhecidas como causadoras de doenças em vegetais (HARIHARAN; PRASANNATH, 2021). Fungos patogênicos conseguem infectar plantas em diferentes fases do desenvolvimento vegetal, desde o estágio de muda até a maturação final, afetando até a germinação das sementes (HARIHARAN; PRASANNATH, 2021). Como representantes dessa classe, pode-se citar *Botrytis cinérea* que causa necrose no tecido vegetal, *Puccinia spp.* que compromete folhagens causando uma espécie de ferrugem nas folhas e *Colletotrichum spp.* que causa antracnose (DEAN *et al.*, 2012). As infecções fúngicas podem ser moduladas por outras infecções concomitantes nas plantas ou por outros estresses de ordem física, como: seca, temperatura e alterações osmóticas (NAZAROV *et al.*, 2020). Assim como as infecções parasitárias supracitadas, essas infecções causam perdas significantes, sejam em termos de quantidade ou qualidade (HARIHARAN; PRASANNATH, 2021; JAIN, A. *et al.*, 2019).

As bactérias, também são reconhecidas como agentes nocivos aos vegetais, causando graves quadros patogênicos, que vão desde manchas nas folhas como padrões de mosaico ou pústulas, até o apodrecimento de tubérculos, responsável pela morte da planta (NAZAROV *et al.*, 2020; PFEILMEIER; CALY; MALONE, 2016; VIDAVER; LAMBRECHT, 2004). Nesses casos, a maioria das bactérias patogênicas são Gram-positivas classificadas no filo *Actinobacteria*, ou Gram-negativas no filo *Proteobacterias* (VIDAVER; LAMBRECHT, 2004). Apesar de se infectarem por uma gama de bactérias patogênicas, as plantas possuem um complexo sistema imunológico composto por proteínas antimicrobianas e outras substâncias imunomoduladoras (NAZAROV *et al.*, 2020; PFEILMEIER; CALY; MALONE, 2016).

Nas plantas, assim como os outros patógenos, os vírus representam uma importante fonte causadora de doenças infecciosas. Alguns relatos mostram que mais da metade das infecções por fitopatógenos são causadas por vírus (MEHETRE *et al.*, 2021). É reportado que mais de 25 famílias de vírus são capazes de infectar uma variedade de espécies agrícolas, causando danos e consequentemente ocasionando perdas (MEHETRE *et al.*, 2021). Algumas famílias de vírus possuem um destaque devido aos danos que podem causar às plantações, uma delas é a família *Potyviridae*, cujo *Potato virus Y (PVY)*, *Tomato mosaic virus (ToMV)* e *Cowpea aphid-borne mosaic virus (CABMV)* pertencem (COLARICCIO *et al.*, 2020; LEFEUVRE *et al.*, 2019).

O eIF4E se liga à região CAP 5' do mRNA, e interage com o eIF4G, que juntos, funcionam como suporte para outros fatores de iniciação (MAYBERRY *et al.*, 2009). O eIF4G é responsável por se ligar com a Proteína de Ligação a Poli-A (PABP), que interage com a porção poli-A 3' do mRNA, formando por fim um sistema onde o mRNA está ligado pelas duas extremidades ao complexo eIF4F (MAYBERRY *et al.*, 2009; PATRICK; BROWNING, 2012).

Dentro do complexo ainda existe o fator eIF4A que atua como helicase desfazendo estruturas de tensão e removendo proteínas anexadas ao mRNA (Figura 5) (MONTERO; PÉREZ-GIL; SAMPIERI, 2019; OBERER; MARINTCHEV; WAGNER, 2005). Essa função de helicase do mRNA, exercida pela eIF4A, é bastante necessária para facilitar a varredura da sequência no ribossomo, principalmente, nos RNAs com uma elevada região 5' não codificante, além disso, é denominada uma estrutura *DEAD-box*, por possuir um motivo conservado de aminoácidos de Asp-Glu-Ala-Asp (*DEAD*) (SANFAÇON, 2015).

Quando as unidades do complexo estão acopladas ao ribossomo e se liga ao mRNA, dá-se início a uma varredura de busca pelo códon de início (AUG), e pôr fim a tradução acontece, até que códons de parada sejam identificados (UAA, UAG e UGA) (BASTET; ROBAGLIA; GALLOIS, 2017; PATRICK; BROWNING, 2012; SAHA; MÄKINEN, 2020).

Além disso, para o início da tradução, faz-se necessário a atividade do complexo 43S, o qual é composto pela subunidade ribossomal 40S, o fator eIF3 e um grupamento ternário, constituído pelo fator eIF2, GTP e metionina-iniciador-tRNA (Met-tRNA_i) (MONTERO; GARCÍA-ROMÁN; MORA, 2015). A função desse agrupamento é ser carregado, junto com a subunidade ribossomal 40S, para o códon iniciador, onde as 2 subunidades irão se ligar para assim formar o ribossomo completo (MONTERO; GARCÍA-ROMÁN; MORA, 2015).

Por fim, as conformações dimensionais da proteína eIF4E, se ligam a extremidade 5' do mRNA formando uma estrutura semelhante a mãos em concha, na qual, a região 5' CAP se encontra localizada entre resíduos conservados de triptofano (Trp), que são capazes de reconhecer a porção 7-methyl-GTP do mRNA (JOSHI *et al.*, 2005).

Apesar da conservação evolutiva, algumas plantas, como a *V. unguiculata*, apresentam isoformas para algumas unidades do complexo eIF4F, sendo estas

eIF(iso)4E e eIF(iso)4G (MAYBERRY *et al.*, 2009; SAHA; MÄKINEN, 2020). (Alguns estudos relatam a formação de complexos mistos de eIF4E a eIF(iso)4G e de eIF(iso)4E a eIF4G, no entanto, com performance enzimática reduzida (COUTINHO DE OLIVEIRA *et al.*, 2019; MAYBERRY *et al.*, 2009; PATRICK; BROWNING, 2012). Fato esse existente, em virtude da capacidade das plantas em se tornar mais susceptível ou resistentes a vírus e outros possíveis patógenos, por meio de mutações pontuais nesses fatores de tradução (ZLOBIN; TARANOV, 2023a).

No entanto, esse elevado número de isoformas existentes não é uma peculiaridade exclusiva das plantas, como exemplo a *Drosophila melanogaster*, que possui 8 diferentes isoformas do fator de tradução eIF4E no seu genoma (LAYANA *et al.*, 2023). Por outro lado, nos mamíferos ocorre o oposto, tal como no humano, onde há uma quantidade reduzida de isoformas do fator eIF4E, quando comparado a outras espécies, na qual há apenas a presença de 2 isoformas localizadas no cromossomo 3, a eIF4E2 e a eIF4E3 (FRYDRYSKOVA *et al.*, 2016). Por este motivo, as plantas da espécie *V. unguiculata*, devido ao seu valor socioeconômico, são ótimos candidatos para esse estudo de identificação e caracterização das isoformas do fator eIF4E.

1.4 Fator de tradução eucariótico eIF4E

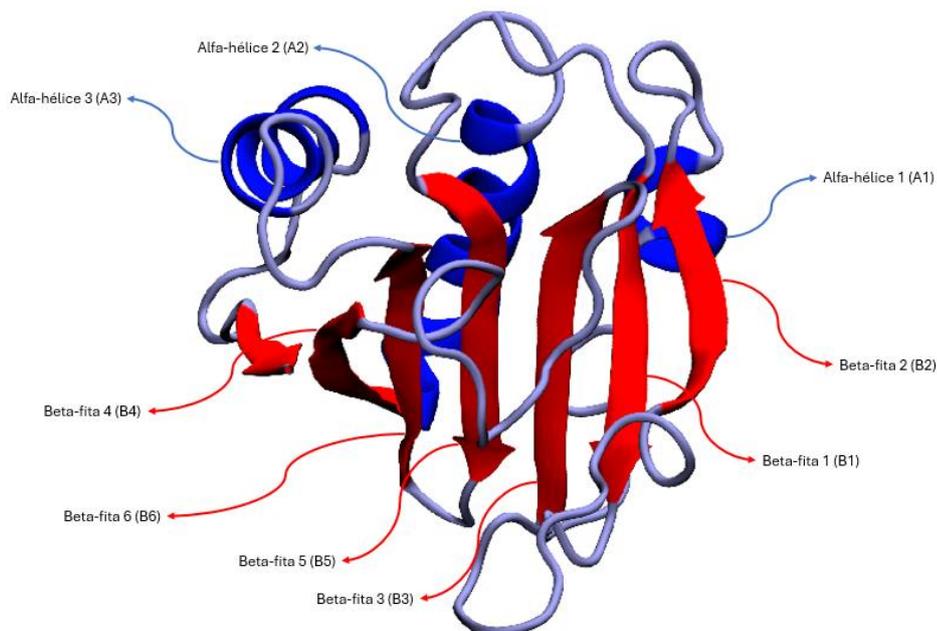
O fator de tradução eucariótico eIF4E apresenta alta conservação estrutural, exibindo um elevado grau de similaridade entre suas sequências. Resíduos críticos para sua função são conservados, como os triptofanos, mesmo entre diferentes espécies. Esses resíduos conservados, são extremamente importantes para o sucesso da ligação, pois, além de serem capazes de reconhecer certas regiões de ancoragem com o mRNA, também auxiliam no equilíbrio de cargas superficiais no sítio de ligação com o RNA mensageiro, corroborando para que a ligação ocorra (HERNÁNDEZ-ALVAREZ *et al.*, 2021)

Assim sendo, os Triptofanos (Trp) são um dos principais resíduos conservados de importância para tal ligação, mesmo sendo aminoácidos aromáticos que possuem carga superficial parcialmente negativa (JENKINS; LARSEN; WILLIAMS, 2005), em conjunto com as Lisinas (Lys) e as Argininas (Arg), que possuem carga superficial positiva (JENKINS; LARSEN; WILLIAMS, 2005), compõem o sítio de ligação com um equilíbrio eletrostático positivo, favorecendo a ligação com o 5' CAP (STACHELSKA

et al., 2002).

A conformação da eIF4E, é constituída por pares de beta-fitas antiparalelas entre si e alfa-hélices posicionadas posteriormente a essas beta-fitas (MIRAS *et al.*, 2017; ROSS-KASCHITZA; ALTMANN, 2020). No caso da *V. unguiculata* (Figura 4), apresenta 6 beta-fitas antiparalelas entre si, que são configuradas na seguinte ordem, da esquerda para a direita: beta-fita 4 (B4), beta-fita 6 (B6), Beta-fita 5 (B5), Beta-fita 3 (B3), beta-fita 1 (B1) e beta-fita 2 (B2), e 3 alfa-hélices, as quais são posicionadas posteriores às beta-fitas. Dessa forma, a primeira alfa-hélice encontra-se posicionada numa região localizada posteriormente às beta-fitas 1 e 2, a segunda alfa-hélice atrás das fitas 3 e 5, e a alfa-hélice 3, atrás das fitas 6 e 4.

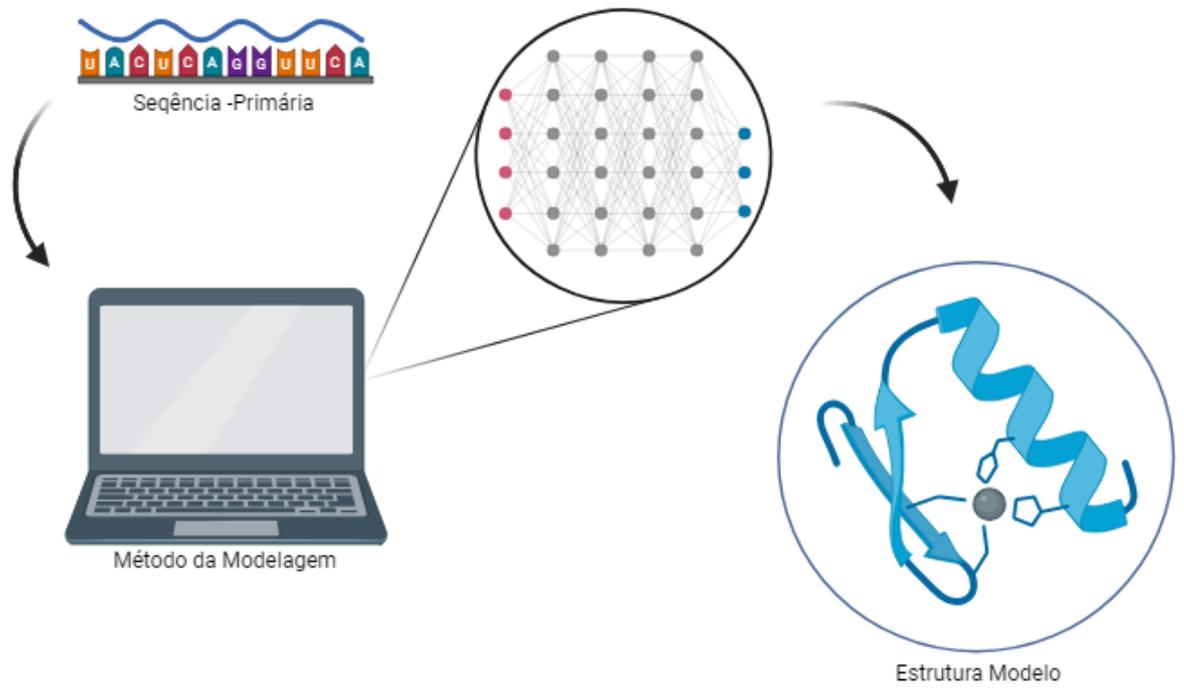
Figura 4. Modelo teórico estrutural do fator eucariótico eIF4E de *Vigna unguiculata*. Representado em cartoon, em azul as Alfa-hélices e em vermelho as Beta-fitas.



1.5 Modelagem molecular

A elucidação de modelos tridimensionais (3D) de proteínas (Figura 5) e seus complexos fornecem informações valiosas, não apenas para a compreensão das bases moleculares de fenômenos biológicos, mas também para triagem e desenvolvimento de novos fármacos (SANTOS-SILVA *et al.*, 2020).

Figura 5. Imagem ilustrativa de métodos computacionais voltados para predição de modelos tridimensionais de proteínas a partir da sequência primária.



Fonte: Criado pelo autor em BioRender.com.

Atualmente é possível classificar a modelagem computacional de proteínas em 4 categorias distintas: modelagem comparativa, segmentação (do inglês *threading*), *ab initio* e modelagem baseada em aprendizado de máquina. Os dois primeiros métodos, modelagem comparativa e segmentação, constroem modelos 3D baseados em alinhamentos de sequências de estruturas tridimensionais já resolvidas e disponíveis para busca (SANTOS-SILVA *et al.*, 2020). Caso os modelos correspondentes às sequências buscadas não estejam disponíveis no *Protein Data Bank* (PDB), os modelos devem ser preditos por modelagem *ab initio*, ou seja, sem referência prévia, baseados apenas em funções energéticas fundamentadas em leis da termodinâmica (SANTOS-SILVA *et al.*, 2020).

Por fim, mas não menos importante, vêm os métodos baseados em aprendizado de máquina, que recentemente vêm chamando atenção na Biologia Estrutural (KOMAROFF, 2021). Este último incorpora novas arquiteturas de redes neurais e procedimentos de treinamento de base com utilizações de informações de restrições evolutivas, físicas e geométricas das estruturas proteicas. Segundo o CASP (do inglês, *Critical Assessment of Protein Structure Prediction*), os algoritmos baseados em aprendizado de máquina conseguem prever e selecionar modelos 3D

com alto valor assertivo quando comparado com os outros métodos (JUMPER *et al.*, 2021; KRYSHTAFOVYCH *et al.*, 2021; KUHLMAN; BRADLEY, 2019).

Alguns estudos já apontam essa técnica como um dos maiores avanços da ciência dos últimos anos, um exemplo é o AlphaFold que foi usado para realizar a predição de parte do proteoma humano (disponível em: <https://alphafold.ebi.ac.uk>) (CALLAWAY, 2020; HUMPHREYS *et al.*, 2021; KOMAROFF, 2021; THORP, 2021; TUNYASUVUNAKOOL *et al.*, 2021).

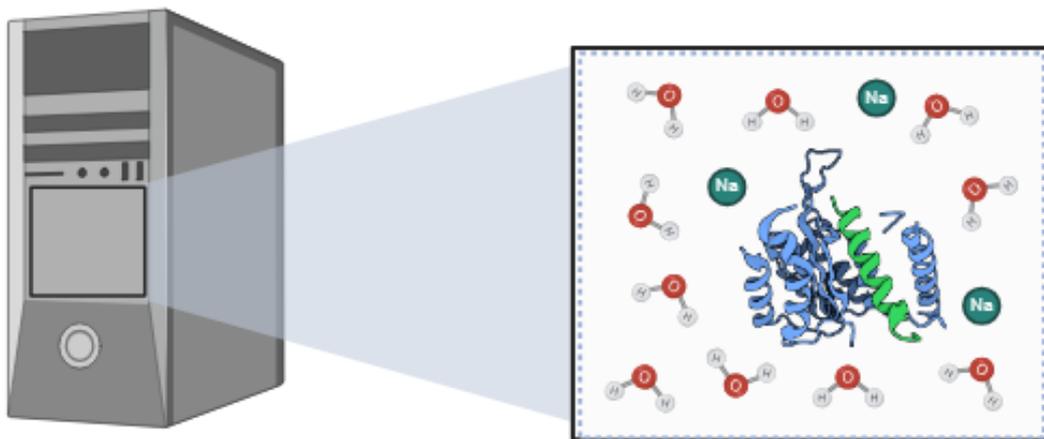
Dessa forma, a modelagem molecular é capaz de analisar e representar as propriedades estáticas da molécula, suas propriedades físico-químicas, como por exemplo, a distribuição das cargas eletrostáticas superficiais (SANT'ANNA, 2009). Devido ao seu fundamental papel na química e na biologia, essa propriedade é de suma importância em variados processos biológicos, como, na identificação das regiões moleculares e nas catalises enzimáticas. Desse modo, métodos computacionais para a avaliação dos potenciais eletrostáticos, tornou-se popular globalmente, principalmente nas áreas da biofísica molecular, biologia estrutural e na química medicinal (YU, B. *et al.*, 2021).

Além disso, para acompanhar a crescente evolução das modelagens moleculares, alguns programas de visualização de estruturas tridimensionais foram criados, entre eles, o Visual Molecular Dynamics (VMD) e o PyMOL ganharam bastante popularidade nos últimos anos. O VMD é um programa de gráficos moleculares, que foi particularmente criado para analisar biopolímeros, como as proteínas e os ácidos nucleicos. Esse programa possui uma variedade de estilos de apresentação e visualização da molécula, assim como, métodos de coloração e a habilidade de animar as trajetórias das simulações das dinâmicas moleculares (HUMPHREY; DALKE; SCHULTEN, 1996). Assim como o VMD, o PyMOL também é um programa de visualização e representação de gráficos moleculares. Entretanto, a nova versão, o PyMOL 3, permite ao usuário uma maior facilidade de utilização da sua interface e, se os mesmos preferirem, podem até realizar todas as etapas do processo de modelagem por homologia, além de possuir uma melhor visualização para aminoácidos específicos, seus recursos de animação das estruturas tornaram-se mais variados (JANSON; PAIARDINI, 2021).

1.5 Simulação em dinâmica molecular

A simulação em dinâmica molecular (MD, do inglês *Molecular Dynamics*) (Figura 6) é uma técnica que permite prever em escala atômica como uma proteína ou outro modelo biológico se moverá ao longo do tempo com base em um modelo regido por leis físicas que conduzem as interações do sistema (LAZIM; SUH; CHOI, 2020).

Figura 6. Imagem conceitual sobre simulações de sistemas proteicos por simulação em dinâmica molecular.



Fonte: Criado pelo autor em BioRender.com.

A MD é capaz de capturar uma gama de variedades em processos biomoleculares, incluindo mudanças conformacionais, e interações interatômicas ou intermoleculares, mostrando as posições atômicas em resolução temporal. É importante ressaltar que essa técnica possui um poder preditivo capaz de evidenciar como proteínas responderão a perturbações como mutações, fosforilação, protonação ou adição/remoção de um ligante (LAZIM; SUH; CHOI, 2020; PAN *et al.*, 2019).

Simulações de MD são frequentemente usadas em combinação com uma ampla variedade de técnicas experimentais de biologia estrutural, incluindo

cristalografia de raio-X, microscopia crioeletrônica (crio-EM) e ressonância magnética nuclear (NMR) (HOLLINGSWORTH; DROR, 2018; PAN *et al.*, 2019). Além disso, a MD também é usada para validação e descrição de fenômenos estruturais em modelos teóricos construídos por técnicas de modelagem molecular (HOLLINGSWORTH; DROR, 2018; PAN *et al.*, 2019).

Simplificando, a ideia básica por trás de uma simulação em MD é direta. Dadas as posições de todos os átomos em um sistema biomolecular (exemplo: uma proteína cercada por água e íons), é possível calcular a força exercida em cada átomo por todos os outros átomos. Pode-se, portanto, usar as leis da mecânica clássica de Newton para prever a posição espacial de cada átomo em função do tempo, o que configura um campo de força (HOLLINGSWORTH; DROR, 2018; LAZIM; SUH; CHOI, 2020; PAN *et al.*, 2019). Por fim, se tem uma trajetória resultante, que é em sua essência, um filme tridimensional que descreve a configuração em nível atômico do sistema em cada ponto durante o intervalo de tempo simulado (HOLLINGSWORTH; DROR, 2018; LAZIM; SUH; CHOI, 2020; PAN *et al.*, 2019).

Nesse aspecto, os dados da dinâmica molecular são muito vantajosos, pois, além de evitar tarefas repetitivas ao usuário, produz resultados rapidamente desejados, através de uma visualização automatizada, porém, flexível do sistema (MARGREITTER; OOSTENBRINK, 2017).

Os resultados dessas trajetórias, podem ser interpretados a partir de 4 principais informações, a Raiz Quadrada do Desvio Quadrático Médio (RMSD, do inglês *Root-Mean-Square Deviation*), que é uma das análises mais comumente usadas no campo de simulações biofísicas, e é uma medida de divergência posicional de um ou mais átomos ao longo de um certo período predeterminado. Essa análise, requer vetores alternados de índices de tempo e valores de RMSD para os seus resultados (MARGREITTER; OOSTENBRINK, 2017). A Raiz Quadrada da Flutuação Quadrática Média (RMSF, do inglês *Root-Mean-Square Fluctuation*), a qual indica o grau da variação posicional de cada átomo da estrutura, sua flexibilidade, ao longo do tempo. Para a obtenção dos resultados, essa análise necessita de uma coluna com os valores de RMSF e outra coluna com todos os resíduos, ou átomos, da molécula a ser analisada (MARGREITTER; OOSTENBRINK, 2017). Outra importante análise da dinâmica molecular, são as Pontes de Hidrogênio (HBond, do inglês *Hydrogen Bond*) intrinsecamente relevantes no contexto das biomoléculas. Nela, observa-se a

formação, destruição ou manutenção das ligações de hidrogênio ao longo de um determinado período (MARGREITTER; OOSTENBRINK, 2017).

Por fim, tem-se a análise do Raio de Giro (RG, do inglês *Radius of Gyration*), a qual reflete o grau de alinhamento de uma molécula, em virtude da taxa de cisalhamento da mesma. Quando essa taxa é alta, pode-se induzir a uma modificação estrutural na molécula, que é refletido no aumento ou na diminuição do RG e, conseqüentemente, corresponder a uma diminuição da viscosidade durante o processo. Esse comportamento, pode ser atribuído a deformação ou alongamento da estrutura, influenciados pelo cisalhamento das moléculas, que foram induzidas pelo fluxo do campo num espaço de tempo predeterminado (LIU, P. *et al.*, 2017).

As trajetórias de MD, trazem informações importantes do sistema como: convergência energética, flexibilidade do sistema, estabilidade, dinâmica estrutural, interações Interatômicas ou intermoleculares etc. (HOLLINGSWORTH; DROR, 2018; LAZIM; SUH; CHOI, 2020). Essas informações são de importância crucial para o entendimento de alguns fenômenos biológicos.

Por isso, alguns pacotes de programas são utilizados para facilitar tais resultados. Um dos pacotes frequentemente usados, é o GROMACS, que é, capaz de criar um sistema fechado e centralizado, introduzir moléculas de água, ajustar a temperatura, a densidade, a pressão e produzir dados de dispersão térmica nas moléculas de H₂O, tanto para os átomos de hidrogênio, quanto para os de oxigênio (HYUNA, 2023). E, envolve transformações de interações não ligantes, para calcular solvatação ou energias livres de ligação, por exemplo (OANCA; ENT, VAN DER; ÅQVIST, 2023).

Esses softwares, podem utilizar variados campos de força em suas análises, em concordância com a necessidade do estudo. A qualidade das DM, são intrinsicamente dependentes do campo de força a ser utilizado, para apresentar as interações entre as partículas no sistema. O GROMOS 53A6, é um dos campos de força utilizados no GROMACS, que possui bons parâmetros em termos de estabilidade da proteína e do DNA, comparado com os campos de força anteriores, e possui vantagens na descrição do equilíbrio entre dobramento e desdobramento dos aminoácidos.(OOSTENBRINK *et al.*, 2005).

Com o avanço da computação e dos conhecimentos em biologia estrutural simulações em MD de sistemas de larga escala como membranas bacterianas,

capsídeos virais e grandes complexos proteicos estão ficando cada vez mais comuns em estudos biocomputacionais (TARASOVA; NERUKH, 2018; WADHWA *et al.*, 2021). Esses adventos tornarão ao longo dos anos as MDs ainda mais precisas, baratas e acessíveis.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Minerar e investigar sequências de eIF4E em genomas de *V. unguiculata*, usando estratégias *in silico*, em escala de sequência primária e atômica, visando compreender assinaturas moleculares associados aos fatores de iniciação do grupo eIF4E em *V. unguiculata*.

2.2 OBJETIVOS ESPECÍFICOS

1. Minerar modelos anotados para fatores de tradução eucarióticos do tipo 4E em genomas de cultivares de *V. unguiculata*;
2. Alinhar e caracterizar as sequências dos genes e proteínas das isoformas de eIF4E encontradas nos genomas de cultivares de *V. unguiculata*;
3. Performar modelagem tridimensional dos candidatos minerados;
4. Realizar simulação em dinâmica molecular dos modelos teóricos e extrair informações de dinâmica estrutural, estabilidade e características físico-químicas;
5. Sugerir estratégias de melhoramento genético vegetal baseado nestes fatores.

3. METODOLOGIA

3.1 MINERAÇÃO

A busca por genes eIF4E foi realizada em genomas anotados de *Phaseolus vulgaris* (ID: 442) e *Vigna unguiculata* (ID: 540), disponíveis no Phytozome. Após a identificação dos genes, as sequências primárias foram extraídas de seus respectivos proteomas putativos. As minerações envolvendo os cultivares BR14, Denkitiliana, IT85F, Santo Inácio e Pingo de Ouro de *Vigna unguiculata* foram conduzidas em genomas em fase de publicação pelo Laboratório de Genética e Biotecnologia Vegetal (LGBV).

3.2 Alinhamento

Visando compreender a conservação do fator eIF4E em organismos eucariotos e entre as cultivares de *V. unguiculata*, uma vez que os genes foram devidamente minerados, os mesmos foram submetidos a etapa de alinhamento através do servidor Clustal Omega (SIEVERS; HIGGINS, 2014) e visualizados pelo Jalview (WATERHOUSE *et al.*, 2009). Para o fator eIF4E de eucariotos, foram usadas as seguintes sequências: *P. vulgaris* (ID: 442), *V. unguiculata* (ID: 540), *Mus musculus* (UniProt ID: P63073), *Arabidopsis thaliana* (UniProt ID: O23252), *H. sapiens* (UniProt ID: P06730) e *Danio rerio* (*Zebrafish*) (UniProt ID: Q9DFS6).

Em seguida, a busca por domínios conservados foi realizada utilizando a ferramenta CD Search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) com a finalidade de buscar os domínios estruturais funcionais das proteínas alvos do estudo (LU, S. *et al.*, 2020).

E para o alinhamento das espécies de *V. unguiculata*, especificamente, foram utilizadas as seguintes cultivares para os homólogos no genoma: BR14, Denkitiliana, IT85F, Santo Inácio, e Pingo de Ouro, comparando com a espécie *P. vulgaris*.

3.3 Modelagem molecular

Para realizar as modelagens tridimensionais, das isoformas do fator eIF4E, das

cultivares de *V. unguiculata*, foi utilizado o pacote AlphaFold2, As informações de entrada utilizadas para as predições dos modelos, foram de acordo com as recomendadas pelos desenvolvedores do AlphaFold2 for Google Colab (AlphaFold-Colab).

Para avaliar a qualidade desses modelos preditos, foi analisada a profundidade e a diversidade do conjunto de alinhamentos de sequências múltiplas (MSAs) e recursos pareados, além das métricas de confiança do próprio AlphaFold2 (pLDDT e PAE) (JUMPER *et al.*, 2021; MIRDITA *et al.*, 2022). De forma complementar para concluir as validações, foram utilizadas as seguintes ferramentas: ProSa-web e os seus valores do Z-score, o PROCHECK e seus gráficos de Ramachandran e, o Swiss-model com seus valores do Qmean-DisCo, com a finalidade de afirmar a qualidade dos seus dobramentos locais e globais, e as suas respectivas estereoquímicas (LASKOWSKI; MACARTHUR; THORNTON, 2012; STUDER *et al.*, 2020; WIEDERSTEIN; SIPPL, 2007).

3.4 Dinâmica molecular

Após serem preditos, os modelos passaram por simulações em dinâmica molecular (DM) utilizando o pacote GROMACS 2022.4 (OANCA; ENT, VAN DER; ÅQVIST, 2023). Desse modo, os modelos tridimensionais foram centralizados em uma caixa cúbica, a uma distância de 1.0 nm, solvatados utilizando água do tipo SPC/E e, as energias dos sistemas foram minimizadas. As simulações então, foram feitas em NVT (Número de partículas, Volume e Temperatura) e NPT (Número de partículas, Pressão e Temperatura).

Assim, para preparar as dinâmicas, foi utilizado o campo de força GROMOS 53A6, com condições periódicas de contorno nos eixos x, y e z, uma vez que, esse campo foi parametrizado para a entalpia livre de hidratação, propriedade essa, diretamente ligada a uma melhor descrição sobre o equilíbrio de dobramento e desdobramento de um peptídeo (OOSTENBRINK *et al.*, 2005). Portanto, as DMs foram realizadas sem restrições, com pressão e temperatura constantes de 1 atm e 300 K respectivamente, por um tempo de total de 100 ns.

Por fim, as simulações dos modelos foram analisadas quanto a sua Raiz Quadrada do Desvio Quadrático Médio (RMSD, do inglês *Root Mean Square*

Deviation), Raiz Quadrada da Flutuação Quadrática Média (RMSF, do inglês *Root Mean Square Fluctuation*), Raio de Giro (RG) e Pontes de Hidrogênio (HBnum), além de avaliações de Potencial da Superfície Eletrostática (PSE).

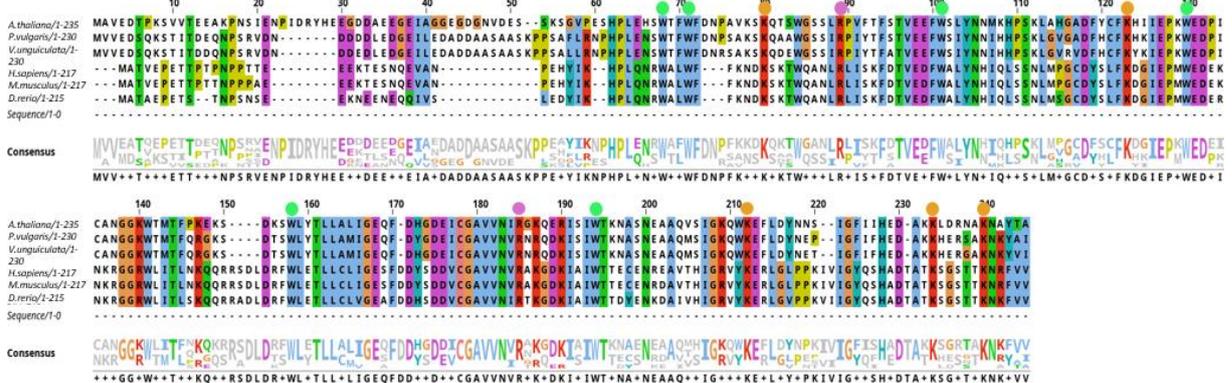
4. RESULTADOS E DISCUSSÃO

Visando uma melhor compreensão na distribuição dos resultados e da discussão, esse estudo irá detalhar as estruturas homólogas do fator de tradução eIF4E identificadas, usando como referência o número do cromossomo em que seu respectivo gene se encontra. Portanto, começando assim, pelo cromossomo 4, em seguida o 7 e, para finalizar, o cromossomo 6.

4.1 Alinhamento das sequências

Desse modo, primeiramente foi obtido os alinhamentos para o fator eIF4E (Figura 7), para tal, foram utilizadas as seguintes sequências: *P. vulgaris* (ID: 442), *V. unguiculata* (ID: 540), *Mus musculus* (UniProt ID: P63073), *Arabidopsis thaliana* (UniProt ID: O23252), *H. sapiens* (UniProt ID: P06730) e *Danio rerio* (UniProt ID: Q9DFS6).

Figura 7. Alinhamento completo das sequências do fator eIF4E de eucariotos. Em verde o Triptofanos conservados (W), em laranja as Lisinas conservadas (K) e em rosa as Argininas conservadas (R).



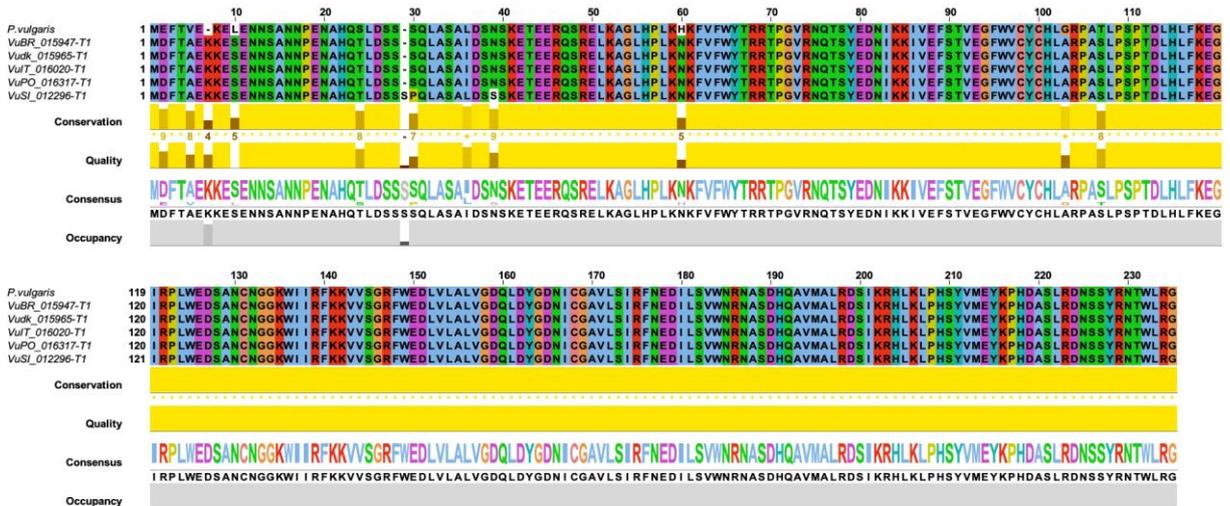
Logo, observamos diferenças entre o fator eIF4E de diferentes espécies, mas, nota-se resíduos conservados de importância para a ligação com o 5'CAP e o eIF4G, como os Triptofanos (W), as Argininas (R) e as Lisinas (K) (MIRAS *et al.*, 2017; ROSS-KASCHITZA; ALTMANN, 2020).

Seguindo essa etapa, os alinhamentos das sequências do fator eIF4E encontradas no cromossomo 4 de *V. unguiculata*, nas cultivares BR14, IT85F, Denkitiliana, Santo Inácio e Pingo de Ouro. Ademais, exclusivamente para fins de uma

melhor observação na comparação e validação, alinharmos também, a sequência do fator eIF4E de *Phaseolus vulgaris* para cada respectivo cromossomo. Essa espécie foi usada a fim de comparação, pois, compartilha de um mesmo ancestral comum e possui uma elevada homologia com a espécie *Vigna unguiculata* (VALE MARTINS, DO *et al.*, 2021).

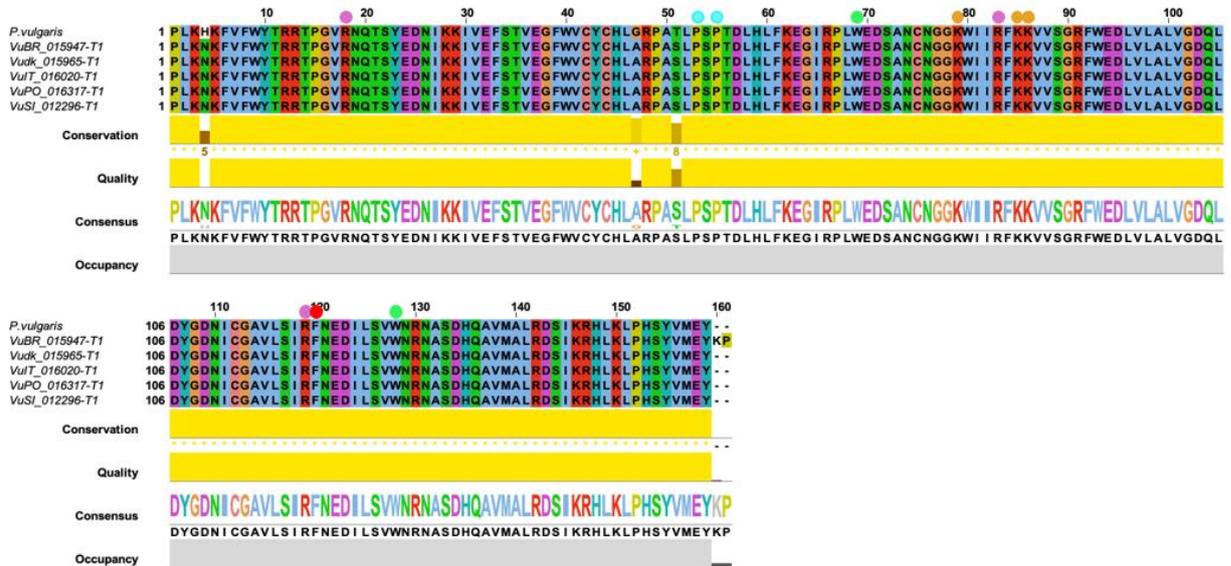
Portanto, primeiramente foi feito o alinhamento (Figura 8) e, observamos um elevado grau de conservação por toda a estrutura das cultivares, até mesmo na região do peptídeo sinal. Onde, apenas a cultivar Santo Inácio apresentou a adição de uma Serina na posição 29, uma troca de uma Serina por uma Prolina na posição 30 e uma troca de uma Asparagina por uma Serina, todas localizadas somente na região do peptídeo sinal, que termina na posição 56.

Figura 8. Alinhamento completo das sequências de eIF4E do cromossomo 4 das cultivares BR14, DEN, IT85F, SI e PO de *V. unguiculata* e da *Phaseolus vulgaris*.



Após isso, realizamos o alinhamento apenas das regiões de domínio conservado do fator de tradução do cromossomo 4 (Figura 9). Observamos no mesmo, uma corroboração do alto grau de conservação genética desses fatores, nas cultivares de *V. unguiculata*, apesar de geralmente ser comum observar algumas mutações pontuais entre cultivares da mesma espécie (MARTÍNEZ-FORTÚN; PHILLIPS; JONES, 2022). Algumas dessas mutações, podem ser fatores auxiliares das características individuais de cada cultivar, como a IT85F, que é conhecida por sua resistência ao CABMV (*Cowpea aphid-borne mosaic virus*) (SILVA, J. A. *et al.*, 2021).

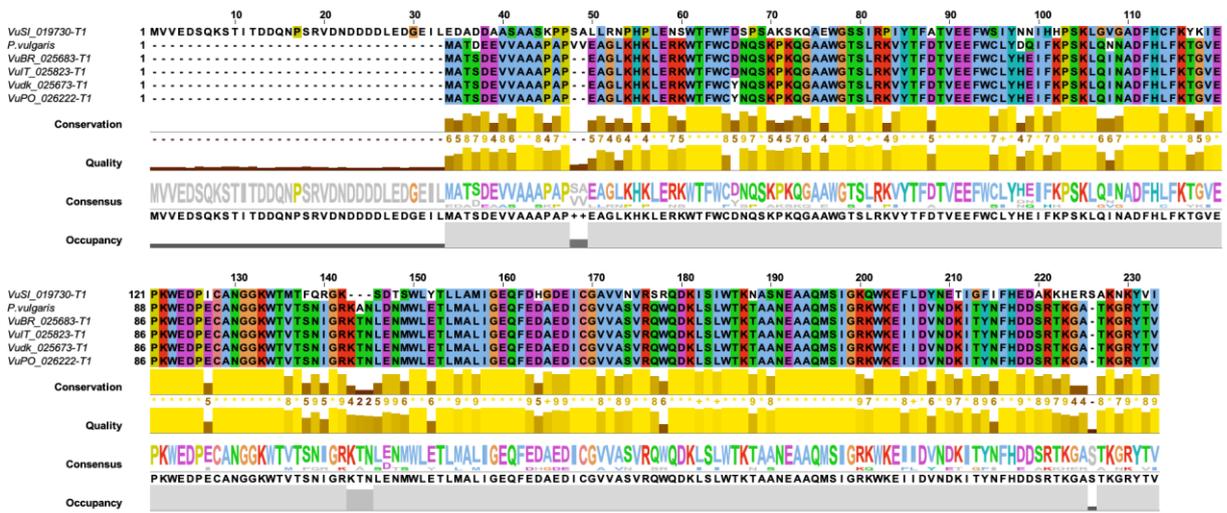
Figura 9. Alinhamento dos domínios das sequências de eIF4E do cromossomo 4 de *Phaseolus vulgaris* e das cultivares BR14, DEN, IT85F, SI e PO de *V. unguiculata*. Em destaque, sinalizados com círculos verdes, estão resíduos de aminoácidos do tipo triptofano (W), em azul resíduos de aminoácidos de prolina (P), em rosa resíduos de arginina (R), em laranja resíduos de lisina (K) e em vermelho resíduos de fenilalanina (F), os quais fazem parte do bolsão catiônico, cruciais para interação dos fosfatos da região 5' CAP.



Entretanto, nos homólogos do cromossomo 4, podemos observar que não há acentuadas mutações entre as isoformas, apresentando estruturas com elevado grau de conservação. Inclusive, com relação aos triptofanos, que são resíduos de aminoácidos preservados, por serem essenciais na atividade de ligação com a região 5'CAP do mRNA (TUPPERWAR *et al.*, 2021).

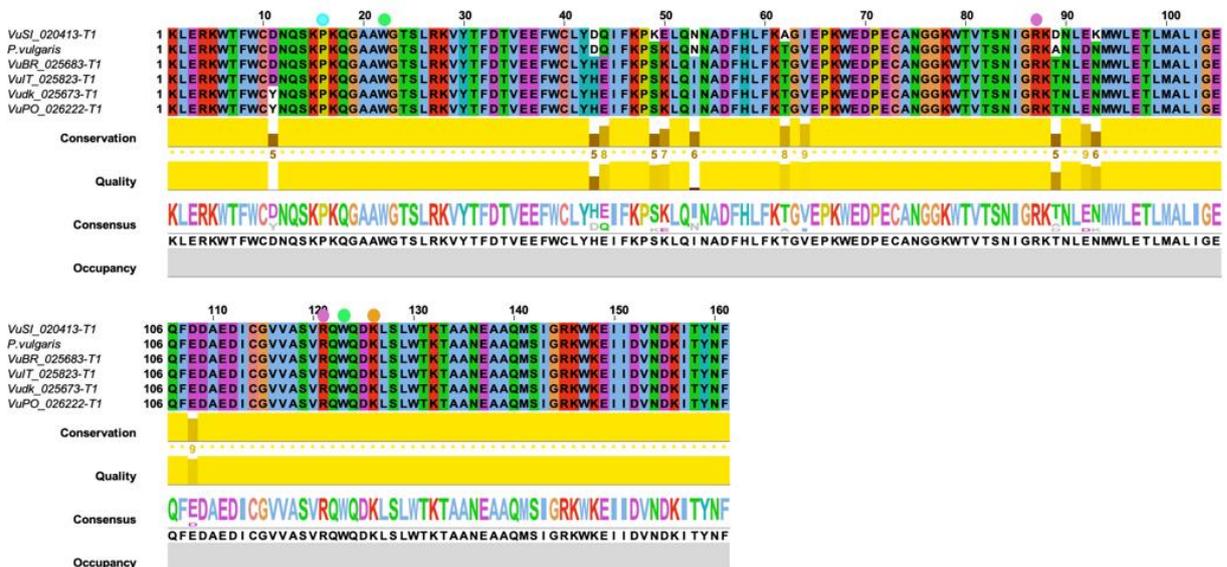
Em seguida, as isoformas do cromossomo 7 foram alinhadas, primeiramente toda a sequência com o peptídeo sinal, que termina na posição 55 (Figura 10) e, depois, o alinhamento apenas das regiões do domínio conservado (Figura 11).

Figura 10. Alinhamento completo das seqüências de eIF4E do cromossomo 7 das cultivares BR14, DEN, IT85F, SI e PO de *V. unguiculata* e da *Phaseolus vulgaris*.



Nesse alinhamento das seqüências completas, podemos observar que a isoforma do cromossomo 7 tem uma conservação menor quando comparada ao cromossomo 4, inclusive, se tratando do peptídeo sinal da cultivar Santo Inácio, a qual apresenta uma variação acentuada no tamanho, quando comparada as outras cultivares. Mas, ainda assim, o grau de correspondência entre essas seqüências de aminoácidos é alto, sendo assim, altamente similares (ZLOBIN; TARANOV, 2023b).

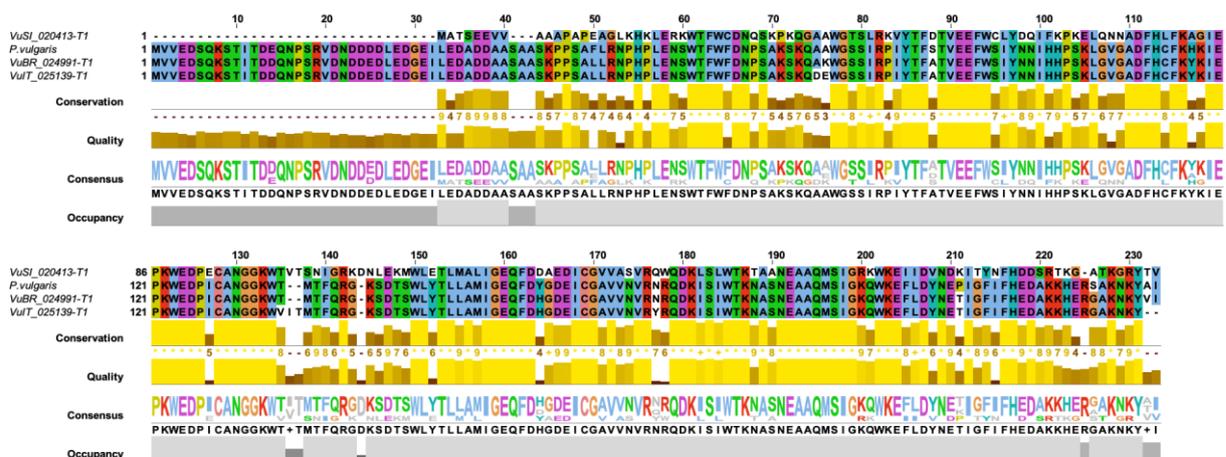
Figura 11. Alinhamento dos domínios das seqüências de eIF4E do cromossomo 7 de *Phaseolus vulgaris* e das cultivares BR14, DEN, IT85F, SI e PO de *V. unguiculata*. Em destaque, sinalizados com círculos verdes, estão resíduos de aminoácidos do tipo triptofano (W), em azul resíduos de aminoácidos de prolina (P), em rosa resíduos de arginina (R) e em laranja resíduos de lisina (K), os quais fazem parte do bolsão catiônico, cruciais para interação dos fosfatos da região 5' CAP.



Esse grau de correspondência, pode ser visualizado com uma maior facilidade no alinhamento das regiões de domínio das isoformas do cromossomo 7 (Figura 11), na qual, apresentam uma similaridade maior que 90%. Essas isoformas, apresentam também, mutações pontuais mais recorrentes que as do cromossomo 4, principalmente a cultivar Santo Inácio, a qual apresenta substituição de uma Histidina por um Ácido Aspártico na posição 43, Glutamina por ácido Glutâmico na posição 44, uma Lisina por uma Serina na posição 49, um Ácido Glutâmico por uma Lisina na posição 50, uma Asparagina por uma Isoleucina na posição 53, uma Alanina por uma Treonina na posição 62, outro Ácido Aspártico por uma Alanina na posição 89 e uma Lisina por uma Asparagina na posição 93. Outras substituições ocorrem também, como as cultivares Denkitiliana e Pingo de Ouro, que compartilham de uma substituição entre um Ácido Aspártico por uma Tirosina na posição 10 do alinhamento do domínio. Entretanto, algumas pequenas trocas ocorrem também, mas não possuem tanto impacto nas propriedades físico-químicas da estrutura, por serem resíduos de aminoácidos com propriedades semelhantes, como por exemplo, o Ácido aspártico e o Ácido glutâmico (HERRINGTON; KELLOGG, 2021) que sofrem tal substituição na posição 108 da Santo Inácio.

Por fim, alinhamos as sequências completas das isoformas recém encontradas no cromossomo 6, das cultivares BR14, IT85F e Santo Inácio (Figura 12).

Figura 12. Alinhamento completo das sequências de eIF4E do cromossomo 6 de *Phaseolus vulgaris* e das cultivares BR14, IT85F e SI de *V. unguiculata*.

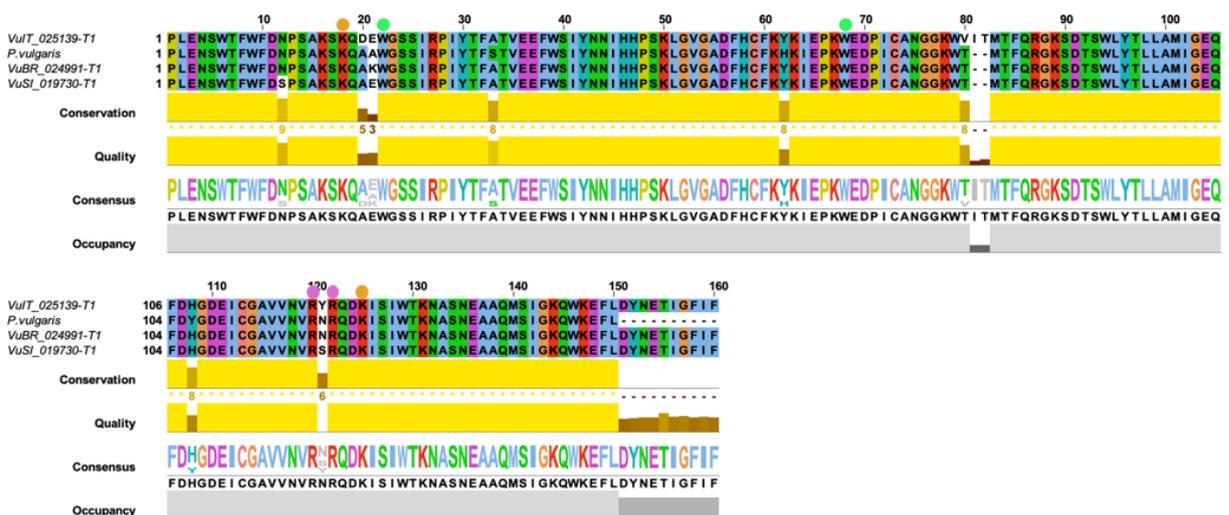


Nota-se, que entre todas as isoformas, essas são as que apresentam mais mutações entre si. Começando pelo peptídeo sinal, que termina na posição 55, na

qual a Santo Inácio possui uma estrutura reduzida, em contrapartida a sua isoforma do cromossomo 7, como também, apresenta mais mutações nessa região, comparado as outras cultivares. As quais, apresentam um elevado grau de homologia da sua sequência na região do peptídeo sinal, inclusive com a *P. vulgaris*.

Em seguida, foi realizado o alinhamento apenas da região de domínio do cromossomo 6 (Figura 13).

Figura 13. Alinhamento dos domínios das sequências de eIF4E do cromossomo 6 de *Phaseolos vulgaris* e das cultivares BR14, IT85F e SI de *V. unguiculata*. Em destaque, sinalizados com círculos verdes, estão resíduos de aminoácidos do tipo triptofano (W), em rosa resíduos de arginina (R) e em laranja resíduos de lisina (K), os quais fazem parte do bolsão catiônico, cruciais para interação dos fosfatos da região 5' CAP.



Essas isoformas, apresentam elevadas mutações pontuais entre si, quando comparado as estruturas dos outros cromossomos, especialmente a cultivar IT85F. Portanto, a primeira modificação pontual foi apresentada pela cultivar Santo Inácio por uma substituição de uma Serina por uma Asparagina na posição 12, seguida pela IT85F que faz a troca de uma Alanina por um Ácido Aspártico na posição 20, uma mutação da BR14 de um Ácido Glutâmico por uma Lisina na posição 21, uma substituição de uma Treonina por uma Valina na posição 80 pela IT85F e, a adição de uma Isoleucina e uma Treonina nas posições 81 e 82 respectivamente, apresentada apenas por essa mesma cultivar. Em especial, essas cultivares apresentaram uma grande diferença entre si, na posição 121 do seu domínio estrutural, onde, cada isoforma possui um resíduo de aminoácido diferente. A IT85F por sua vez, contém uma tirosina, a Santo Inácio apresenta uma Serina, e a BR14, em concordância com

a *P. vulgaris*, exibe uma Asparagina nessa região.

4.3 Modelagem molecular

Para realizar a predição dos modelos estruturais das cultivares, foi utilizado o servidor do AlphaFold2, conforme dito anteriormente. Estruturalmente, todos os modelos possuem uma visível similaridade, sem grandes alterações conformacionais, mesmo entre modelos de diferentes cromossomos.

Assim, temos as estruturas modeladas do cromossomo 4, começando com o modelo da cultivar BR14 (Figura 14), que será usado como referência para a visualização dos outros modelos desse cromossomo (Figura 15).

Figura 14. Modelo teórico estrutural do cromossomo 4 da cultivar BR14 representado em cartoon. Em azul as Alfa-hélices e em vermelho as Beta-fitas.

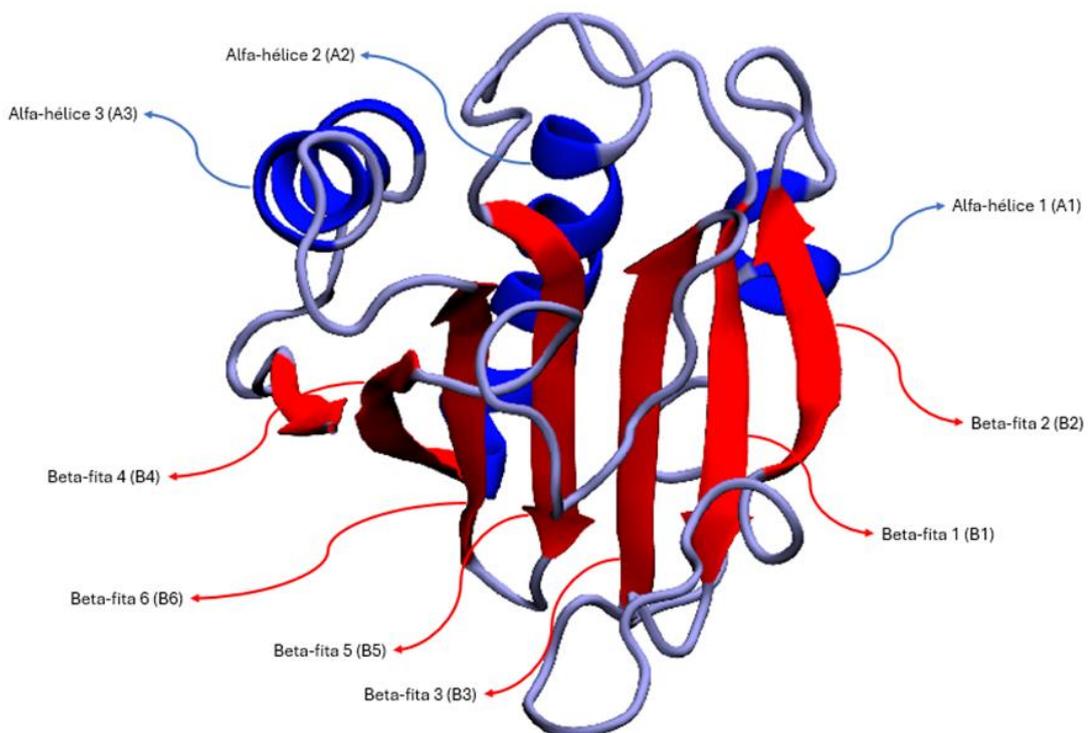
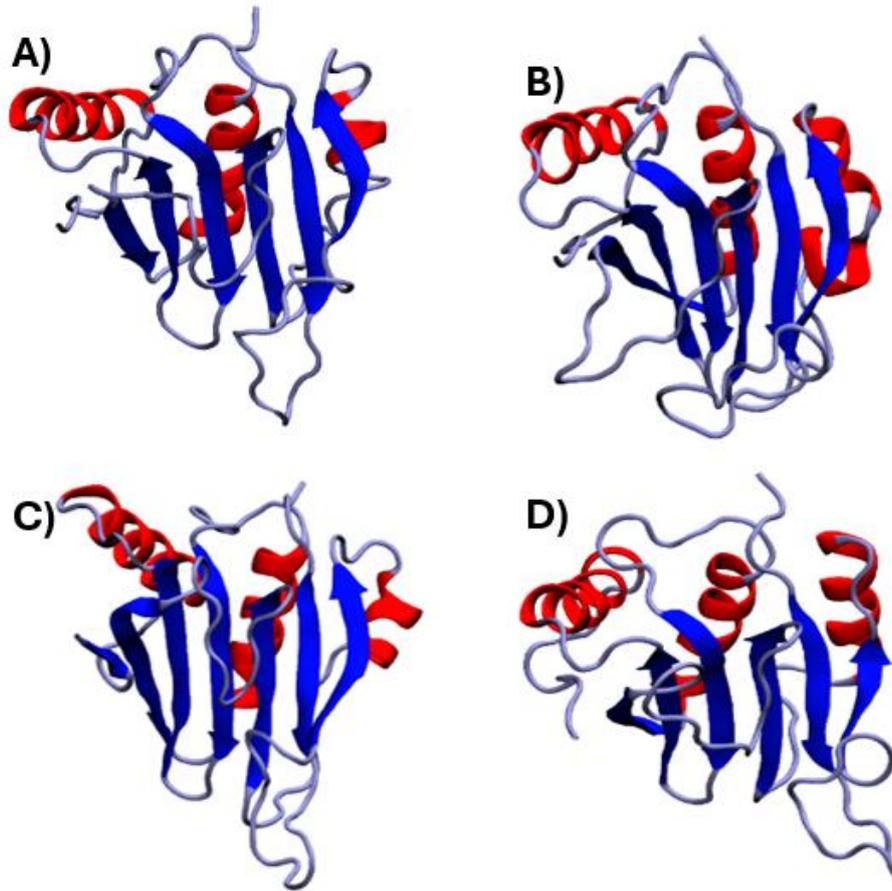


Figura 15. Modelos estruturais do cromossomo 4. Em (A) a cultivar Denkitiliana, (B) a cultivar IT85F, (C) a cultivar Pingo de Ouro e (D) a cultivar Santo Inácio. Representado em vermelho as Alfa-hélices e em azul as Beta-fitas.



Em seguida, as estruturas modeladas do cromossomo 7, começando pela estrutura da cultivar BR14 (Figura 16), de referência para as outros modelos desse cromossomo.

Figura 16. Modelo estrutural do cromossomo 7 da cultivar BR14. Representado em azul as Alfa-hélices e em vermelho as Beta-fitas.

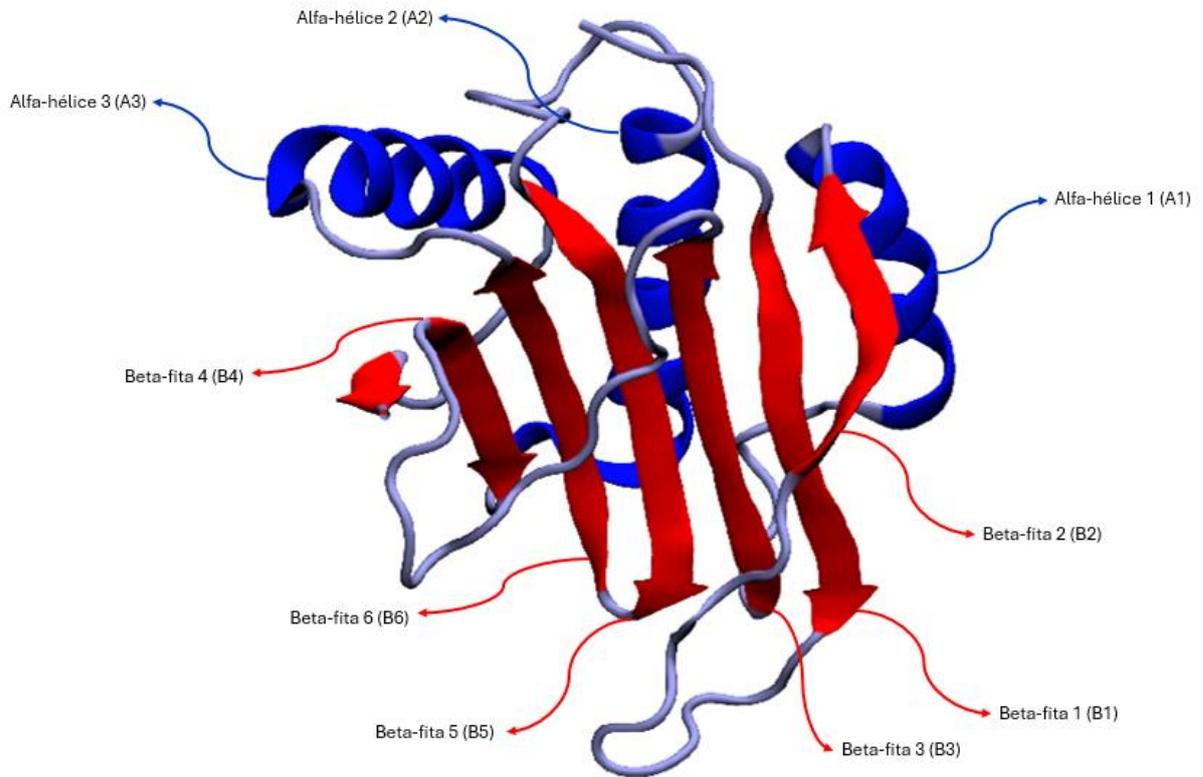
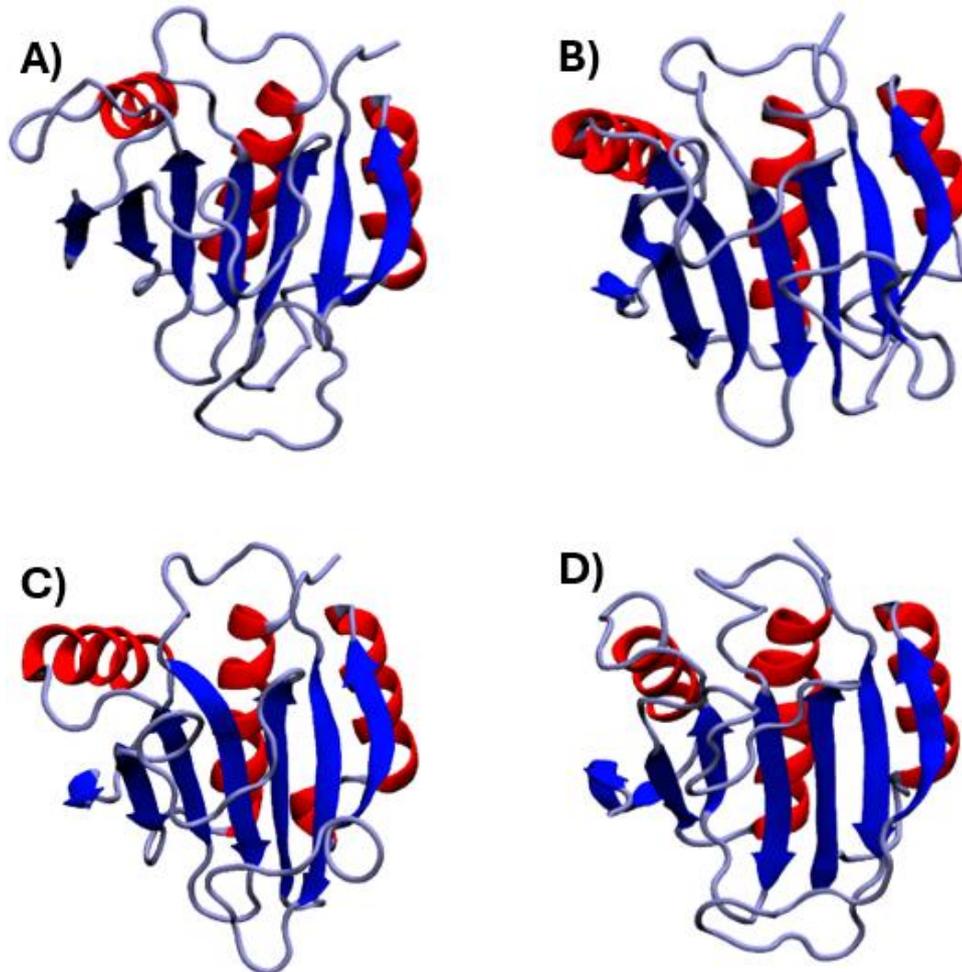
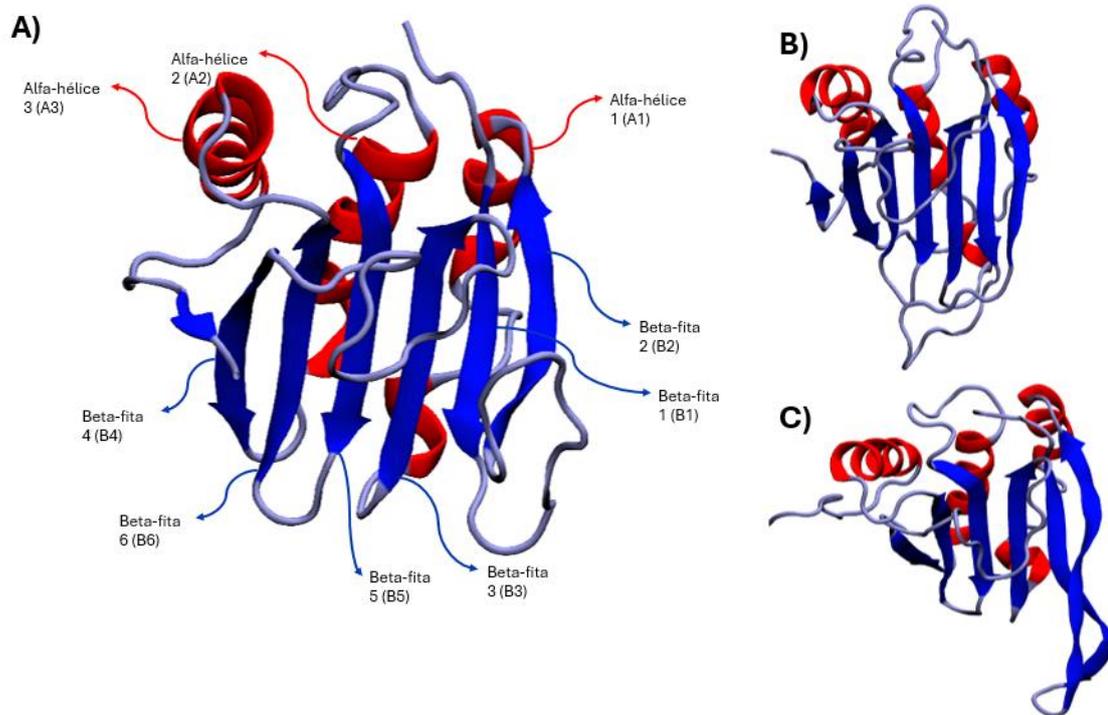


Figura 17. Modelos estruturais do cromossomo 7. Em (A) a cultivar Denkitiliana, (B) a cultivar IT85F, (C) a cultivar Pingo de Ouro e (D) a cultivar Santo Inácio. Representado em vermelho as Alfa-hélices e em azul as Beta-fitas.



E, para finalizar, tem-se as modelagens do cromossomo 6 das cultivares BR14, IT85 e Santo Inácio de *Vigna unguiculata*. Essas estruturas seguem o mesmo padrão dos cromossomos anteriores, no entanto, a cultivar Santo Inácio (Figura 18 “C”) apresentou uma característica distinta, um alongamento das Beta-fitas 1 e 2, nas regiões ocupadas pelos resíduos de número 12 ao 15 e 19 a 25. Alongamento esse, provavelmente acarretado pela mutação da serina pela asparagina na posição 12 e da mutação do ácido glutâmico na posição 21, com a manutenção da alanina na posição 20, diferentemente da cultivar IT85F, que faz a substituição por um ácido aspártico nessa posição.

Figura 18. Modelo estrutural do cromossomo 6 da cultivar BR14 (A), IT85F (B) e Santo Inácio (C). Representado em vermelho as Alfa-hélices e em azul as Beta-fitas.

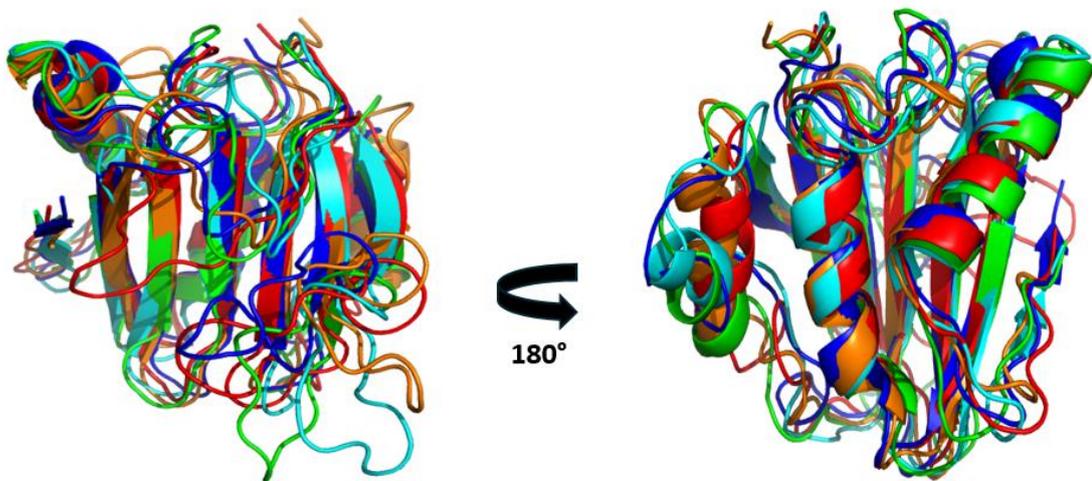


4.2 Validação dos Modelos Estruturais

Para um melhor entendimento da conservação das sequências alinhadas anteriormente, foram realizadas as modelagens e os alinhamentos de todas as estruturas de eIF4E, presentes nos respectivos cromossomos das cultivares de *Vigna unguiculata*. Tais modelagens foram realizadas com base nos resultados obtidos pelo software CD-Search, incluindo, apenas o intervalo de resíduos de aminoácidos entre 56-217 no cromossomo 4 e, 55-216 nos cromossomos 6 e 7, região de domínio estrutural, excluindo a região N terminal não estruturada. Apesar das mutações evolutivas que podem ocorrer naturalmente, quando alinharmos as estruturas, é possível observar uma similaridade elevada entre os modelos do cromossomo 4 (Figura 19), assim como mostram os resultados dos alinhamentos de sequência, apresentando como base global da estrutura da proteína eIF4E, 3 alfa-hélices e 6 beta-fitas. Visto que, a estrutura é mais conservada do que a sequência, corroborando com a já mencionada conservação estrutural nesse fator de tradução eucariótico (COUTINHO DE OLIVEIRA *et al.*, 2019; KINKELIN *et al.*, 2012). Devido a essa alta

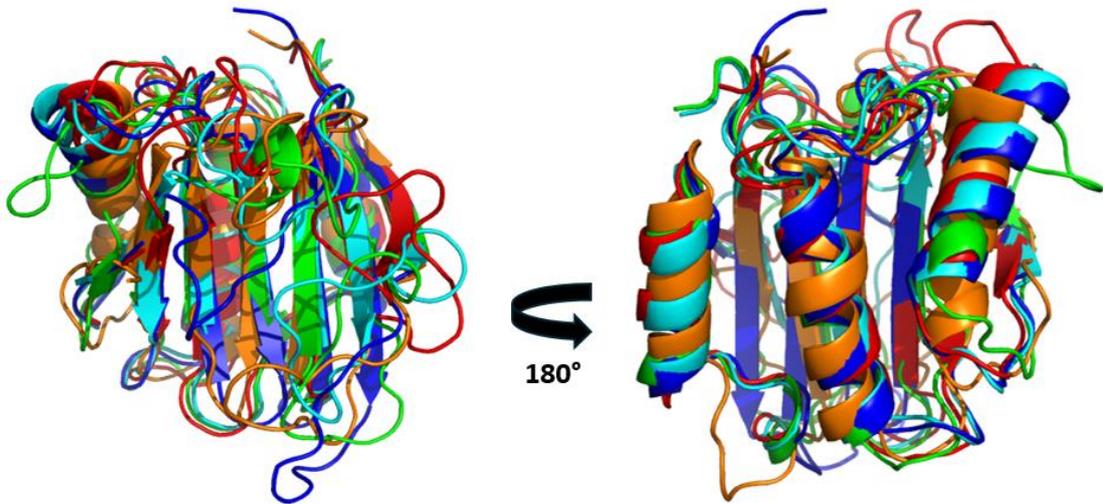
conservação, a estrutura da eIF4E entre os eucariotos, é comumente descrita na literatura como “mão em concha” (do inglês, cupped hand), e adquire diferenças mínimas geralmente em regiões desordenadas (sem estrutura secundária definida) (KINKELIN *et al.*, 2012; VOLPON *et al.*, 2006). Como, por exemplo, a diferença na flutuação entre as suas alças.

Figura 19. Representação em cartoon do alinhamento dos modelos do cromossomo 4 das cultivares BR14 em azul, Denkitiliana em verde, IT85F em vermelho, Pingo de Ouro em turquesa e Santo Inácio em laranja.



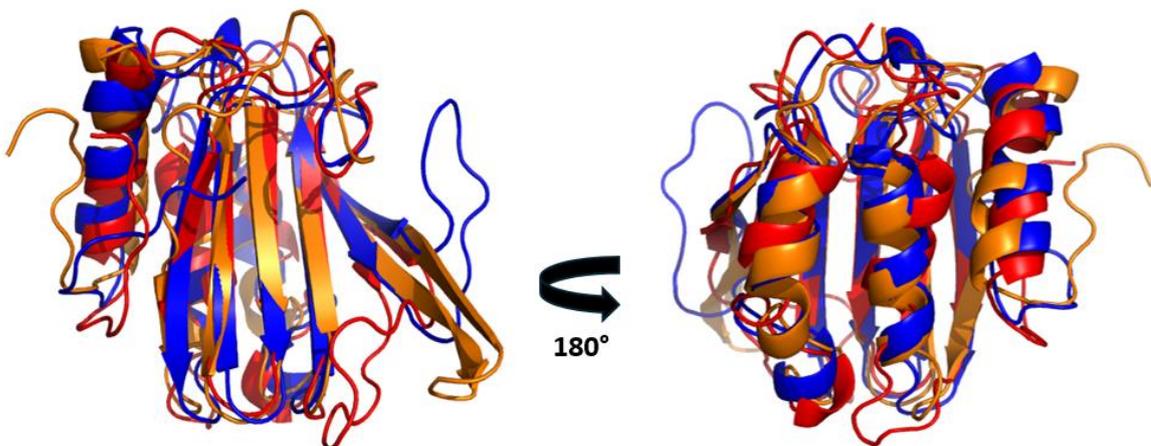
Por conseguinte, as estruturas do cromossomo 7, também mostraram elevado grau de conservação entre os modelos e, algumas pequenas diferenças de flexibilidade das alças, observadas também no cromossomo anterior, como podemos ver no alinhamento entre elas (Figura 20). No entanto, em comparação com o cromossomo 4, as estruturas do cromossomo 7 apresentam elevada similaridade quanto a sua conformação estrutural.

Figura 20. Representação em cartoon do alinhamento dos modelos do cromossomo 7 das cultivares BR14 em azul, Denkitiliana em verde, IT85F em vermelho, Pingo de Ouro em turquesa e Santo Inácio em laranja.



Por fim, os modelos alinhados do cromossomo 6 das cultivares BR14, IT85F e Santo Inácio (Figura 21). Essas estruturas, apresentam da mesma forma, um elevado grau de similaridade entre elas, entretanto, por possuírem mais mutações pontuais entre as suas sequencias, podemos observar algumas diferenças entre os modelos. Tais diferenças, são principalmente observadas nas conformações distintas das flutuações entre as suas alças e as suas estruturas secundárias. Além de que, a cultivar Santo Inácio, particularmente, apresenta as suas 2 primeiras beta-fitas, com um maior grau de alongamento, quando comparadas com as cultivares anteriores.

Figura 21. Representação em cartoon do alinhamento dos modelos do cromossomo 6 das cultivares BR14 em azul, IT85F em vermelho e Santo Inácio em laranja.



Visando avaliar a qualidade dos modelos gerados anteriormente, foram avaliados três principais parâmetros, sendo estes: a distribuição dos ângulos de torção pelo gráfico de Ramachandran (PROCHECK), o Z-Score que avalia a qualidade do modelo pela energia total da estrutura (PRO-SA) e o QMEANDisCo score, que avalia a qualidade geométrica local e global dos modelos (SwissModel) (LASKOWSKI; MACARTHUR; THORNTON, 2012; STUDER *et al.*, 2020; WIEDERSTEIN; SIPPL, 2007).

Os resultados obtidos pelo PROCHECK, revelam que todas as modelagens do cromossomo 4 e 6, possuem valores acima de 94%, sobre os resíduos preditos em regiões onde os ângulos de torção são termodinamicamente favoráveis. Os modelos do cromossomo 4, foram os que apresentaram maior índice nessa avaliação, com as cultivares IT85F, DEN, SI e PO empatadas com um valor de 96.5% e, entre esses 2 cromossomos, o modelo com menor valor foi do cromossomo 6 pela cultivar IT85F, com 94.4%. Já, para as modelagens do cromossomo 7, o maior valor obtido foi pela cultivar Santo Inácio, com o índice de 93.2%, e, esses modelos, apresentaram valores acima 90% para essa validação, com exceção da cultivar Pingo de Ouro, a qual evidenciou um valor de 89.7%, o menor valor entre todos os modelos preditos. Os respectivos valores do PROCHECK, estão disponíveis na tabela 1 e, os gráficos de Ramachandran, estão disponíveis em Anexo.

As análises observadas no PRO-SA, para todos os modelos, revelaram pontuações de Z-Score entre -6.34 a -5.41, que indicam estruturas confiáveis e, dentro do intervalo de pontuações comumente encontrado, de 0 a -10, para proteínas de tamanhos semelhantes, mais de 100 resíduos (WIEDERSTEIN; SIPPL, 2007). Os respectivos valores de Z-Score dos modelos, estão disponíveis na tabela 1 e, as suas representações gráficas, disponíveis no apêndice.

Por fim, os dados resultantes das análises do QMEAN-DisCo (Tabela 1), mostraram que todos as estruturas modeladas, apresentam qualidade na sua construção geométrica local e global, corroborando com os dados dos gráficos de remancharam. Esses dados, revelam um intervalo de 0.76 a 0.80 para os cromossomos 4 e 7, e, 0.83 a 0.87 entre as estruturas do cromossomo 6, as quais se espera uma qualidade levemente maior dos modelos, segundo esse parâmetro, por apresentar valores que mais se aproximam de 1 (STUDER *et al.*, 2020). Os gráficos resultantes do QMEAN-DisCo, estão disponíveis em Anexo.

Tabela 1. Descrição de scores de validação das estruturas modeladas. Na primeira coluna encontram-se a identificação das estruturas, na segunda coluna os *scores* do QMEAN-DisCO, na terceira coluna os valores retomados pelo PRO-SA e por fim os valores obtidos no PROCHECK.

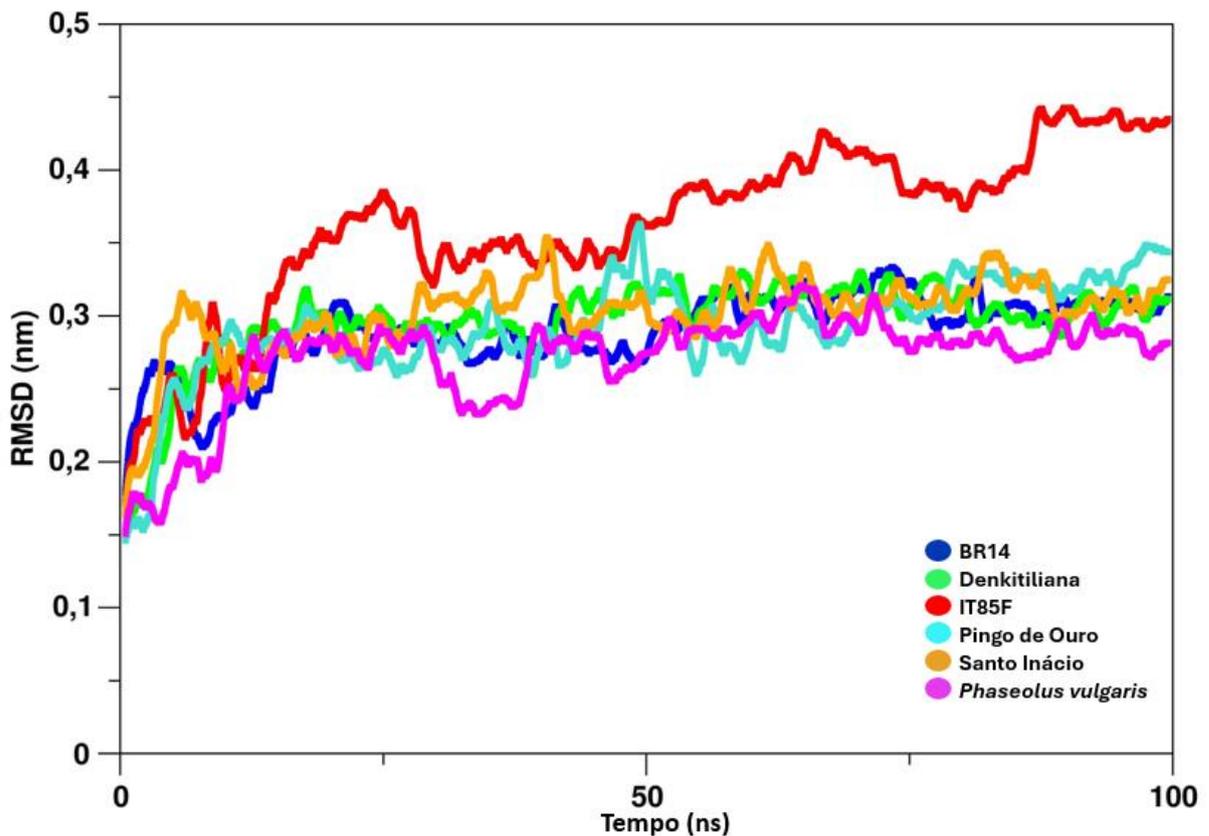
| BR14 | QMEAN-DisCo | Z-Score | PROCHECK |
|----------------------|--------------------|----------------|-----------------|
| -CH4- | 0.80 | -5.57 | 95.8 |
| -CH6- | 0.87 | -6.28 | 96.4 |
| -CH7- | 0.77 | -5.49 | 92.5 |
| IT85F | QMEAN-DisCo | ZScore | PROCHECK |
| -CH4- | 0.79 | -5.46 | 96.5 |
| -CH6- | 0.83 | -6.17 | 94.4 |
| -CH7- | 0.76 | -5.51 | 92.5 |
| Santo Inácio | QMEAN-DisCo | Z-Score | PROCHECK |
| -CH4- | 0.79 | -5.38 | 96.5 |
| -CH6- | 0.86 | -6.34 | 95.7 |
| -CH7- | 0.77 | -5.93 | 93.2 |
| Denkitiliana | QMEAN-DisCo | Z-Score | PROCHECK |
| -CH4- | 0.79 | -5.42 | 96.5 |
| -CH7- | 0.76 | -5.48 | 91.1 |
| Pingo de Ouro | QMEAN-DisCo | Z-Score | PROCHECK |
| -CH4- | 0.79 | -5.41 | 96.5 |
| -CH7- | 0.76 | -5.50 | 89.7 |

4.4 Dinâmica molecular

A etapa de dinâmica molecular, envolve 4 análises distintas, RMSD, RMSF, HB

e RG, para cada isoforma encontrada nos respectivos cromossomos das cultivares. Desse modo, as análises foram feitas pelo agrupamento das isoformas de cada cromossomo respectivamente. Portanto, para iniciar, temos o gráfico do RMSD das isoformas do cromossomo 4 (Figura 22).

Figura 22. Gráfico de RMSD do cromossomo 4. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



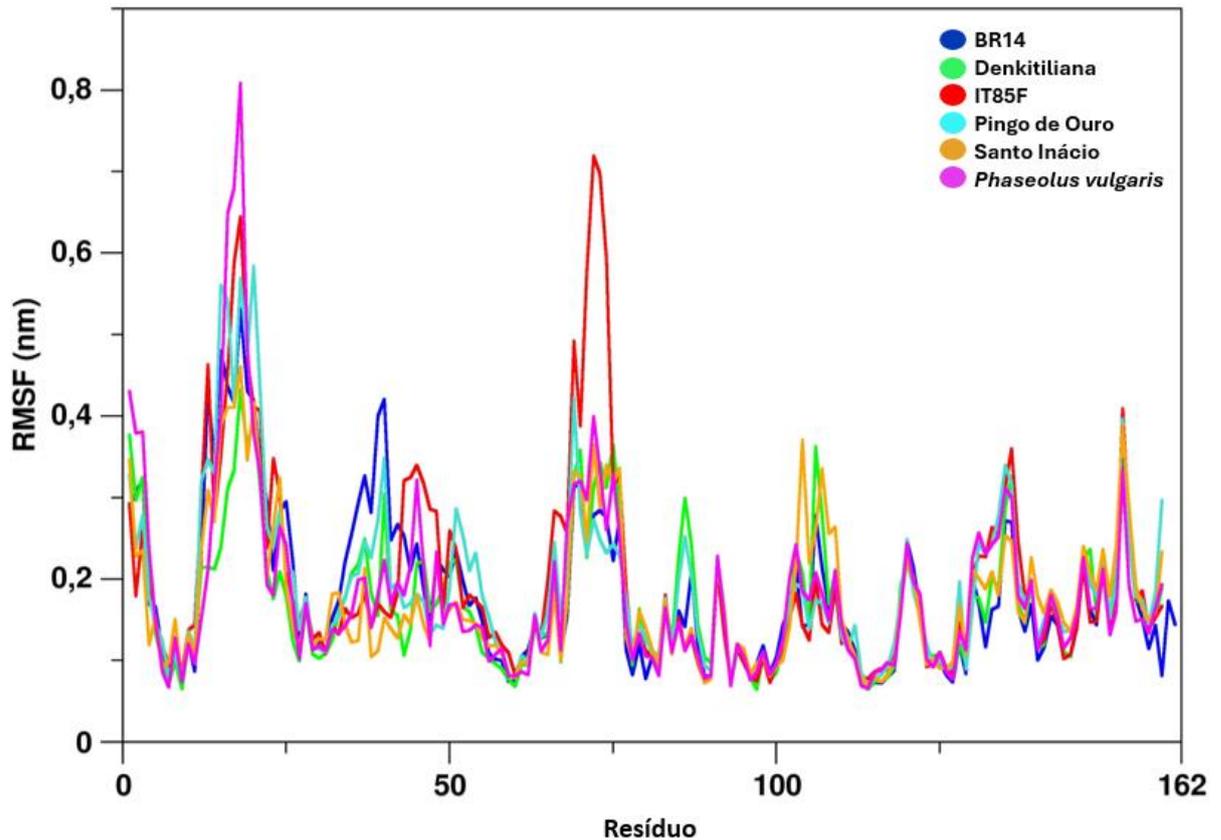
A partir desse gráfico, podemos observar que todas as isoformas conseguiram alcançar um grau de convergência estrutural por volta dos 50 ns, com uma generalidade de variação de aproximadamente 1,5 nm, confirmando assim a qualidade dos modelos teóricos preditos, e, com isso, sem extremas modificações nas suas estruturas. No entanto, a cultivar IT85F apresentou particularidades levemente diferentes das outras cultivares, como por exemplo, um número mais elevado de rearranjos estruturais para alcançar a sua convergência, com uma variação de aproximadamente 2,5 nm no período de 100 ns.

Por conseguinte, temos a análise da dinâmica do RMSF (Figura 23) das estruturas do cromossomo 4. Esse gráfico permite a visualização da flexibilidade de

cada aminoácido da estrutura, permitindo uma análise geral e pontual.

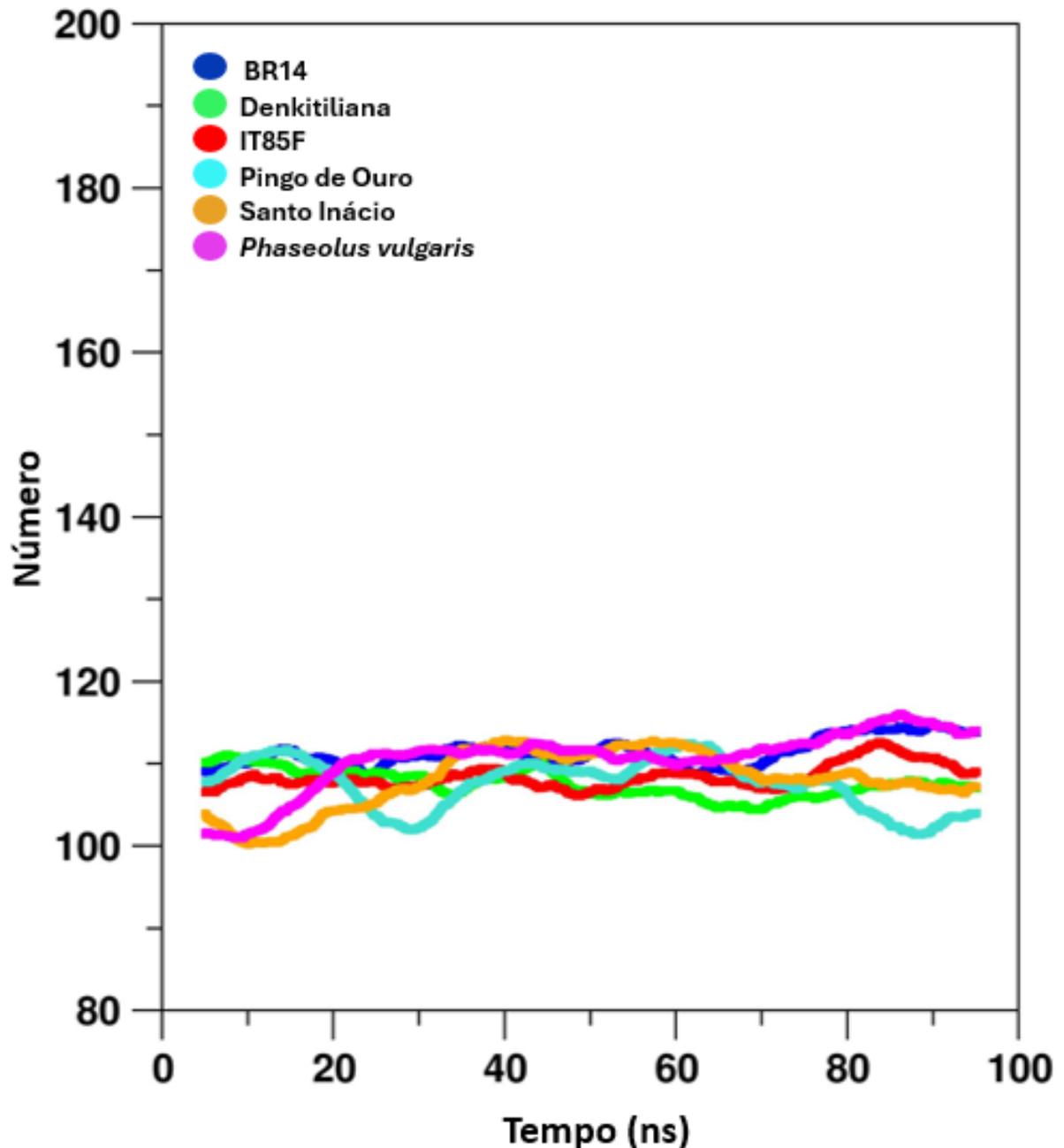
Portanto, para as características gerais apresentadas pelas cultivares, pode-se destacar que todos os sistemas mostraram perfis de flexibilidade semelhantes, apresentando os maiores pontos de flutuação localizados nas alças dessas proteínas, principalmente, na primeira alça da estrutura, posição 18, entre a primeira beta-fita (B1) e a segunda beta-fita (B2), e, na alça da posição 69 – 70, a qual também é o segundo pico de flexibilidade para elas. Importante também, um olhar especial para a Arginina conservada nessa posição 18, a qual é presente em todas as cultivares, com um elevado grau de flexibilidade. Outra característica comum dessas proteínas, é o Triptofano conservado na posição 40 de todas elas e que apresenta um certo grau elevado de flexibilidade nas estruturas da BR14, Denkitiliana e Pingo de Ouro. Mesmo as sequências dessas proteínas sendo idênticas, elas ainda possuem algumas características individuais, como, a cultivar BR14, a qual possui seu maior pico de flutuação, com seus respectivos resíduos, na posição R-13, P-15, R-18, que formam a região de positividade para a ligação com o 5'CAP, e o triptofano na posição 40, também presente nessa região de ligação. Para a cultivar Denkitiliana, os resíduos de importância mais flexíveis nessa região positiva de ligação, são os resíduos R-18, Q-20 e o W-40, completamente conservados. Já a cultivar IT85F, é a que possui maior grau de flexibilidade entre as isoformas de *V. unguiculata*, apresentando os resíduos de flutuação comum entre elas, na região de positividade, R-13, R-18 e, possui uma região de alça, ainda mais flexível que a primeira, composta pelos resíduos Serina na posição 72 e Alanina na posição 73. Essa região não carrega nenhuma mutação quando comparado com as outras cultivares, sendo este fenômeno influenciado por algum resíduo da vizinhança. Uma importante contribuição entrópica vem da maior flexibilidade de um sistema, o que conseqüentemente interfere caso o mesmo partisse de interações com outros sistemas (ALVAREZ-GARCIA; BARRIL, 2014; AMARAL *et al.*, 2017). Por conseguinte, os resíduos mais flexíveis da cultivar Pingo de Ouro são, além da Arginina na posição 18, a P-15, Q-20, W-40 e o W-69. Por último, a cultivar Santo Inácio, contém os seguintes resíduos de maior flexibilidade: R-18, Q-20, E-24, W-69 e Q-104. No geral, esses dados de flexibilidade corroboram com o que está descrito na literatura, onde mostra que a estrutura das eIF4E possuem um domínio estrutural altamente estável (ESSEGIAN *et al.*, 2021; PAPADOPOULOS *et al.*, 2014; ROMAGNOLI *et al.*, 2021).

Figura 23. Gráfico de RMSF do cromossomo 4. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Em seguida, para corroborar com os fatos anteriores, temos as análises das Pontes de Hidrogênio (HB) (Figura 24), na qual podemos observar uma manutenção dessas ligações, em todas as cultivares. As mesmas, mantiveram um nível de estabilidade em relação a síntese e decomposição das pontes de hidrogênio, no decorrer dos 100 ns.

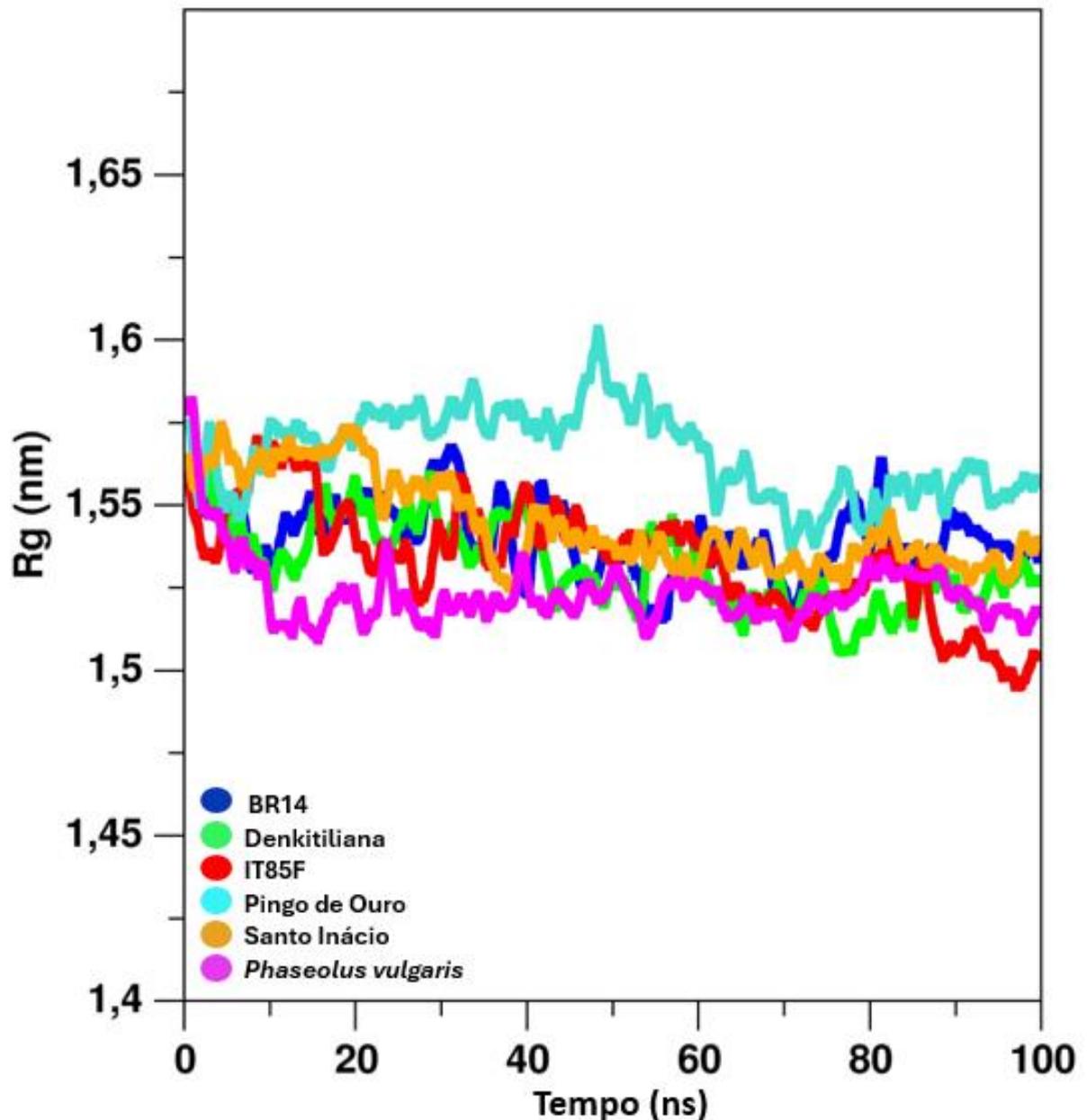
Figura 24. Gráfico de Pontes de Hidrogênio do cromossomo 4. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Para finalizar as análises do cromossomo 4, temos o gráfico do Raio de Giro (RG) (Figura 25). Nele, conseguimos avaliar que todas as estruturas das cultivares desse cromossomo, não sofreram grandes alterações conformacionais, sendo a sua forma inicial muito similar à sua forma final, de convergência estrutural, com uma

variação de no máximo 0,075nm. Ademais, todas as isoformas mostraram um nível elevado de compactação, em comparação com a estrutural inicial, corroborando com os gráficos anteriores, de uma estabilidade elevada das estruturas.

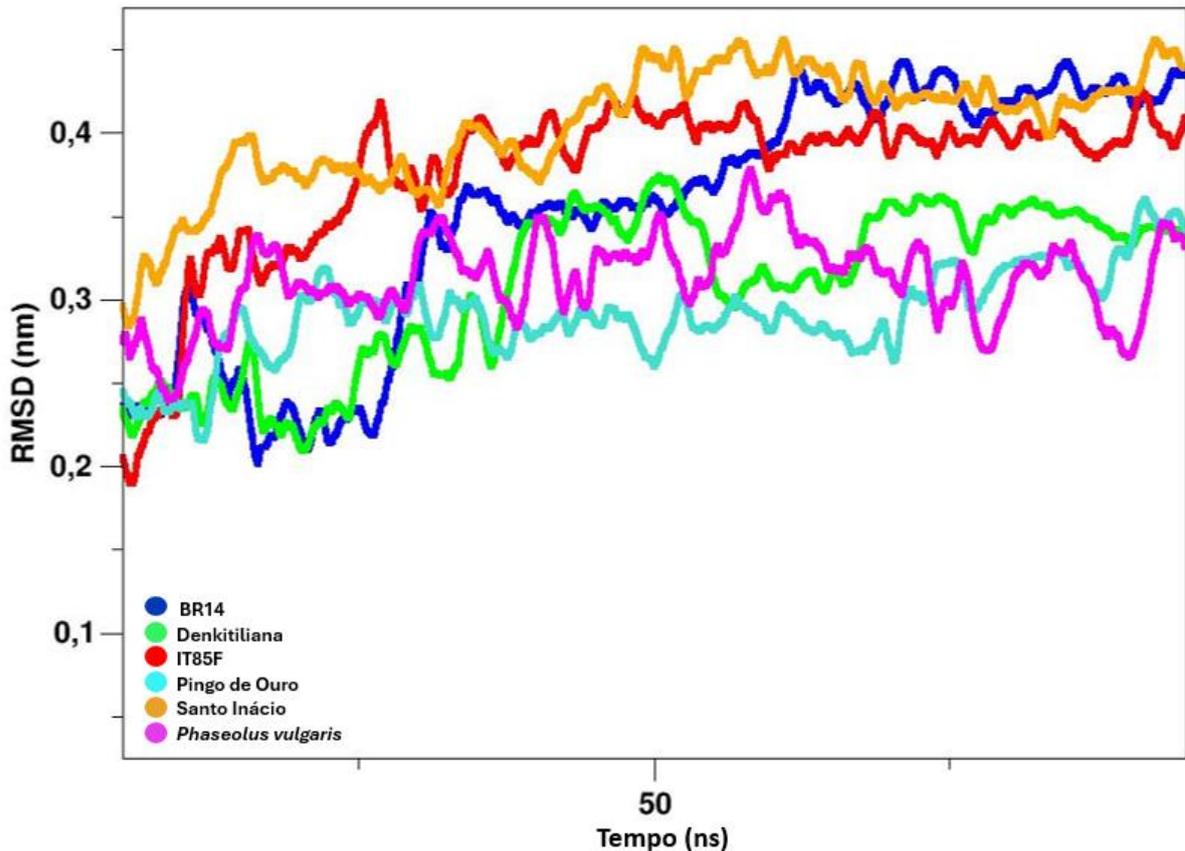
Figura 25. Gráfico de Raio de Giro do cromossomo 4. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Seguidamente, as avaliações das dinâmicas do cromossomo 7, começando pelo gráfico do RMSD (Figura 26). A partir do mesmo, podemos observar que todas as cultivares desse cromossomo, alcançaram um nível de convergência estrutural, por

volta dos 50 ns, como as do cromossomo 4. Entretanto, diferentemente das isoformas anteriores, as isoformas BR14 e SI do cromossomo 7, tiveram um grau mais elevado de rearranjos estruturais para conseguir se convergir.

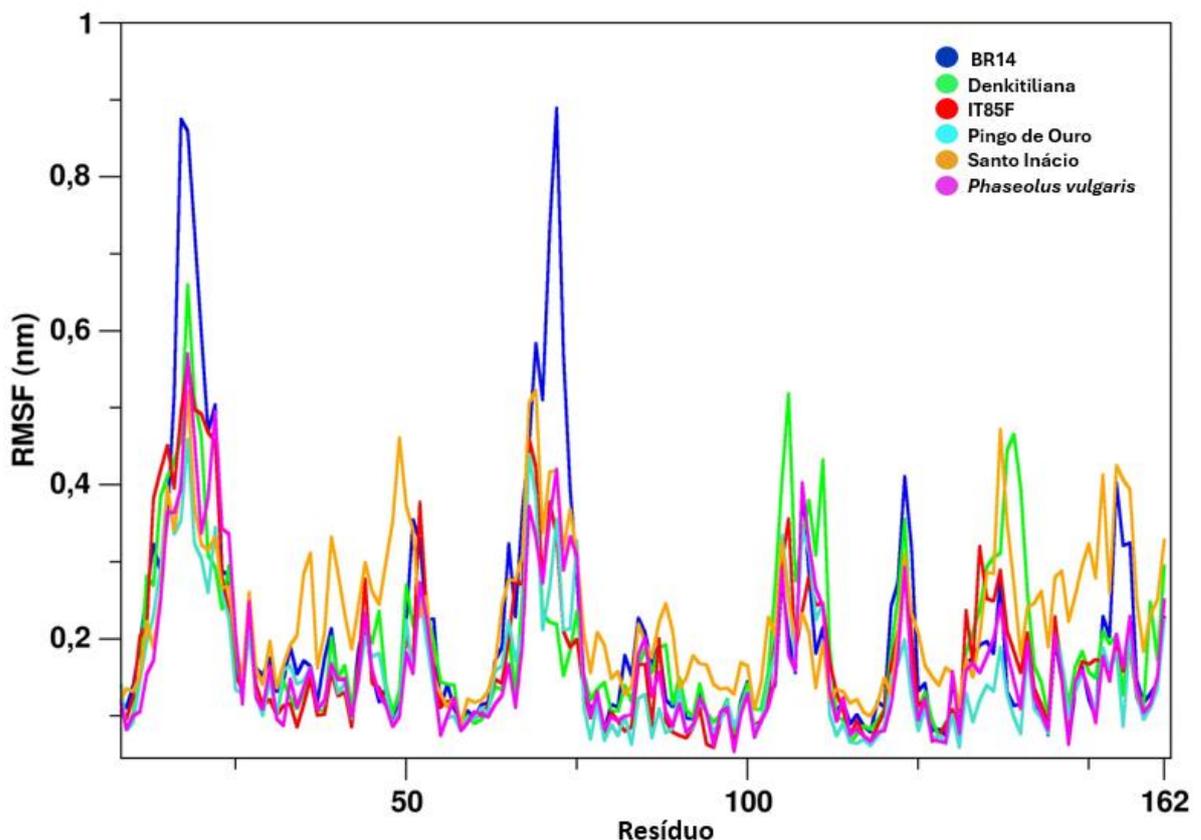
Figura 26. Gráfico de RMSD do cromossomo 7. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Logo após, o gráfico das análises do RMSF (Figura 27). De modo geral, as cultivares do cromossomo 7, se mostraram mais flexíveis que as do cromossomo 4, as quais apresentam 4 regiões de alta flexibilidade acima de 0,4 nm, enquanto, as cultivares do cromossomo 7, apresentam 6 regiões. Assim como as cultivares anteriores, essas estruturas apresentam seus pontos de maior flutuação nas alças, principalmente, entre a primeira e a segunda beta-fita (B1) e (B2), respectivamente. Essas isoformas, também compartilham algumas regiões de pico de flexibilidade entre si, como, a Glutamina na posição 18, presente em todas essas estruturas e, a Lisina na posição 15, com elevada flexibilidade nas cultivares IT85F, PO e SI. Mas, diferentemente do cromossomo anterior, a estrutura mais flexível do cromossomo 7 é a cultivar BR14, a qual possui seus picos de flexibilidade nas posições K-17, Q-18, E-

69 e E-72, que é o resíduo com o maior grau de flutuação dessa estrutura. Em seguida, a isoforma da cultivar Denkitiliana, mostrou-se com elevada flexibilidade, principalmente nos pontos Q-18, W-68, Q-106 e E-111. Após, a isoforma do cromossomo 7 da IT85F, mostrou-se como uma das estruturas menos flexíveis entre todas, em contrapartida com a isoforma do cromossomo 4, e, possui seus pontos mais flexíveis nas posições K-15, Q-18, Q-52 e W-68. Seguidamente, a cultivar Pingo de Ouro, tem seus pontos de maior flexibilidade nas posições K -15, Q-18, W-22 e W-68. Por fim, a isoforma da cultivar Santo Inácio, possui pontos de maior flutuação nas posições K-15, Q-18, K-49 e W-68.

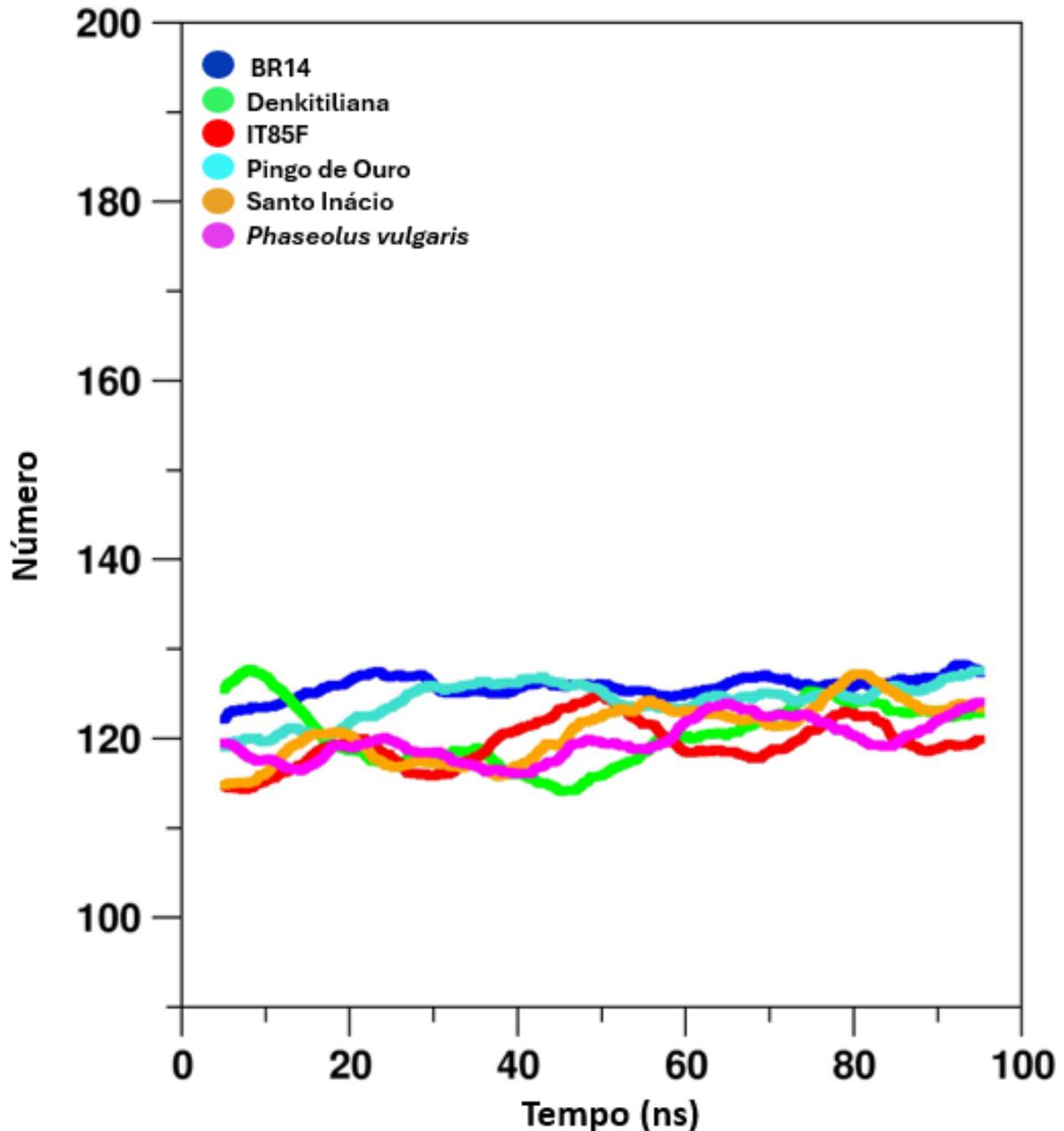
Figura 27. Gráfico de RMSF do cromossomo 7. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Prontamente, as análises das Pontes de Hidrogênio (HB) das isoformas do cromossomo 7, podem ser observadas no gráfico abaixo (Figura 28). A partir do mesmo, pode-se avaliar que essas isoformas, assim como as anteriores, apresentam um grau de estabilidade em relação as pontes de hidrogênio e, tiveram uma geral conservação da quantidade dessas ligações, em meio aos processos de síntese e

decomposição das mesmas, ao longo de 100 ns. No entanto, uma característica comum nas isoformas do cromossomo 7, é um número global maior de pontes de hidrogênio, quando comparadas às isoformas do cromossomo 4.

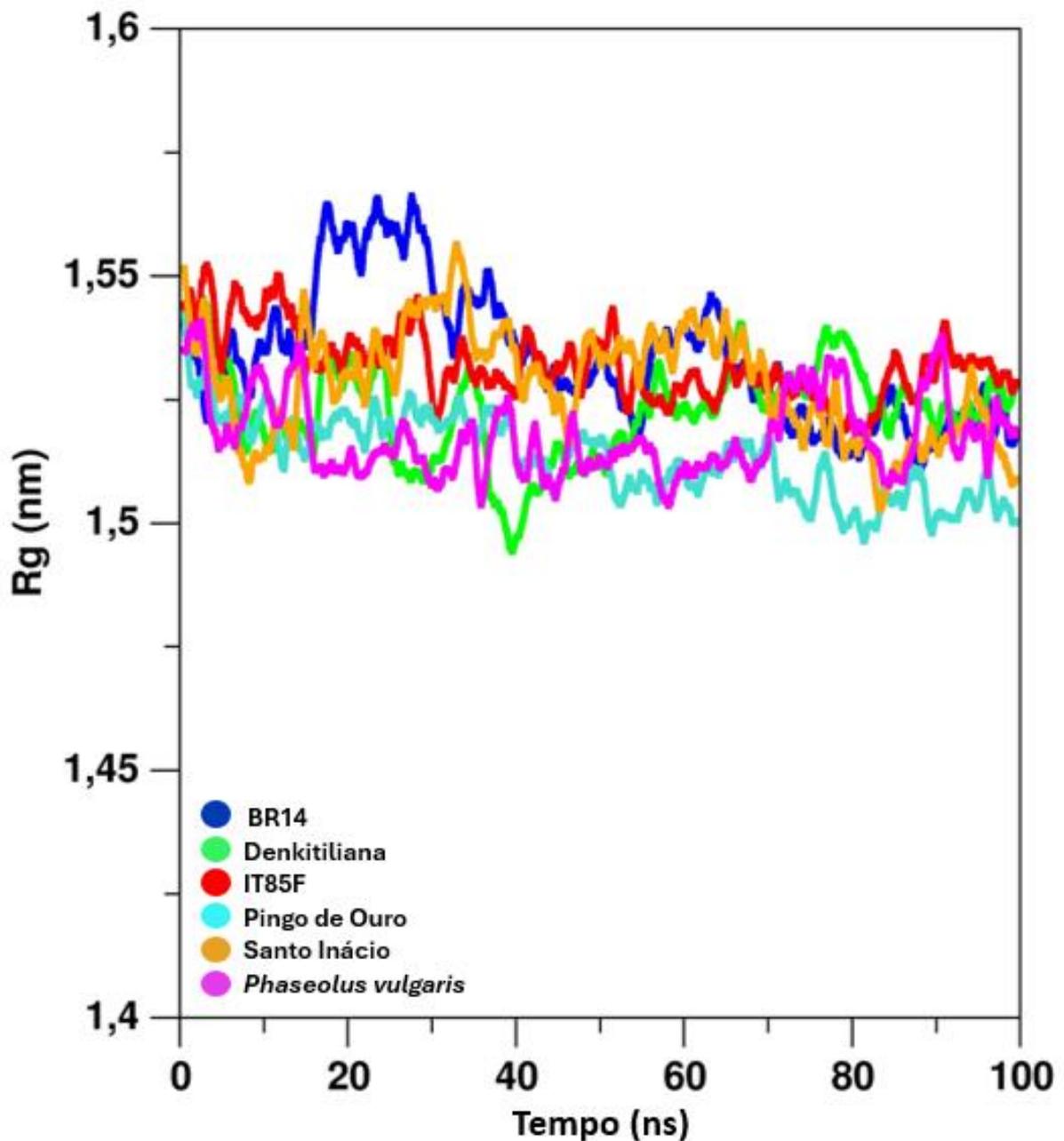
Figura 28. Gráfico de Pontes de Hidrogênio do cromossomo 7. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Para concluir, tem-se o gráfico dos Raios de Giro (RG) das isoformas do cromossomo 7 (Figura 29). Com base nele, avalia-se que todas as estruturas das

cultivares, não apresentaram grandes variações até alcançar a sua forma mais estável, de convergência estrutural. E, assim como as isoformas do cromossomo anterior, apresentam um maior grau de compactação, quando comparadas a sua estrutura anterior, com uma variação menor que 0,05 nm, validando assim, a estabilidade das estruturas, em concordância com as análises anteriores.

Figura 29. Gráfico do Raio de Giro do cromossomo 7. Representado em azul a análise da cultivar BR14, em verde a Denkitiliana, em vermelho a IT85F, em turquesa a Pingo de Ouro, em laranja a Santo Inácio e em magenta a *P. vulgaris*.

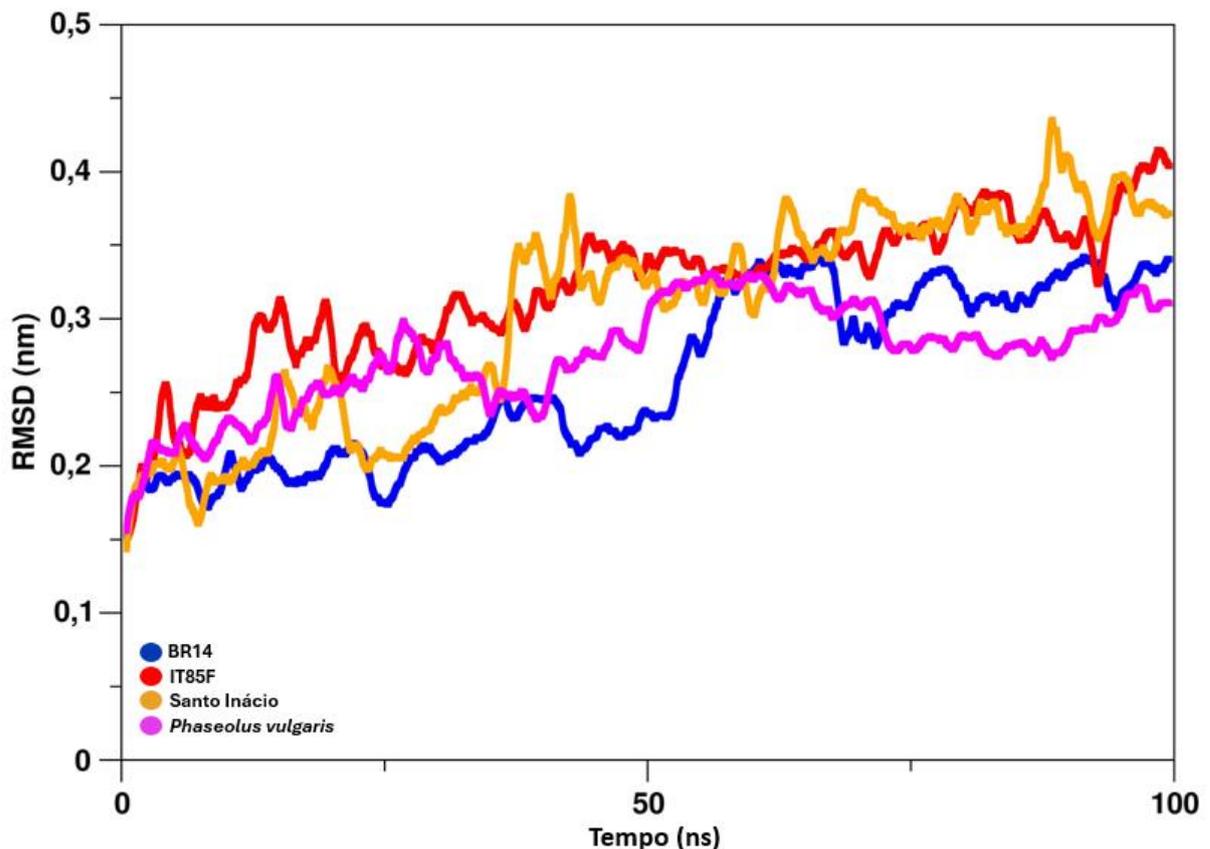


Para finalizar a etapa das dinâmicas, as isoformas mais recentes do

cromossomo 6, também foram submetidas às análises da dinâmica molecular.

Começando assim, pela avaliação do RMSD (Figura 30) das estruturas do cromossomo 6, das cultivares BR14, IT85F e Santo Inacio, pois, são as únicas que apresentam tais isoformas nessa região genômica, além da *Phaseolus vulgaris*, utilizada a fim de comparação. Baseado no gráfico, pode-se observar que as recentes isoformas, assim como as anteriores, alcançam suas respectivas convergências estruturais, principalmente após os 50 ns, com possíveis perdas de cadeias laterais, mas sem exacerbadas modificações, confirmando assim, a qualidade dos modelos teóricos preditos.

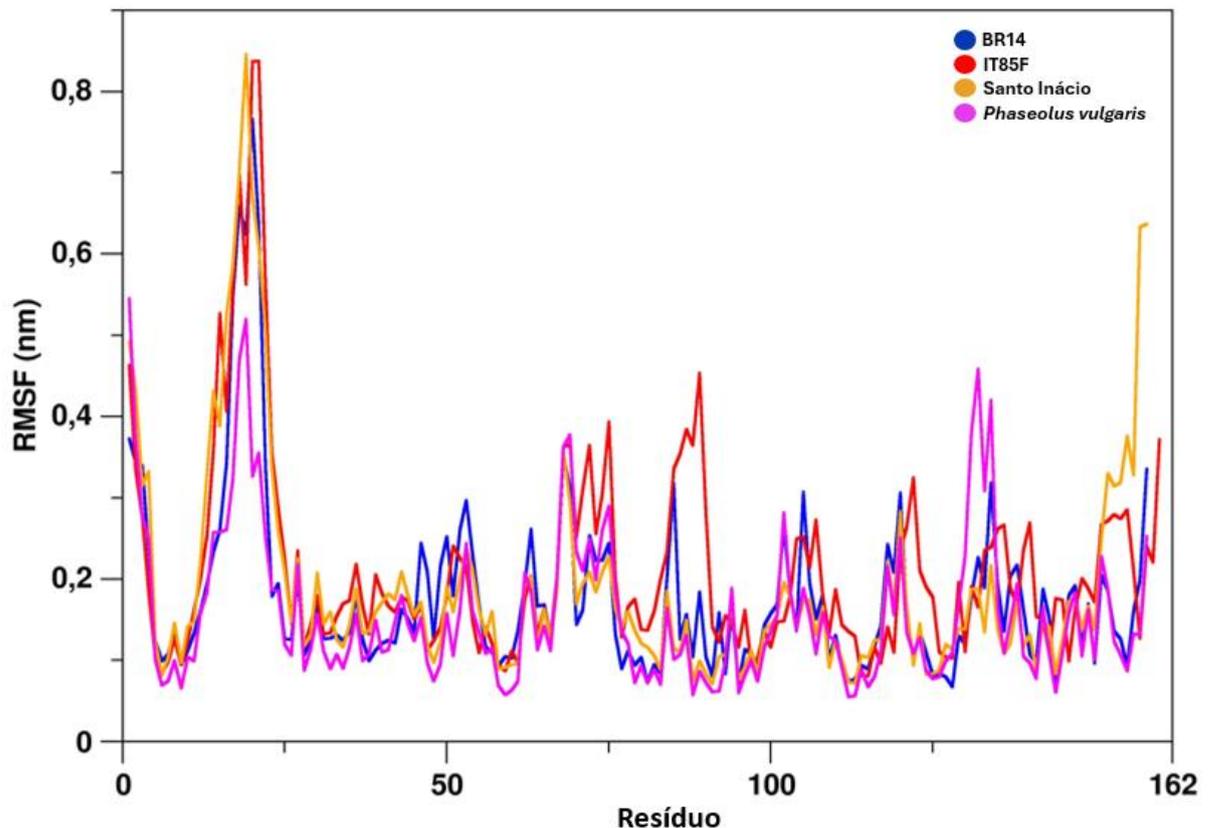
Figura 30. Gráfico de RMSD do cromossomo 6. Representado em azul a análise da cultivar BR14, em vermelho a IT85F, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Em seguida, foi realizada a dinâmica do RMSF das estruturas do cromossomo 6, como mostra o gráfico abaixo (Figura 31). Embasado nele, podemos verificar algumas características comuns entre elas, como, o triptofano conservado em todas, na posição 68, o qual, possui uma certa flexibilidade, mas levemente abaixo de 0,4 nm. De modo geral, essas isoformas são mais rígidas que as anteriores,

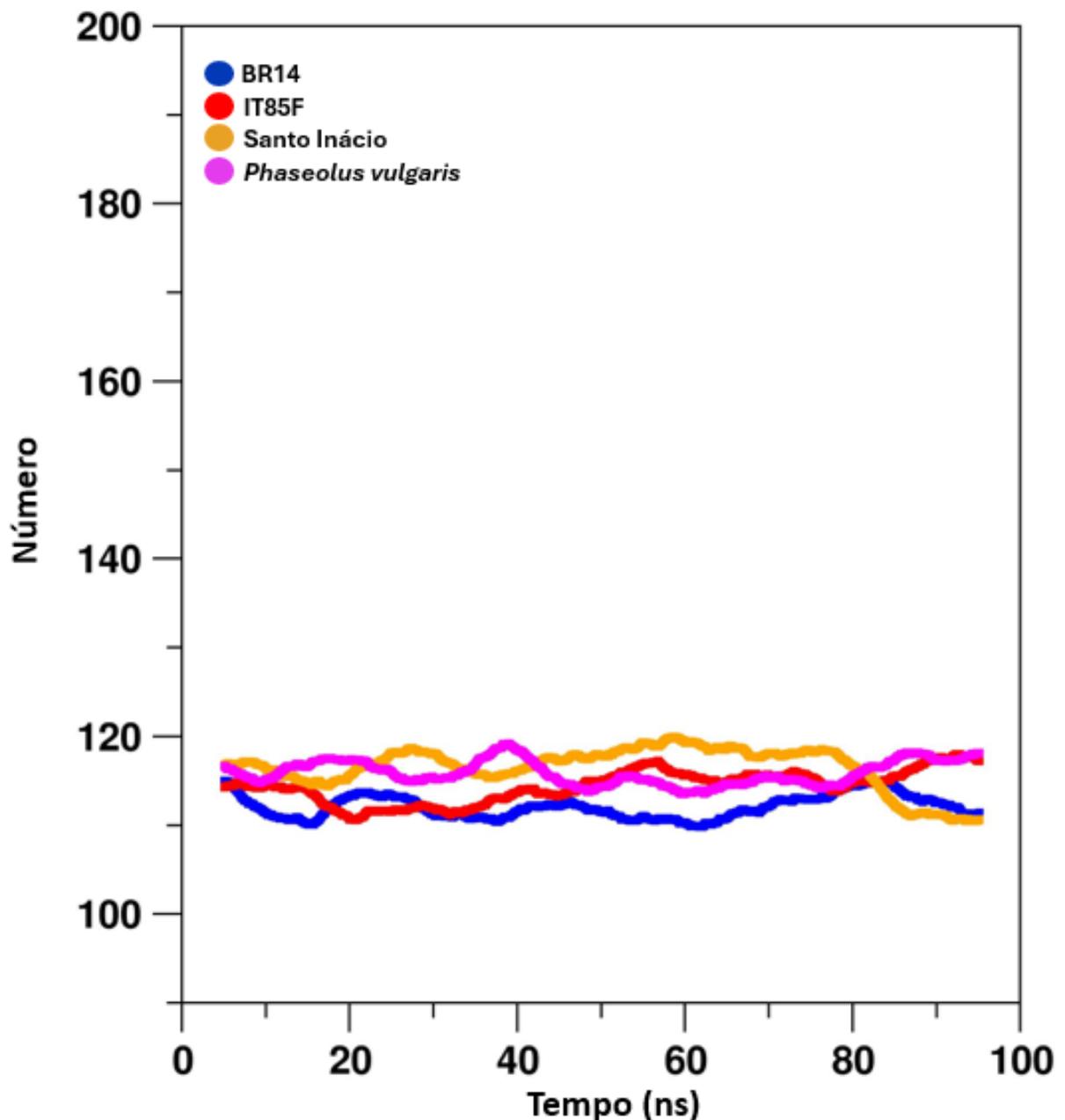
demonstrando apenas 1 região de pico de flexibilidade maior que 0,4 nm. Com exceção da IT85F, a qual pode ser considerada a cultivar mais flexível dentre essas, com uma segunda região de pico, levemente mais alta que 0,4 nm, na posição 89, ocupada por um resíduo de Lisina. Portanto, para as características individuais, a BR14 possui seus resíduos flexíveis de maior importância, nas seguintes posições: K-18, A-20 e W-68, a IT85F nas posições K-18, D-20/E-21, W-68 e K-89. Já a Santo Inácio, pode ser considerada a mais rígida dentre todas, possuindo seus pontos de maior flexibilidade nas posições Q-19, sendo esse o maior pico dentre todas desse cromossomo, e, o triptofano na posição 68. Baseado nesses resultados, as isoformas do cromossomo 6, provavelmente, não se ligam facilmente a outras moléculas, em virtude da rigidez das suas estruturas, pois, a flexibilidade das proteínas permite que elas adotem múltiplas conformações, o que é crucial para a ligação, em vista que, proteínas mais flexíveis podem se ajustar melhor à forma do ligante, facilitando o encaixe no sítio de ligação (AMARAL *et al.*, 2017).

Figura 31. Gráfico de RMSF do cromossomo 6. Representado em azul a análise da cultivar BR14, em vermelho a IT85F, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



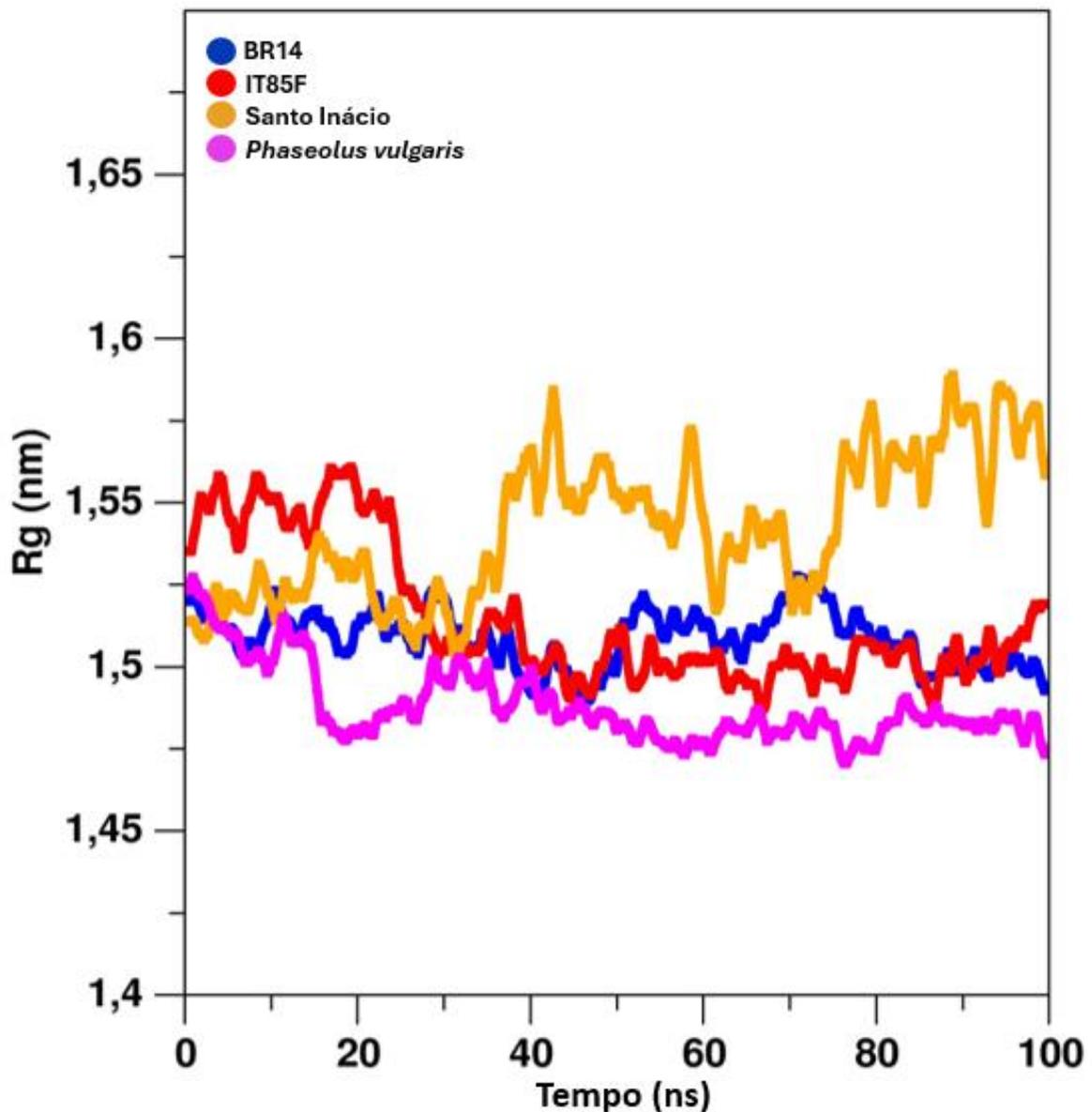
A seguir, as análises das Pontes de Hidrogênio (HB) das isoformas do cromossomo 6 (Figura 32). No geral, as isoformas mostraram uma estabilidade na manutenção dessas ligações, no decorrer dos 100 ns, com uma leve diminuição do número de HB nas cultivares BR14 e SI, e, um leve aumento na IT85F. Quando comparadas às isoformas dos cromossomos anteriores, essas demonstram mais pontes de hidrogênio que as isoformas do cromossomo 4, entretanto, menos que as isoformas do cromossomo 7.

Figura 32. Gráfico das Pontes de Hidrogênio do cromossomo 6. Representado em azul a análise da cultivar BR14, em vermelho a IT85F, em laranja a Santo Inácio e em magenta a *P. vulgaris*.



Por fim, tem-se as análises do Raio de Giro (RG) dessas isoformas mais recentes (Figura 33). Assim como as isoformas anteriores, as do cromossomo 6 não apresentaram grandes modificações estruturais nos 100 ns, demonstrando um maior nível de compactação das suas estruturas finais, em comparação com a suas estruturas iniciais. Com exceção da cultivar Santo Inácio, a qual demonstrou um comportamento de expansão da sua estrutura, com uma variação de aproximadamente 0,05 nm, em contrapartida com a sua menor flexibilidade evidenciada anteriormente.

Figura 33. Gráfico do Raio de Giro do cromossomo 6. Representado em azul a análise da cultivar BR14, em vermelho a IT85F, em laranja a Santo Inácio e em magenta a *P. vulgaris*.

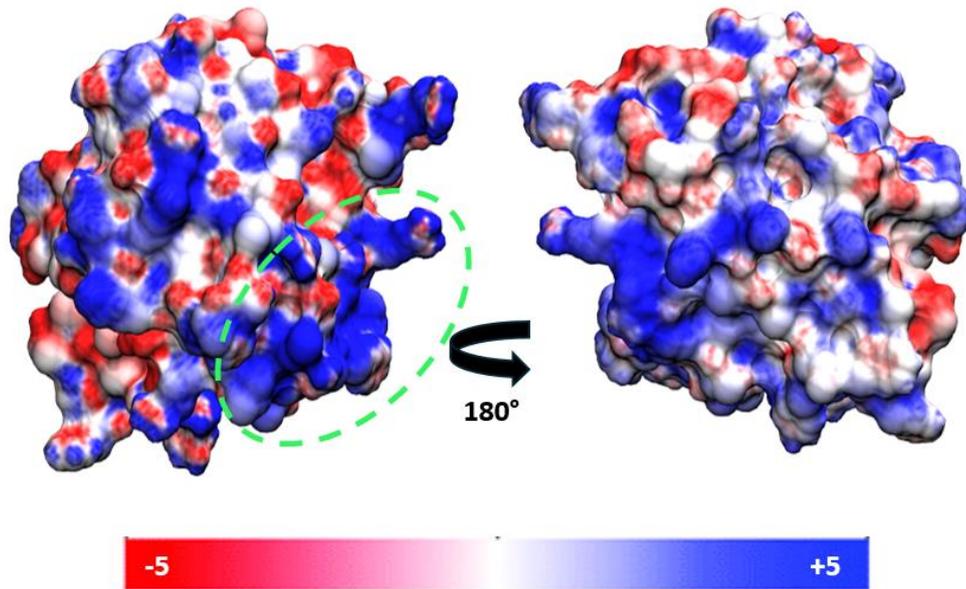


4.5 Perfil de Superfície Eletrostática

O último parâmetro observado nesse estudo, foi a análise das cargas eletrostáticas superficiais, das cultivares de *Vigna unguiculata*. Dessa forma, características individuais de cada cromossomo e, características comuns entre as isoformas de diferentes cromossomos, foram notadas. Uma dessas características globais observadas, foi um bolsão de cargas positivas no sítio de ligação com a região 5' CAP do mRNA, característica essa, comum para todas as isoformas do fator eIF4E, com exceção das isoformas do cromossomo 6, apresentadas pelas cultivares BR14, IT85F e SI, as quais, apresentaram cargas superficiais mais negativas. Entretanto, as isoformas do cromossomo 7, apesar de apresentar a região de maior positividade, tem um total de cargas superficiais negativas, maior que as do cromossomo 4. Esse bolsão de positividade, apresentado pela maioria das isoformas, é um importante fator para a ligação com o mRNA, pois, essa estrutura é carregada negativamente e, com o auxílio das estruturas ribossomais, esse agrupamento geralmente segue o potencial eletrostático positivo ao longo da superfície da proteína (STEITZ; MOORE, 2003; VARANI; NAGAI, 1998). Outra característica notada, foi o total de cargas superficiais apresentadas pelas isoformas do cromossomo 6, as quais tem uma negatividade superior as suas cargas positivas, apresentando um bolsão negativo.

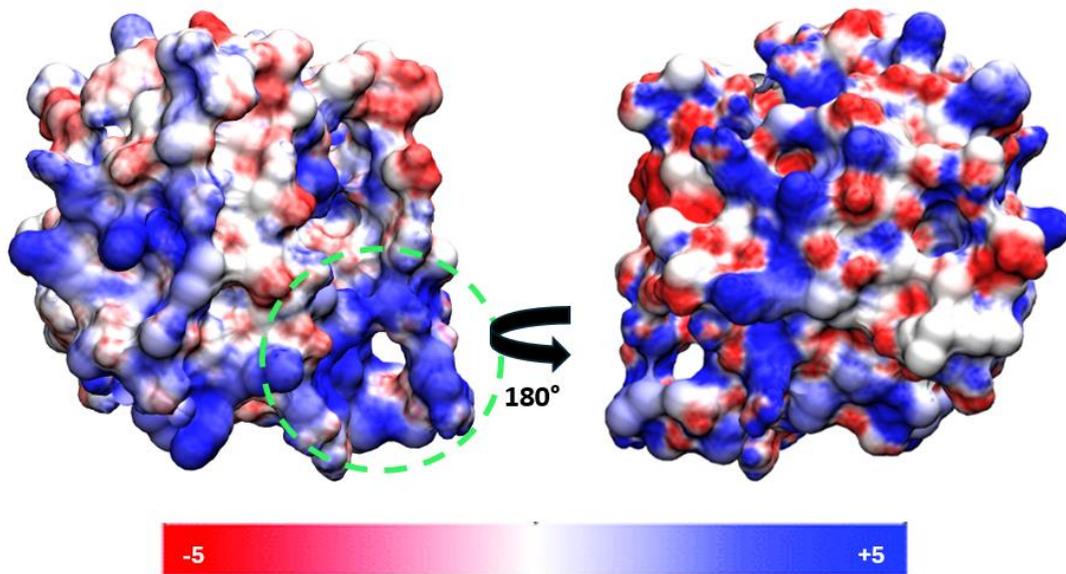
Portanto, para uma visão mais minuciosa das características de cada isoforma, por cromossomo respectivo, primeiramente as isoformas do cromossomo 4. As isoformas desse cromossomo, apresentaram características comuns nos seguintes resíduos de elevada positividade, presente nas regiões dos bolsões: R18, R83, R119, P53, P55, K79, K85, K86 e F120, além dos seu triptofanos de importância, conservados nas posições W69 e W128, que mesmo com as suas cargas mais negativas, auxiliam nas ligações com o 5'CAP. Assim, para as individualidades de cada isoforma, na estrutura do cromossomo 4 da BR14 (Figura 34), pode-se observar uma região altamente positiva, que contém, além dos resíduos comuns falados anteriormente, os resíduos R12, R13, R48, R66, R130 e P15.

Figura 34. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 4 da cultivar BR14. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



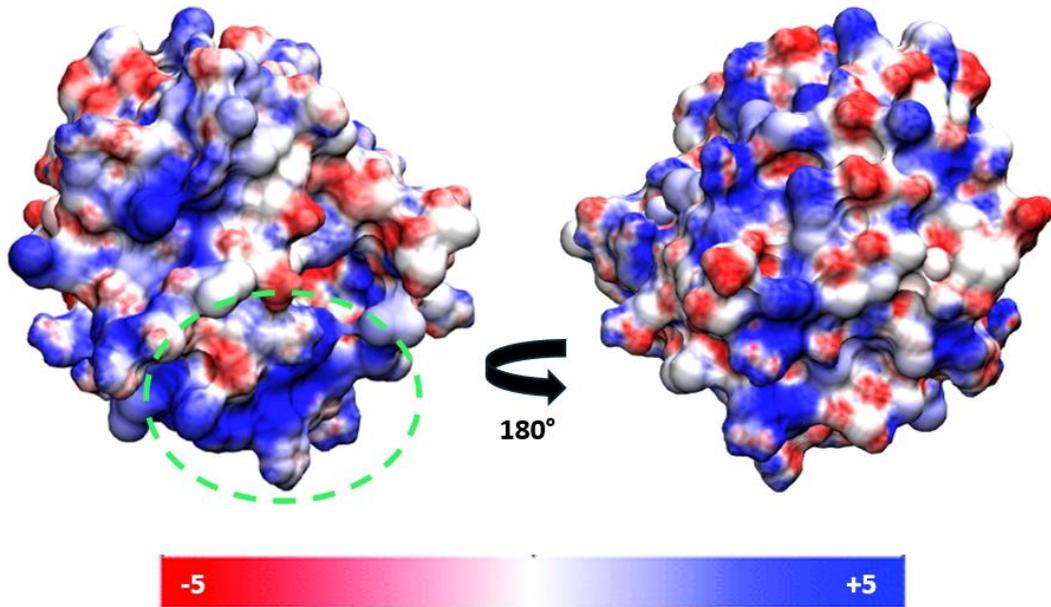
Por conseguinte, a estrutura do cromossomo 4 da cultivar Denkitiliana (Figura 35). Além dos resíduos comuns entre as cultivares, nessa região de elevada positividade, essa isoforma apresenta o resíduo R13, assim como a BR14.

Figura 35. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 4 da cultivar Denkitiliana. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



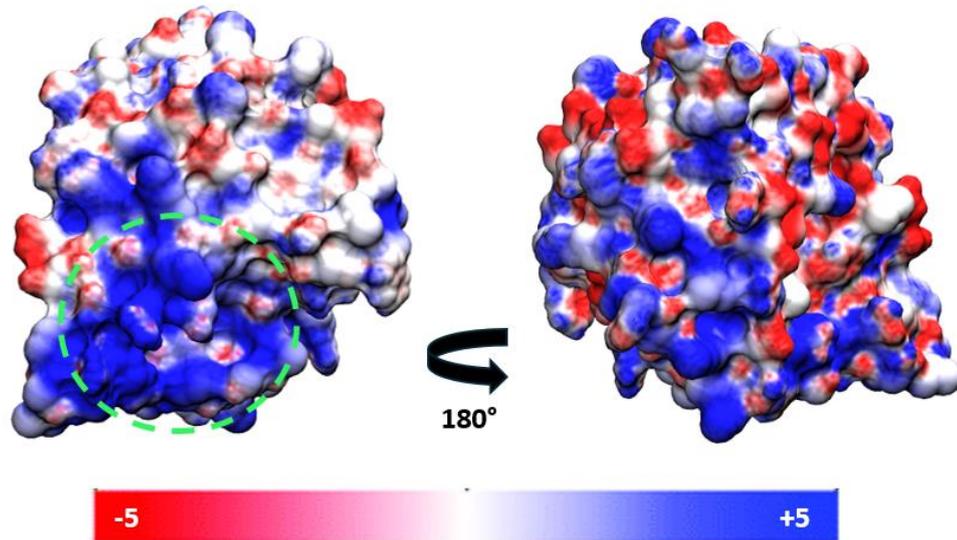
Após, a estrutura do cromossomo 4 da cultivar IT85F (Figura 36). Observa-se uma disposição menor de cargas positivas ao redor da superfície da proteína, em comparação com as cultivares anteriores, entretanto, ainda assim essa isoforma possui a sua região de maior positividade, apresentando os resíduos R66 e P15 como característica adicional dessa região.

Figura 36. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 4 da cultivar IT85F. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



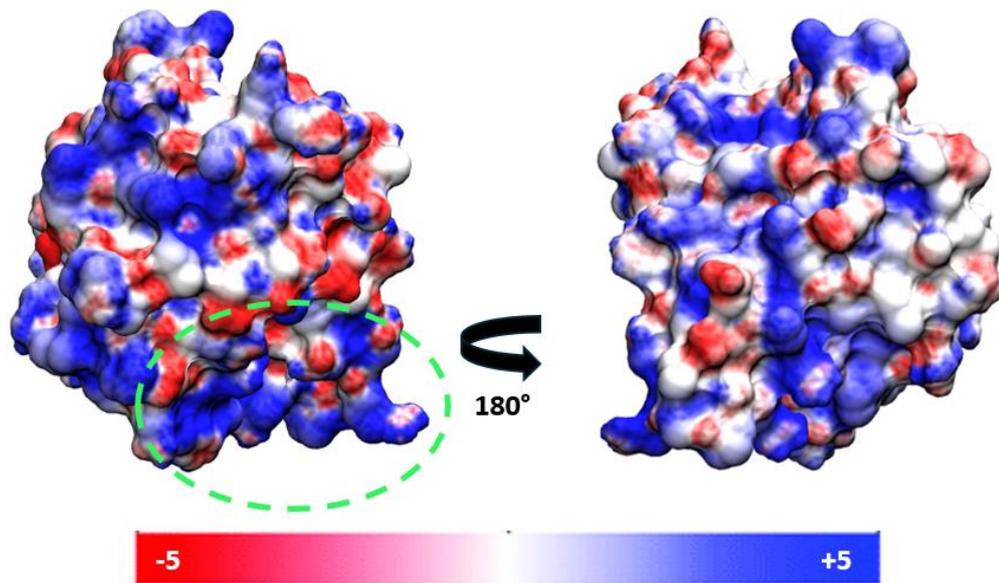
Em seguida, a disposição superficial das cargas do cromossomo 4 da cultivar Pingo de Ouro (Figura 37). Nota-se, que essa isoforma segue o padrão de disposição de cargas desse cromossomo, com uma superfície eletrostática de elevada positividade, com a sua região do bolsão possuindo os resíduos R12, R13, além dos resíduos comuns das isoformas desse cromossomo.

Figura 37. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 4 da cultivar Pingo de Ouro. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



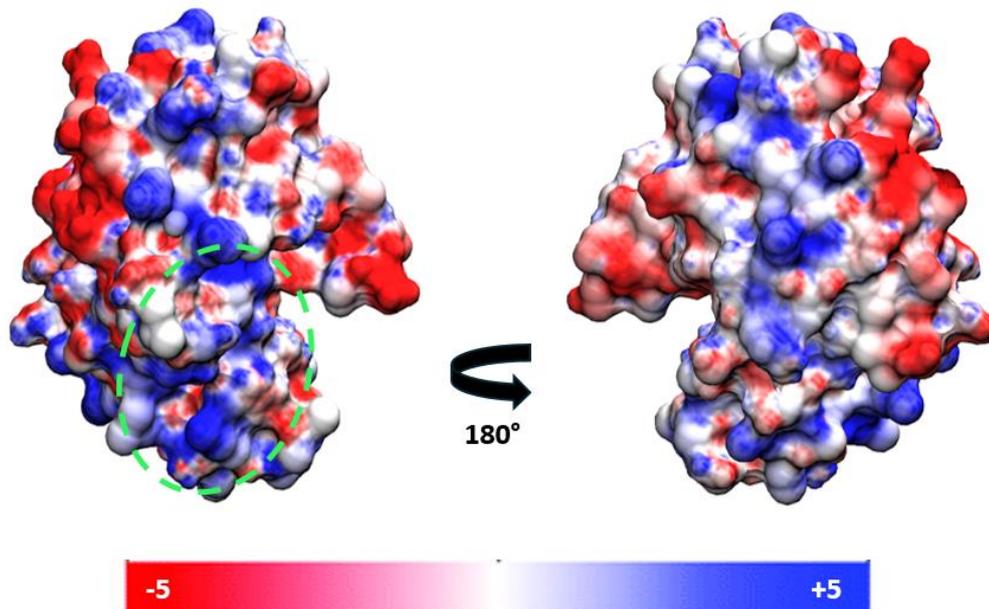
Por fim, as cargas superficiais do cromossomo 4 da cultivar Santo Inácio (Figura 38). Seguindo o mesmo padrão, com o seu bolsão nos resíduos comuns e possui R12, R13, P15 e P67, como característica a mais dessa região.

Figura 38. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 4 da cultivar Santo Inácio. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



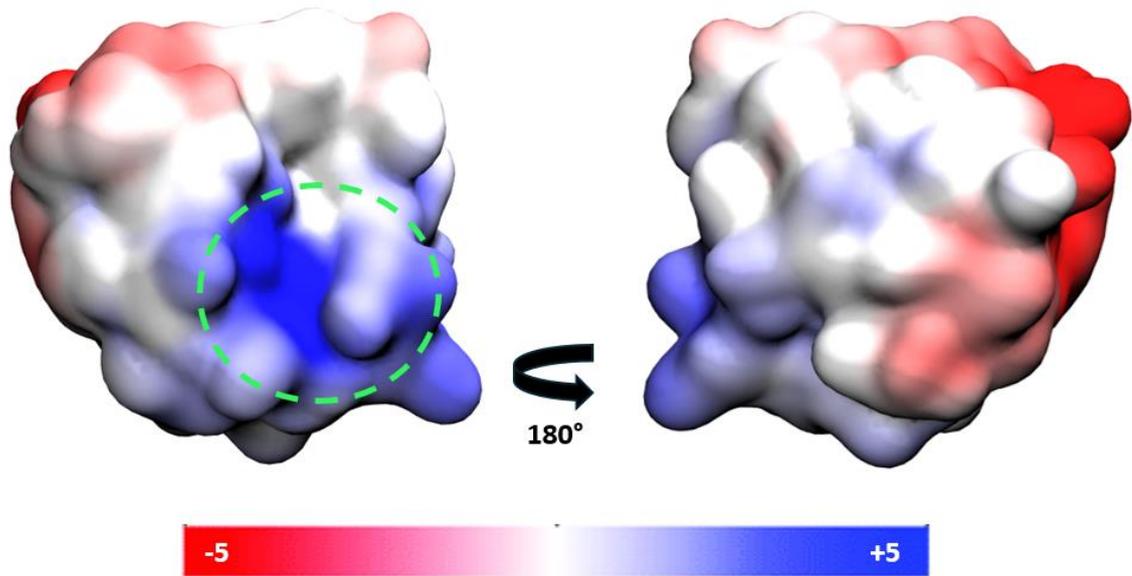
Para dar continuidade, as isoformas do cromossomo 7, as quais apresentam nos seus bolsões de positividade, resíduos comuns como: R87, R121 e K126, e os seus triptofanos de importância na posição W22. Possuem também, na posição 16, um Prolina fazendo parte dessa região, com exceção da cultivar BR14. Outra exceção nessa região de positividade, é na posição 78, onde as cultivares apresentam uma Lisina de importância para essa região, menos a cultivar Santo Inácio. Essa cultivar, compartilha de uma característica única com a cultivar Pingo de Ouro, apresentando um Triptofano de importância no bolsão de positividade, na posição 68. Então, para as características individuais, tem-se a cultivar BR14 (Figura 39). Essa estrutura, possui um volume global de cargas negativas maior que as isoformas do cromossomo 4, entretanto, possui também, embora não tão positiva quanto a isoforma do cromossomo anterior, uma região de positividade, sendo formada pelos resíduos, além dos compartilhados, comentados anteriormente, K17, P71 e um Triptofano na posição 123.

Figura 39. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 7 da cultivar BR14. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



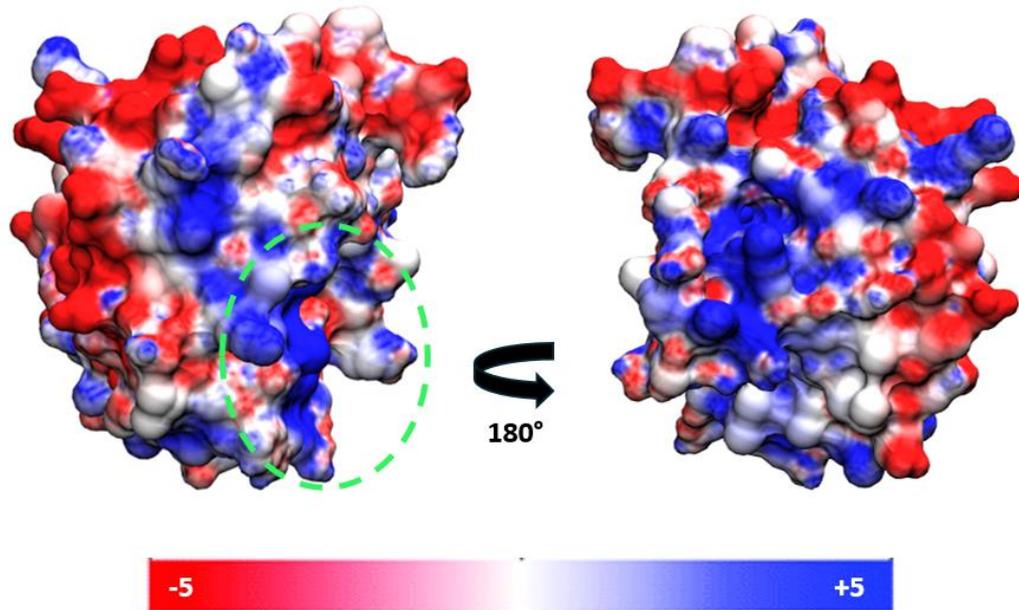
Por conseguinte, a isoforma do cromossomo 7 da cultivar Denkitiliana (Figura 40). Essa cultivar, além dos resíduos comuns, compartilha dos resíduos K17 e W123 em seu bolsão, como a BR14.

Figura 40. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 7 da cultivar Denkitiliana. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde. Por motivos técnicos essa modelagem foi feita usando o parâmetro “QuickSurf” em contrapartida das outras representações eletrostáticas que foram representadas no “Surf”.



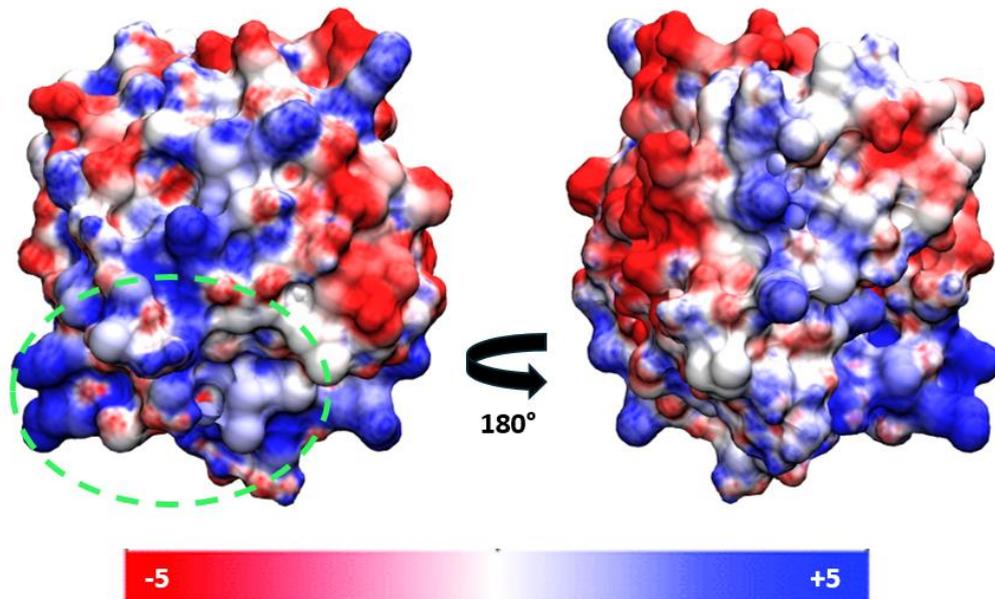
Logo após, as cargas superficiais do cromossomo 7 da cultivar IT85F (Figura 41). A qual, possui também, o W123, além de uma Lisina na posição 156 e uma Prolina na posição 66, compondo sua região de elevada positividade.

Figura 41. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 7 da cultivar IT85F. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



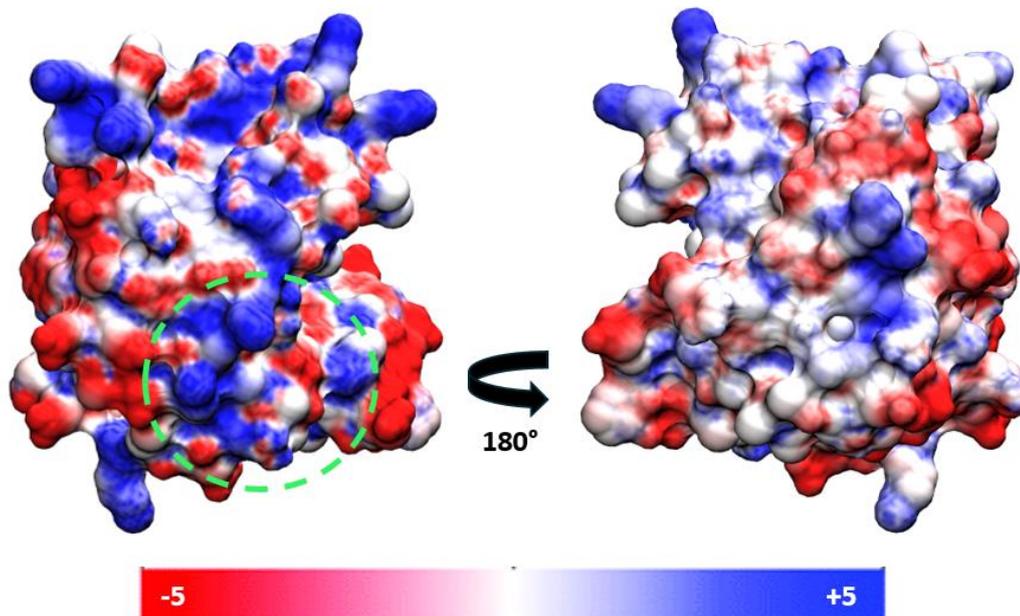
A seguir, a estrutura do cromossomo 7 da cultivar Pingo de ouro (Figura 42). Essa isoforma, como dito anteriormente, compartilha com a cultivar SI, um Triptofano da posição 68 de alta positividade, compõem também, os seguintes resíduos: K15, K17 e K156.

Figura 42. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 7 da cultivar Pingo de Ouro. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



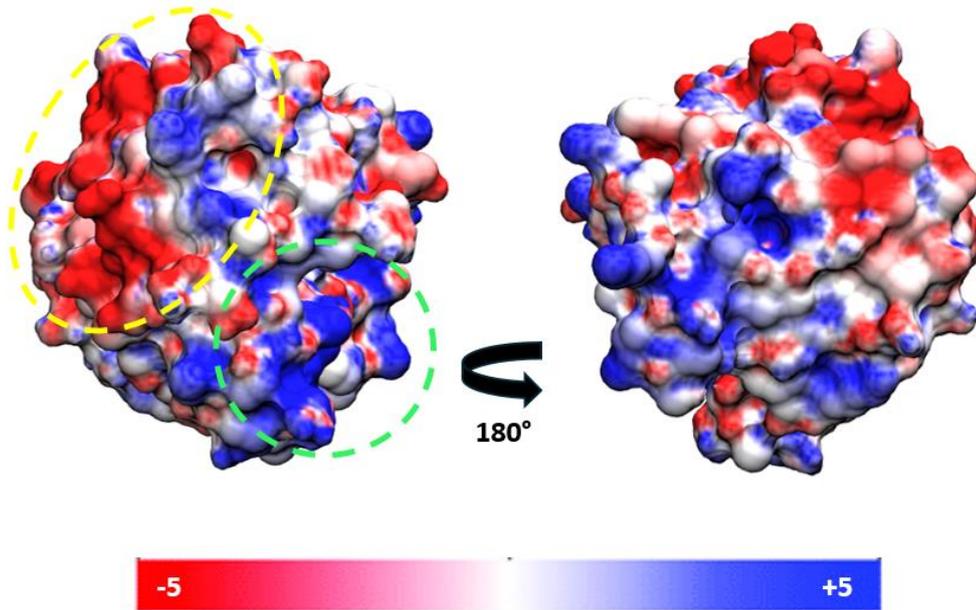
Para finalizar as estruturas desse cromossomo, a isoforma da cultivar Santo Inácio (Figura 43). Essa isoforma, além do seu W68, comporta também um Triptofano na posição 9, Prolina na posição 66 e Lisina na posição 67, como características únicas da sua região de elevada positividade.

Figura 43. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 7 da cultivar Pingo de Ouro. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde.



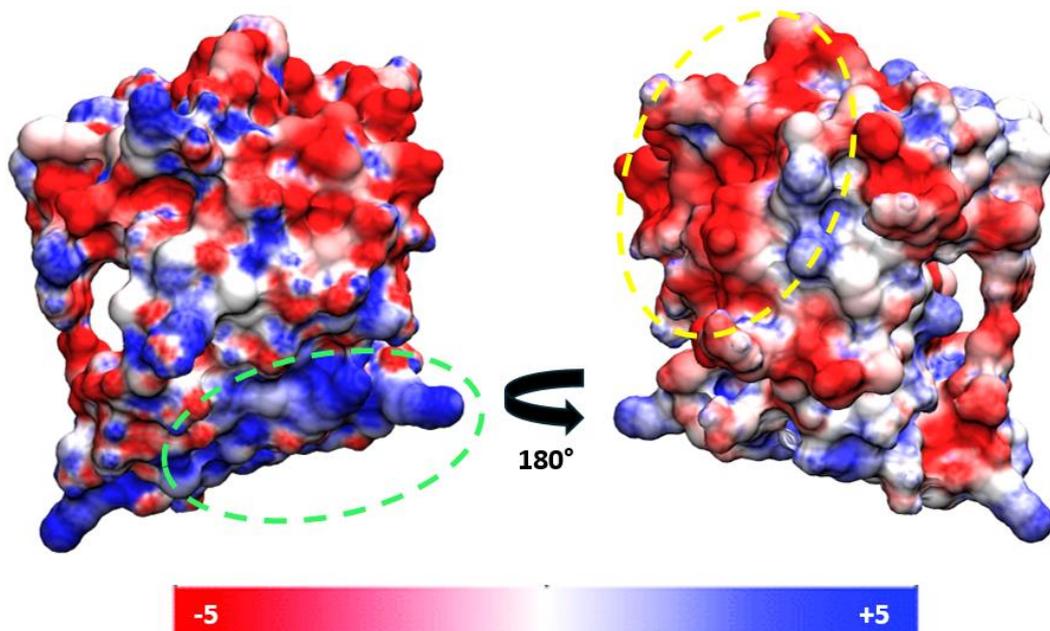
Para concluir, as isoformas do cromossomo 6, que apresentam nos seus bolsões de positividade, os resíduos comuns R120, R122, K18 e K123, e os seus triptofanos de importância nas posições W22 e W68. Iniciando com a cultivar BR14 (Figura 44), que apresenta como característica individual, os resíduos adicionais: R85 e W127 na sua região de positividade, mas, ainda possui uma distribuição de cargas negativas superficiais, elevadamente maior que as cargas positivas.

Figura 44. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 6 da cultivar BR14. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde e a região de alta negatividade no círculo amarelo



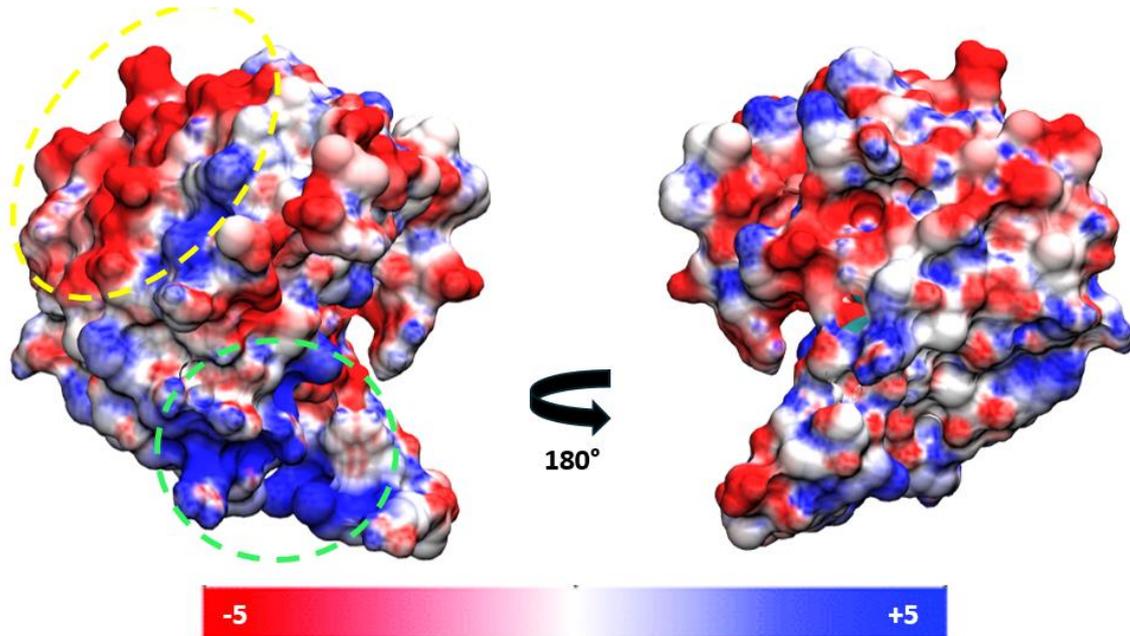
Logo após, a isoforma do cromossomo 6 da cultivar IT85F (Figura 45).

Figura 45. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 6 da cultivar IT85F. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde e a região de alta negatividade no círculo amarelo



Em conclusão, a isoforma do cromossomo 6 da cultivar Santo Inácio (Figura 46).

Figura 46. Resíduos de destaque para a diferenciação e perfil de superfície eletrostática do cromossomo 6 da cultivar Santo Inácio. Em azul os resíduos com carga positiva e em vermelho os com carga negativa. O bolsão de positividade representado no círculo verde e a região de alta negatividade no círculo amarelo



5. CONCLUSÃO

Em suma, este trabalho obteve resultados promissores sobre as isoformas de eIF4E presentes no genoma de *Vigna unguiculata*. As análises revelaram informações importantes sobre a localização cromossômica e assinaturas moleculares das eIF4E. As isoformas do cromossomo 4 mostraram uma elevada similaridade estrutural entre si, caracterizadas por bolsões de positividade e sítios de flexibilidade, sugerindo que são potenciais alvos para estudos de ligação com a região 5'CAP do mRNA, sendo as cultivares BR14, IT85F e PO, alvos promissores para estudos de susceptibilidade nessa região, e, a cultivar SI, um promissor alvo de resistência, em virtude da sua baixa flexibilidade.

No cromossomo 7, observou-se um aumento nas cargas superficiais negativas em relação ao cromossomo 4, com bolsões de positividade e flexibilidade em resíduos-chave como W68 e K15, favorecendo o reconhecimento e a ligação com o 5'CAP. Nesse cromossomo, promissores alvos de estudos para a susceptibilidade de ligação com a região 5'CAP, são as cultivares BR14, DEN e SI, com sua elevada flexibilidade nessa região de ligação, e, para estudos de resistência, a cultivar IT85F, em virtude da sua maior rigidez e distribuição de cargas negativas na superfície da proteína.

Já as 3 isoformas do cromossomo 6, apresentaram as cargas negativas mais elevadas e uma menor região de positividade, sendo mais rígidas, mas com flexibilidade conservada no triptofano W68, o que as torna promissoras para estudos de resistência à ligação com o 5'CAP ou interações com eIF4G.

Em resumo, este estudo mostrou que as isoformas do eIF4E de *V. unguiculata*, distribuídas nos cromossomos 4, 6 e 7, possuem características estruturais únicas que podem influenciar na sua capacidade de interação com componentes essenciais da tradução, como a região 5'CAP e o eIF4G. Essas diferenças estruturais podem ter implicações no desenvolvimento de resistência a patógenos e no aprimoramento da resposta ao estresse em plantas, além de fornecer uma base sólida para futuros estudos sobre a funcionalidade e adaptação evolutiva dessas proteínas. Os resultados gerados contribuem para a compreensão dos mecanismos moleculares envolvidos na resistência e na regulação da tradução em plantas, oferecendo potenciais alvos para biotecnologia e melhoramento genético.

REFERÊNCIAS

ALA-POIKELA, M.; RAJAMÄKI, M.-L.; VALKONEN, J. P. T. A Novel Interaction Network Used by Potyviruses in Virus–Host Interactions at the Protein Level. **Viruses**, 14 dez. 2019. v. 11, n. 12, p. 1158.

ALVAREZ-GARCIA, D.; BARRIL, X. Relationship between Protein Flexibility and Binding: Lessons for Structure-Based Drug Design. **Journal of Chemical Theory and Computation**, 10 jun. 2014. v. 10, n. 6, p. 2608–2614.

AMARAL, M. *et al.* Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. **Nature Communications**, 22 dez. 2017. v. 8, n. 1, p. 2276.

ANDERSEN, E. J. *et al.* Disease Resistance Mechanisms in Plants. **Genes**, 4 jul. 2018. v. 9, n. 7, p. 339.

BASHIR, M.; AHMAD, Z.; GHAFOR, A. *Cowpea aphid-borne mosaic potyvirus: A review.* **International Journal of Pest Management**, 26 jan. 2002. v. 48, n. 2, p. 155–168.

BASTET, A.; ROBAGLIA, C.; GALLOIS, J.-L. eIF4E Resistance: Natural Variation Should Guide Gene Editing. **Trends in Plant Science**, maio. 2017. v. 22, n. 5, p. 411–419.

BORGES, A. N. C. **Estresses em feijão-caupi : análise da resposta transcricional a injúria, viroses ou desidratação radicular.** Recife: Universidade Federal de Pernambuco, 2021.

BOUKAR, O. *et al.* Cowpea (*Vigna unguiculata*): Genetics, genomics and breeding. **Plant Breeding**, 9 ago. 2019. v. 138, n. 4, p. 415–424.

BRYANT, P.; ELOFSSON, A. Decomposing Structural Response Due to Sequence Changes in Protein Domains with Machine Learning. **Journal of Molecular Biology**, jul. 2020. v. 432, n. 16, p. 4435–4446.

CALLAWAY, E. 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. **Nature**, 10 dez. 2020. v. 588, n. 7837, p. 203–204.

CARVALHO, M. *et al.* Cowpea: a legume crop for a challenging environment. **Journal of the Science of Food and Agriculture**, 17 out. 2017. v. 97, n. 13, p. 4273–4284.

CHEMIM, F. Cenário brasileiro do Feijão-caupi. **Instituto Brasileiro do Feijão e Pulses (IBRAFE)**, 2022. Disponível em: <<https://www.ibrafe.org>>. Acesso em: 16 set. 2024.

COLARICCIO, A. *et al.* Characterization of *Cowpea aphid-borne mosaic virus* in *Passiflora edulis* 'Catarina' in the southern coast of Santa Catarina State, Brazil. **Arquivos do Instituto Biológico**, 2020. v. 87.

COUTINHO DE OLIVEIRA, L. *et al.* Structural studies of the eIF4E–VPg complex reveal a direct competition for capped RNA: Implications for translation. **Proceedings of the National Academy of Sciences**, 26 nov. 2019. v. 116, n. 48, p. 24056–24065.

DEAN, R. *et al.* The Top 10 fungal pathogens in molecular plant pathology. **Molecular plant pathology**, maio. 2012. v. 13, n. 4, p. 414–30.

DHANASEKAR, P.; SOUFRAMANIEN, J.; SUPRASANNA, P. Breeding Cowpea for Quality Traits: A Genetic Biofortification Perspective. **Breeding for Enhanced Nutrition and Bio-Active Compounds in Food Legumes**. Cham: Springer International Publishing, 2021, p. 157–179.

ESSEGIAN, D. J. *et al.* Molecular Dynamics Simulations Identify Tractable Lead-like Phenyl-Piperazine Scaffolds as eIF4A1 ATP-competitive Inhibitors. **ACS Omega**, 28 set. 2021. v. 6, n. 38, p. 24432–24443.

FREITAS, R. L.; TEIXEIRA, A. R.; FERREIRA, R. B. Characterization of the Proteins from *Vigna unguiculata* Seeds. **Journal of Agricultural and Food Chemistry**, 1 mar. 2004. v. 52, n. 6, p. 1682–1687.

FROTA, K. De M. G.; SOARES, R. A. M.; ARÊAS, J. A. G. Composição química do feijão caupi (*Vigna unguiculata* L. Walp), cultivar BRS-Milênio. **Ciência e Tecnologia de Alimentos**, jun. 2008. v. 28, n. 2, p. 470–476.

FRYDRYSKOVA, K. *et al.* Distinct recruitment of human eIF4E isoforms to processing bodies and stress granules. **BMC Molecular Biology**, 30 dez. 2016. v. 17, n. 1, p. 21.

GOMES, A. M. F. *et al.* Genetic Diversity among Cowpea (*Vigna unguiculata* (L.) Walp.) Landraces Suggests Central Mozambique as an Important Hotspot of Variation. **Agronomy**, 30 nov. 2020. v. 10, n. 12, p. 1893.

GONÇALVES, A. *et al.* Cowpea (*Vigna unguiculata* L. Walp), a renewed multipurpose crop for a more sustainable agri-food system: nutritional advantages and constraints. **Journal of the Science of Food and Agriculture**, jul. 2016. v. 96, n. 9, p. 2941–2951.

HARIHARAN, G.; PRASANATH, K. Recent Advances in Molecular Diagnostics of Fungal Plant Pathogens: A Mini Review. **Frontiers in Cellular and Infection Microbiology**, 11 jan. 2021. v. 10.

HERNÁNDEZ-ALVAREZ, L. *et al.* Computational study on the allosteric mechanism of *Leishmania major* IF4E-1 by 4E-interacting protein-1: Unravelling the determinants of m7GTP cap recognition. **Computational and Structural Biotechnology Journal**, 2021. v. 19, p. 2027–2044.

HERRINGTON, N. B.; KELLOGG, G. E. 3D Interaction Homology: Computational Titration of Aspartic Acid, Glutamic Acid and Histidine Can Create pH-Tunable Hydrophobic Environment Maps. **Frontiers in Molecular Biosciences**, 3 nov. 2021. v. 8.

HOLLINGSWORTH, S. A.; DROR, R. O. Molecular Dynamics Simulation for All. **Neuron**, set. 2018. v. 99, n. 6, p. 1129–1143.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual molecular dynamics. **Journal of Molecular Graphics**, fev. 1996. v. 14, n. 1, p. 33–38.

HUMPHREYS, I. R. *et al.* Computed structures of core eukaryotic protein complexes. **Science**, 10 dez. 2021. v. 374, n. 6573.

HYUNA, H. D. H. K. S. G. H. Generation of Thermal Neutron Scattering Library for Light Water Using Molecular Dynamics Simulation code, GROMACS. 2023.

JAIN, A. *et al.* A review of plant leaf fungal diseases and its environment speciation. **Bioengineered**, 1 jan. 2019. v. 10, n. 1, p. 409–424.

JANSON, G.; PAIARDINI, A. PyMod 3: a complete suite for structural bioinformatics in PyMOL. **Bioinformatics**, 16 jun. 2021. v. 37, n. 10, p. 1471–1472.

JAYATHILAKE, C. *et al.* Cowpea: an overview on its nutritional facts and health benefits. **Journal of the Science of Food and Agriculture**, 30 out. 2018. v. 98, n. 13, p. 4793–4806.

JENKINS, A. L.; LARSEN, R. A.; WILLIAMS, T. B. Characterization of amino acids using Raman spectroscopy. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, maio. 2005. v. 61, n. 7, p. 1585–1594.

JONES, J. T. *et al.* Top 10 plant-parasitic nematodes in molecular plant pathology. **Molecular Plant Pathology**, dez. 2013. v. 14, n. 9, p. 946–961.

JOSHI, B. *et al.* Phylogenetic analysis of eIF4E-family members. **BMC Evolutionary Biology**, 28 dez. 2005. v. 5, n. 1, p. 48.

JUMPER, J. *et al.* Highly accurate protein structure prediction with AlphaFold. **Nature**, 26 ago. 2021. v. 596, n. 7873, p. 583–589.

KINKELIN, K. *et al.* Crystal structure of a minimal eIF4E–Cup complex reveals a general mechanism of eIF4E regulation in translational repression. **RNA**, set. 2012. v. 18, n. 9, p. 1624–1634.

KOMAROFF, A. L. Breakthrough Discovery in Protein Structure Prediction and the Promise of New Treatments. **JAMA**, 12 out. 2021. v. 326, n. 14, p. 1369–1370.

KRYSHTAFOVYCH, A. *et al.* Critical assessment of methods of protein structure prediction (CASP)—Round <scp>XIV</scp>. **Proteins: Structure, Function, and Bioinformatics**, 7 dez. 2021. v. 89, n. 12, p. 1607–1617.

KUHLMAN, B.; BRADLEY, P. Advances in protein structure prediction and design. **Nature Reviews Molecular Cell Biology**, 15 nov. 2019. v. 20, n. 11, p. 681–697.

LASKOWSKI, R. A.; MACARTHUR, M. W.; THORNTON, J. M. *PROCHECK*: validation of protein-structure coordinates. [S.l.]: [s.n.], 2012, p. 684–687.

LAYANA, C. *et al.* *Drosophila* Me31B is a Dual eIF4E-Interacting Protein. **Journal of Molecular Biology**, mar. 2023. v. 435, n. 5, p. 167949.

LAZIM, R.; SUH, D.; CHOI, S. Advances in Molecular Dynamics Simulations and Enhanced Sampling Methods for the Study of Protein Systems. **International Journal of Molecular Sciences**, 1 set. 2020. v. 21, n. 17, p. 6339.

LEFEUVRE, P. *et al.* Evolution and ecology of plant viruses. **Nature Reviews Microbiology**, 16 out. 2019. v. 17, n. 10, p. 632–644.

LEITE, Rubson Da Costa; LEITE, Robson Da Costa; CRAVO, M. Da S. Second-crop cowpea under residual phosphorus doses in the Brazilian Amazon. **Pesquisa Agropecuária Tropical**, 2020. v. 50.

LIU, P. *et al.* Lubricant shear thinning behavior correlated with variation of radius of gyration via molecular dynamics simulations. **The Journal of chemical physics**, 28 ago. 2017. v. 147, n. 8, p. 084904.

LU, S. *et al.* CDD/SPARCLE: the conserved domain database in 2020. **Nucleic Acids Research**, 8 jan. 2020. v. 48, n. D1, p. D265–D268.

MARGREITTER, C.; OOSTENBRINK, C. MDplot: Visualise Molecular Dynamics. **The R journal**, 10 maio. 2017. v. 9, n. 1, p. 164–186.

MARTÍNEZ-FORTÚN, J.; PHILLIPS, D. W.; JONES, H. D. Natural and artificial sources of genetic variation used in crop breeding: A baseline comparator for genome editing. **Frontiers in Genome Editing**, 22 ago. 2022. v. 4.

MAYBERRY, L. K. *et al.* Evidence for Variation in the Optimal Translation Initiation Complex: Plant eIF4B, eIF4F, and eIF(iso)4F Differentially Promote Translation of mRNAs. **Plant Physiology**, ago. 2009. v. 150, n. 4, p. 1844–1854.

MEHETRE, G. T. *et al.* Current Developments and Challenges in Plant Viral Diagnostics: A Systematic Review. **Viruses**, 5 mar. 2021. v. 13, n. 3, p. 412.

MINISTÉRIO DA AGRICULTURA (MAPA). Perspectiva aponta para novo aumento na área de arroz e feijão na safra 2024/2025. **Ministério da Agricultura, Pecuária e Abastecimento (MAPA)**, 17 set. 2024. Disponível em: <<https://www.conab.gov.br/ultimas-noticias/5735-perspectiva-aponta-para-novo-aumento-na-area-de-arroz-e-feijao-na-safra-2024-2025>>. Acesso em: 16 set. 2024.

MIRAS, M. *et al.* Structure of eIF4E in Complex with an eIF4G Peptide Supports a Universal Bipartite Binding Mode for Protein Translation. **Plant Physiology**, jul. 2017. v. 174, n. 3, p. 1476–1491.

MIRDITA, M. *et al.* ColabFold: making protein folding accessible to all. **Nature methods**, jun. 2022. v. 19, n. 6, p. 679–682.

MONTERO, H.; GARCÍA-ROMÁN, R.; MORA, S. eIF4E as a Control Target for Viruses. **Viruses**, 16 fev. 2015. v. 7, n. 2, p. 739–750.

_____; PÉREZ-GIL, G.; SAMPIERI, C. L. Eukaryotic initiation factor 4A (eIF4A) during viral infections. **Virus Genes**, 22 jun. 2019. v. 55, n. 3, p. 267–273.

NAZAROV, P. A. *et al.* Infectious plant diseases: etiology, current status, problems and prospects in plant protection. **Acta Naturae**, 27 out. 2020. v. 12, n. 3, p. 46–59.

OANCA, G.; ENT, F. VAN DER; ÅQVIST, J. Efficient Empirical Valence Bond Simulations with GROMACS. **Journal of Chemical Theory and Computation**, 12 set. 2023. v. 19, n. 17, p. 6037–6045.

OBERER, M.; MARINTCHEV, A.; WAGNER, G. Structural basis for the enhancement of eIF4A helicase activity by eIF4G. **Genes & Development**, 15 set. 2005. v. 19, n. 18, p. 2212–2223.

OOSTENBRINK, C. *et al.* Validation of the 53A6 GROMOS force field. **European Biophysics Journal**, 1 jun. 2005. v. 34, n. 4, p. 273–284.

PAN, A. C. *et al.* Atomic-level characterization of protein–protein association. **Proceedings of the National Academy of Sciences**, 5 mar. 2019. v. 116, n. 10, p. 4244–4249.

PAPADOPOULOS, E. *et al.* Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. **Proceedings of the National Academy of Sciences**, 5 ago. 2014. v. 111, n. 31.

PATRICK, R. M.; BROWNING, K. S. The eIF4F and eIFiso4F Complexes of Plants: An Evolutionary Perspective. **Comparative and Functional Genomics**, 2012. v. 2012, p. 1–12.

PFEILMEIER, S.; CALY, D. L.; MALONE, J. G. Bacterial pathogenesis of plants: future challenges from a microbial perspective. **Molecular Plant Pathology**, 4 out. 2016. v. 17, n. 8, p. 1298–1313.

POVEDA, J.; ABRIL-URIAS, P.; ESCOBAR, C. Biological Control of Plant-Parasitic Nematodes by Filamentous Fungi Inducers of Resistance: *Trichoderma*, *Mycorrhizal* and *Endophytic Fungi*. **Frontiers in Microbiology**, 25 maio. 2020. v. 11.

ROMAGNOLI, A. *et al.* Control of the eIF4E activity: structural insights and pharmacological implications. **Cellular and Molecular Life Sciences**, 19 nov. 2021. v. 78, n. 21–22, p. 6869–6885.

ROSS-KASCHITZA, D.; ALTMANN, M. eIF4E and Interactors from Unicellular Eukaryotes. **International Journal of Molecular Sciences**, 21 mar. 2020. v. 21, n. 6, p. 2170.

RUBIO, L.; GALIPIENSO, L.; FERRIOL, I. Detection of Plant Viruses and Disease Management: Relevance of Genetic Diversity and Evolution. **Frontiers in Plant Science**, 17 jul. 2020. v. 11.

Safety Assessment of Foods and Feeds Derived from Transgenic Crops, Volume 3. [S.I.]: OECD, 2019.

SAHA, S.; MÄKINEN, K. Insights into the Functions of eIF4E-Binding Motif of VPg in Potato Virus A Infection. **Viruses**, 11 fev. 2020. v. 12, n. 2, p. 197.

SANFAÇON, H. Plant Translation Factors and Virus Resistance. **Viruses**, 24 jun. 2015. v. 7, n. 7, p. 3392–3419.

SANT'ANNA, C. M. R. Molecular modeling methods in the study and design of bioactive compounds: An introduction. **Revista Virtual de Química**, 2009. v. 1, n. 1.

SANTOS-SILVA, C. A. Dos *et al.* Plant Antimicrobial Peptides: State of the Art, In Silico Prediction and Perspectives in the Omics Era. **Bioinformatics and Biology Insights**, 2 jan. 2020. v. 14, p. 117793222095273.

SIDDIQUE, S.; GRUNDLER, F. M. Parasitic nematodes manipulate plant development to establish feeding sites. **Current Opinion in Microbiology**, dez. 2018. v. 46, p. 102–108.

SIEVERS, F.; HIGGINS, D. G. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. [S.l.]: [s.n.], 2014, p. 105–116.

SILVA, J. A. *et al.* Inheritance of cowpea resistance to *Cowpea aphid-borne mosaic virus*. **Pesquisa Agropecuária Brasileira**, 2021. v. 56.

SINDHU, M. *et al.* Current advances and future directions in genetic enhancement of a climate resilient food legume crop, cowpea (*Vigna unguiculata L. Walp.*). **Plant Cell, Tissue and Organ Culture (PCTOC)**, 23 dez. 2019. v. 139, n. 3, p. 429–453.

STACHELSKA, A. *et al.* Interaction of three *Caenorhabditis elegans* isoforms of translation initiation factor eIF4E with mono- and trimethylated mRNA 5' cap analogues. **Acta biochimica Polonica**, 2002. v. 49, n. 3, p. 671–82.

STEITZ, T. A.; MOORE, P. B. RNA, the first macromolecular catalyst: the ribosome is a ribozyme. **Trends in Biochemical Sciences**, ago. 2003. v. 28, n. 8, p. 411–418.

STUDER, G. *et al.* QMEANDisCo—distance constraints applied on model quality estimation. **Bioinformatics**, 1 mar. 2020. v. 36, n. 6, p. 1765–1771.

TARASOVA, E.; NERUKH, D. All-Atom Molecular Dynamics Simulations of Whole Viruses. **The Journal of Physical Chemistry Letters**, 4 out. 2018. v. 9, n. 19, p. 5805–5809.

THANGADURAI, D. CHEMICAL COMPOSITION AND NUTRITIONAL POTENTIAL OF VIGNA UNGUICULATA SSP. CYLINDRICA (FABACEAE). **Journal of Food Biochemistry**, fev. 2005. v. 29, n. 1, p. 88–98.

THORP, H. H. Proteins, proteins everywhere. **Science**, 17 dez. 2021. v. 374, n. 6574, p. 1415–1415.

TIAN LU. Plotting electrostatic potential colored molecular surface map with ESP surface extrema via Multiwfn and VMD. 2014.

TUNYASUVUNAKOOL, K. *et al.* Highly accurate protein structure prediction for the human proteome. **Nature**, 26 ago. 2021. v. 596, n. 7873, p. 590–596.

TUPPERWAR, N. *et al.* Characterization of an Atypical eIF4E Ortholog in Leishmania, LeishIF4E-6. **International Journal of Molecular Sciences**, 24 nov. 2021. v. 22, n. 23, p. 12720.

VALE MARTINS, L. DO *et al.* BAC- and oligo-FISH mapping reveals chromosome evolution among *Vigna angularis*, *V. unguiculata*, and *Phaseolus vulgaris*. **Chromosoma**, 28 set. 2021. v. 130, n. 2–3, p. 133–147.

VARANI, G.; NAGAI, K. RNA RECOGNITION BY RNP PROTEINS DURING RNA PROCESSING. **Annual Review of Biophysics and Biomolecular Structure**, jun. 1998. v. 27, n. 1, p. 407–445.

VIDAVER, A. K.; LAMBRECHT, P. A. Bacteria as plant pathogens. **The Plant Health Instructor**, 2004.

VOLPON, L. *et al.* Cap-free structure of eIF4E suggests a basis for conformational regulation by its ligands. **The EMBO Journal**, 1 nov. 2006. v. 25, n. 21, p. 5138–5149.

WADHWA, R. *et al.* Molecular dynamics simulations and experimental studies reveal differential permeability of withaferin-A and withanone across the model cell membrane. **Scientific Reports**, 27 jan. 2021. v. 11, n. 1, p. 2352.

WALTER, J. *et al.* Spectroscopic Investigation of the Kinetic Mechanism Involved in the Association of Potyviral VPg with the Host Plant Translation Initiation Factor eIF4E. **International Journal of Molecular Sciences**, 5 ago. 2020. v. 21, n. 16, p. 5618.

WATERHOUSE, A. M. *et al.* Jalview Version 2—a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, 1 maio. 2009. v. 25, n. 9, p. 1189–1191.

WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, 8 maio. 2007. v. 35, n. Web Server, p. W407–W410.

XU, M. *et al.* Translation Initiation Factor eIF4E and eIFiso4E Are Both Required for Peanut stripe virus Infection in Peanut (*Arachis hypogaea* L.). **Frontiers in Microbiology**, 10 mar. 2017. v. 8.

YU, B. *et al.* De novo determination of near-surface electrostatic potentials by NMR. **Proceedings of the National Academy of Sciences**, 22 jun. 2021. v. 118, n. 25.

YUAN, S.; CHAN, H. C. S.; HU, Z. Using <sc>PyMOL</sc> as a platform for computational drug design. **WIREs Computational Molecular Science**, 5 mar. 2017. v. 7, n. 2.

ZLOBIN, N.; TARANOV, V. Plant eIF4E isoforms as factors of susceptibility and resistance to potyviruses. **Frontiers in Plant Science**, 10 fev. 2023a. v. 14.

_____; _____. Plant eIF4E isoforms as factors of susceptibility and resistance to potyviruses. **Frontiers in Plant Science**, 10 fev. 2023b. v. 14.

APÊNDICES

Figura Suplementar 1. Validação do modelo do cromossomo 4 da cultivar BR14. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

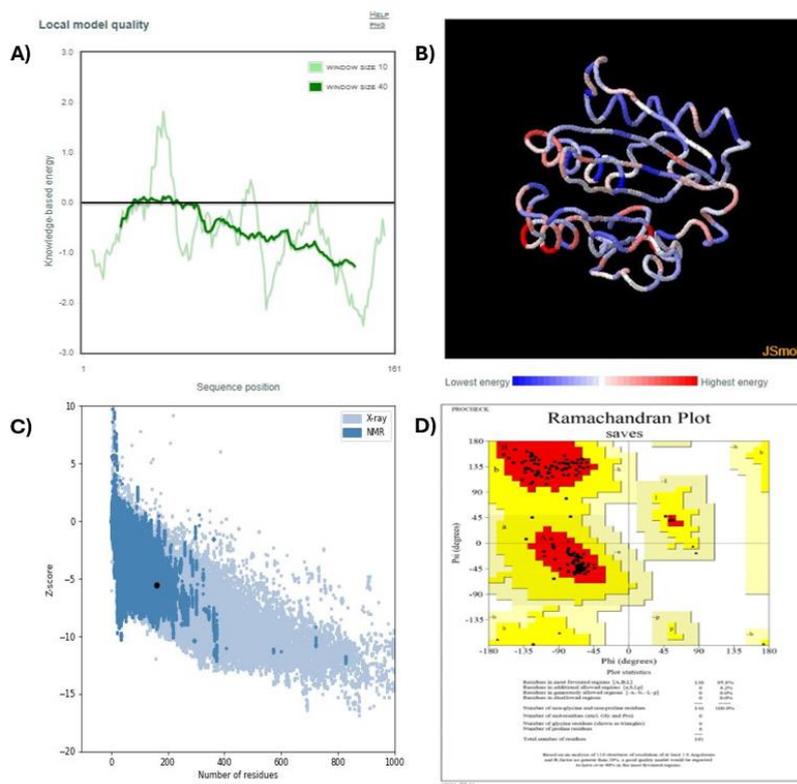


Figura Suplementar 2. Validação do modelo do cromossomo 4 da cultivar DEN. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

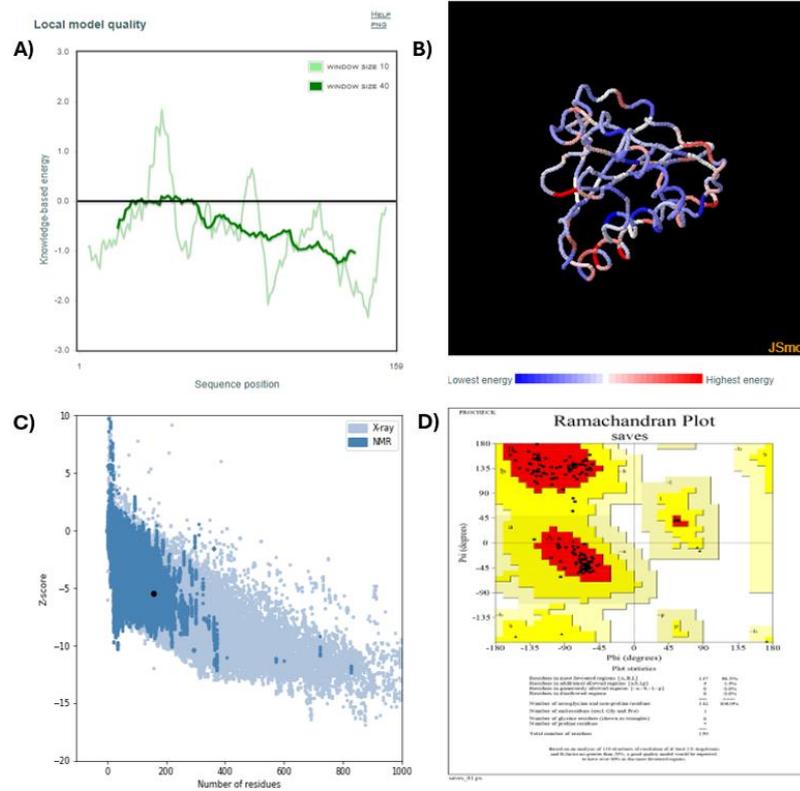


Figura Suplementar 3. Validação do modelo do cromossomo 4 da cultivar IT85F. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

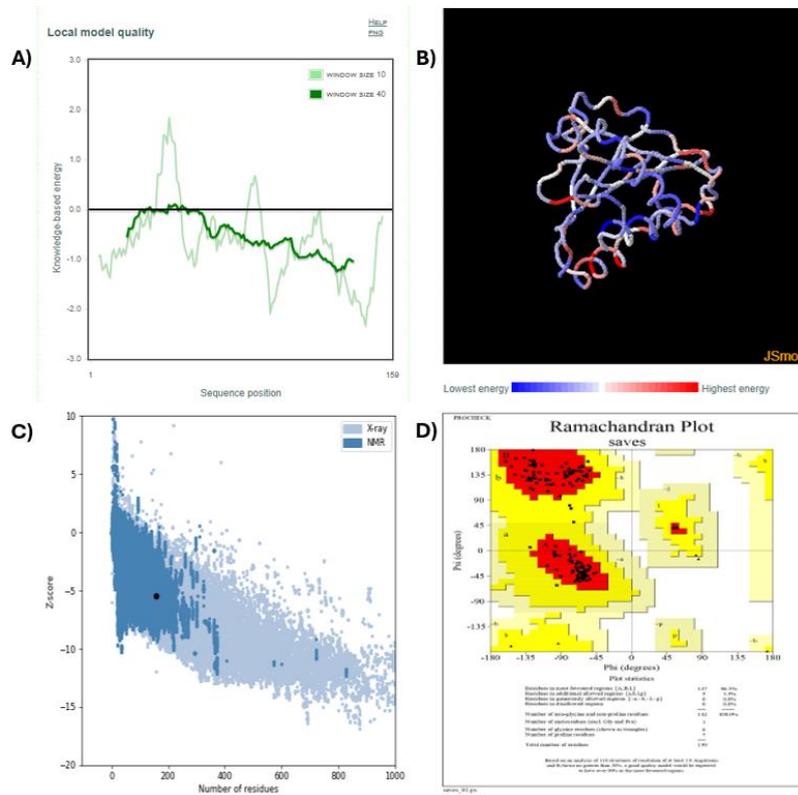


Figura Suplementar 5. Validação do modelo do cromossomo 4 da cultivar SI. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

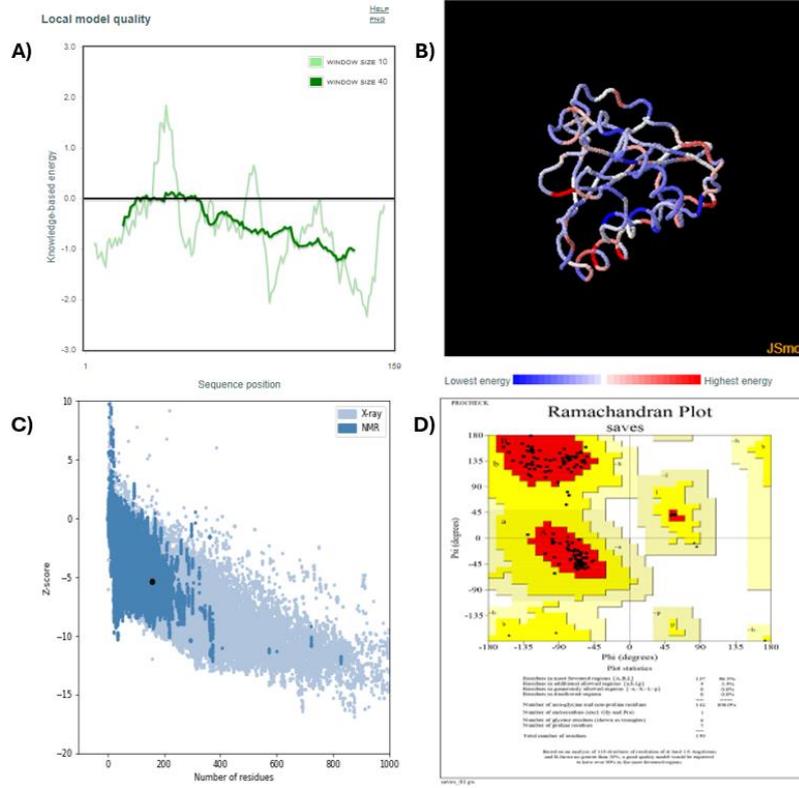


Figura Suplementar 6. Validação do modelo do cromossomo 7 da cultivar BR14. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

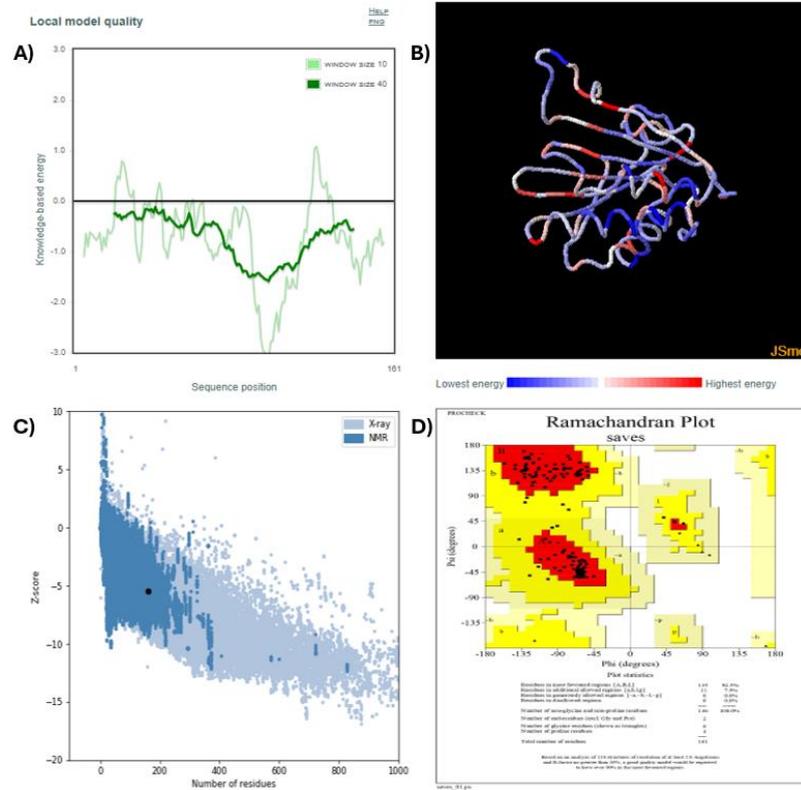


Figura Suplementar 7. Validação do modelo do cromossomo 7 da cultivar DEN. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

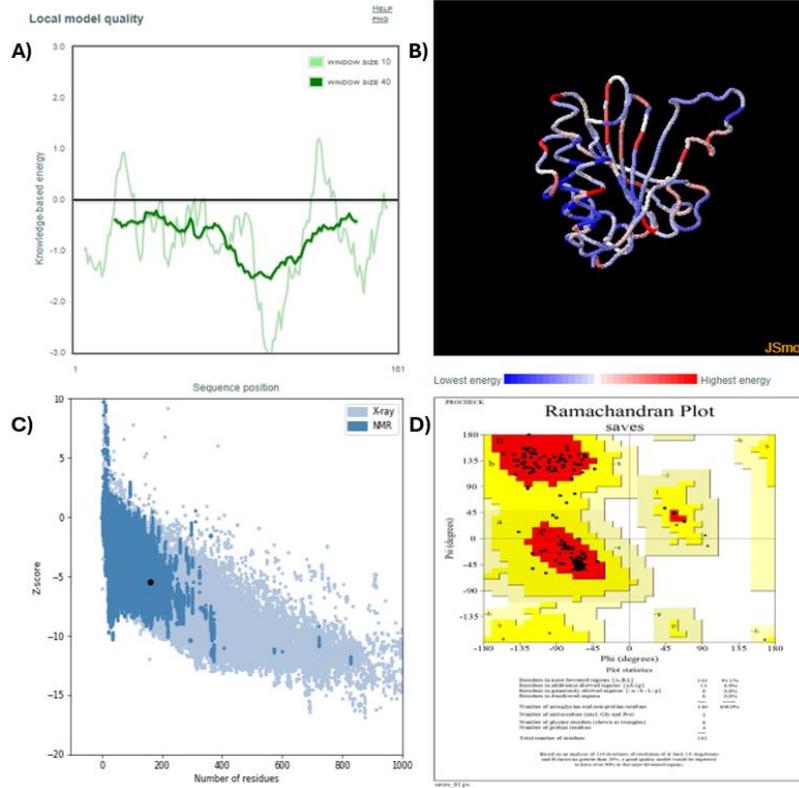


Figura Suplementar 8. Validação do modelo do cromossomo 7 da cultivar IT85F. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

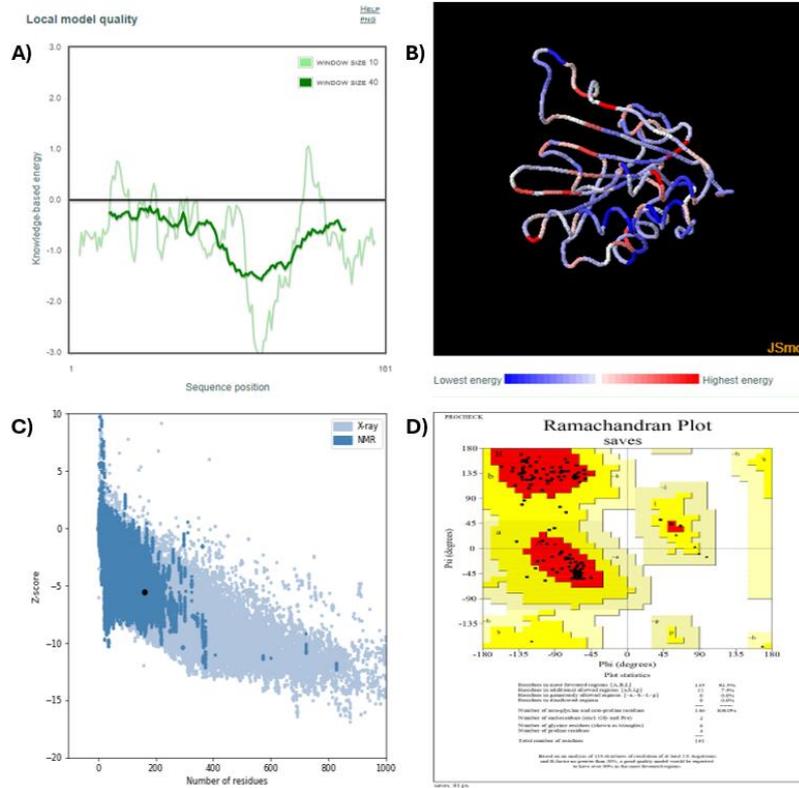


Figura Suplementar 10. Validação do modelo do cromossomo 7 da cultivar SI. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

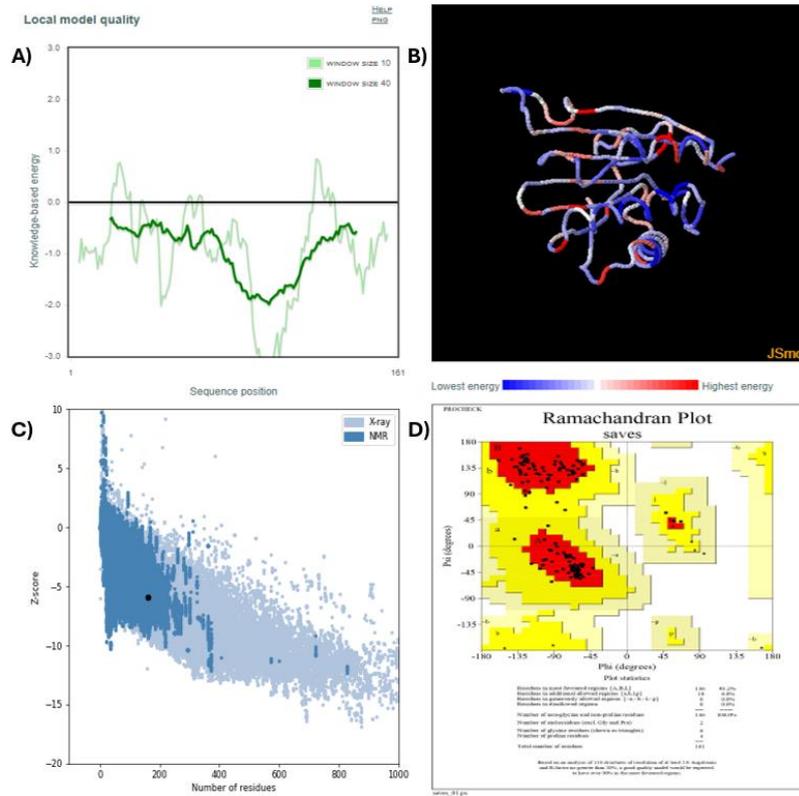


Figura Suplementar 11. Validação do modelo do cromossomo 6 da cultivar BR14. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

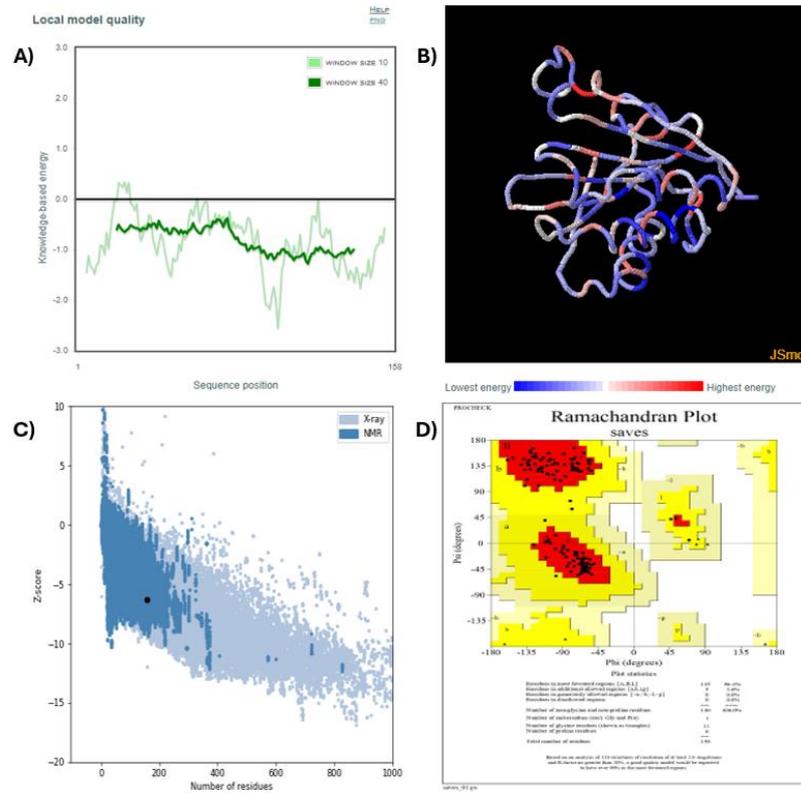


Figura Suplementar 12. Validação do modelo do cromossomo 6 da cultivar IT85F. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

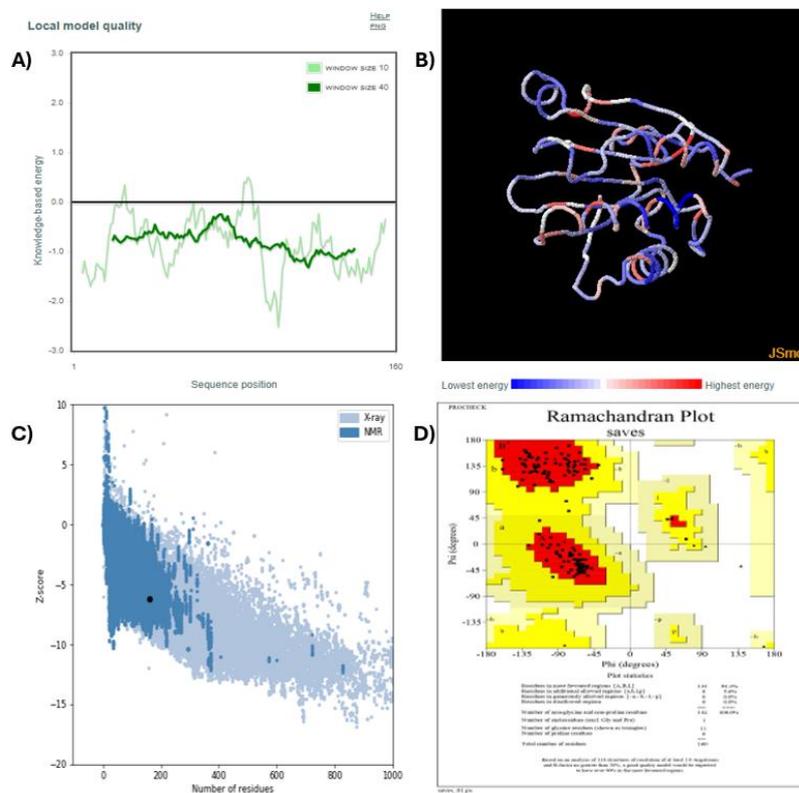


Figura Suplementar 13. Validação do modelo do cromossomo 6 da cultivar SI. A) Gráfico de cobertura de sequência B) Representação tridimensional do melhor modelo C) Z-score do melhor modelo, D) Gráfico de Ramachandran do melhor modelo.

