



**UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
CENTRO DE BIOCÊNCIAS  
CURSO DE BIOMEDICINA**

**GABRIEL VINÍCIUS ROLIM SILVA**

**USO DE TÉCNICAS *AB INITIO* NO ESTUDO DE INTERAÇÕES ENTRE DOIS  
POTENCIAIS COMPOSTOS FARMACOLÓGICOS E A PROTEASE DA CEPA  
MUTANTE DO VÍRUS HIV-1**

**NATAL/RN**

**2024**

GABRIEL VINÍCIUS ROLIM SILVA

**USO DE TÉCNICAS *AB INITIO* NO ESTUDO DE INTERAÇÕES ENTRE DOIS  
POTENCIAIS COMPOSTOS FARMACOLÓGICOS E A PROTEASE DA CEPA  
MUTANTE DO VÍRUS HIV-1**

Monografia apresentada ao curso de graduação em Biomedicina, da Universidade Federal do Rio Grande do Norte, como requisito parcial à obtenção do título de Bacharel em Biomedicina.

Orientadora: Prof. Dr. Umberto Laino Fulco  
Co-orientadora: MSc. Gabriela de Lima Menezes

NATAL/RN  
2024



Esta obra está licenciada com uma licença Creative Commons Atribuição 4.0 Internacional. Permite que outros distribuam, remixem, adaptem e desenvolvam seu trabalho, mesmo comercialmente, desde que creditem a você pela criação original. Link dessa licença: [creativecommons.org/licenses/by/4.0/legalcode](https://creativecommons.org/licenses/by/4.0/legalcode)

Universidade Federal do Rio Grande do Norte - UFRN  
Sistema de Bibliotecas - SISBI  
Catalogação de Publicação na Fonte. UFRN - Biblioteca Central Zila Mamede

Silva, Gabriel Vinicius Rolim.

Uso de técnicas ab initio no estudo de interações entre dois potenciais compostos farmacológicos e a protease da cepa mutante do vírus HIV-1 / Gabriel Vinicius Rolim Silva. - 2024.

56 f.: il.

Monografia (graduação) - Universidade Federal do Rio Grande do Norte, Centro de Biociências, Curso de Biomedicina. Natal, RN, 2024.

Orientação: Prof. Dr. Umberto Laino Fulco.

Coorientação: Me. Gabriela de Lima Menezes.

1. Dinâmica Molecular - Monografia. 2. Ab initio - Monografia. 3. Teoria do Funcional da Densidade - DFT - Monografia. 4. Fracionamento Molecular com Capas Conjugadas - MFCC - Monografia. 5. Interações Proteína-ligante - Monografia. 6. Inibidor da Protease - Monografia. I. Fulco, Umberto Laino. II. Menezes, Gabriela de Lima. III. Título.

RN/UF/BCZM

CDU 544.112:578.828

GABRIEL VINÍCIUS ROLIM SILVA

USO DE TÉCNICAS *AB INITIO* NO ESTUDO DE INTERAÇÕES ENTRE DOIS  
POTENCIAIS COMPOSTOS FARMACOLÓGICOS E A PROTEASE DA CEPA  
MUTANTE DO VÍRUS HIV-1

Monografia apresentada ao curso de  
graduação em Biomedicina, da  
Universidade Federal do Rio Grande do  
Norte, como requisito parcial à obtenção  
do título de Bacharel em Biomedicina.

Aprovada em: \_\_\_\_/\_\_\_\_/\_\_\_\_

BANCA EXAMINADORA

---

Prof. Dr. Umberto Laino Fulco  
Departamento de Biofísica e Farmacologia - UFRN

---

MSc. Gabriela de Lima Menezes  
Departamento de Bioinformática - IMD/UFRN

---

Dr<sup>a</sup>. Katyanna Sales Bezerra  
Unicamp

Dedico este trabalho a todos que participaram da minha trajetória acadêmica e de vida: família, amigos, professores, colegas, e todos aqueles que deixaram sua marca em meu caminho.

## AGRADECIMENTOS

Primeiramente, à espiritualidade que sempre amparou e guiou meus passos e da minha família: Oxalá, dono do meu ori; Iemanjá, de minha mãe Ana Karine; Oxum, da minha tia Eyrijane; Obaluaê, da minha irmã Mayana; Xangô, do meu pai Edmar e minha avó Antônia; Oxóssi, do meu primo Rafael e tio Edivan. E por fim, ao meu saudoso avô, Zé do Zé.

Agradeço a minha mãe Ana Karine e meu pai Edmar, que tanto lutaram e me apoiaram desde meus primeiros momentos nesse mundo. Todas as minhas conquistas e as muitas outras que virão são graças a vocês, e dedicadas a vocês.

A minha avó Antônia - “Branquinha” -, por todo o apoio, a ajuda e em especial o carinho que você sempre me ofereceu

Aos meus tios Edivan e Eyrijane, meus segundos pais que me acolheram de braços abertos em suas vidas e ajudaram a moldar a pessoa que sou hoje.

A minha irmã Mayana e meu “irmão” Rafael, por toda a companhia e momentos divertidos que compartilhei com vocês, e todos os PdL por vir no futuro!

Aos meus orientadores e professores Umberto e Gabriela, por todo o acolhimento desde o início de minha trajetória e suas preciosas orientações e apoio que levarei durante toda a caminhada acadêmica que o futuro ainda me reserva!

A todos os outros membros do laboratório que tão bem me acolheram por esses anos: professores Jonas e Katy, colega de pesquisa Klaus, e todo o resto da equipe de pós-doutorandos, doutorandos, mestrandos e ICs.

Aos amigos que tanto marcaram minha vida na UFRN: João Paulo, Rian, Bruna, Joama e Thayane. Jamais esquecerei todos os momentos que passamos rindo (de alegria ou de desespero), todos os perrengues que enfrentamos juntos, e todo o apoio que damos diariamente uns aos outros para superarmos esta etapa de nossas vidas. Obrigado por tudo, e que nossa amizade dure por muitos anos.

Aos amigos de longa data do IFRN: Fernando, Douglas, Robson, Kennedy, Navarro e Rodolfo, uma amizade que já dura mais de uma década, e que com certeza durará outras várias.

A minha parceira Eloísa, por todo o amor, carinho e apoio incondicional que você me oferece diariamente, alegrando e animando cada dia meu. Muito obrigado pela sua presença e por fazer parte da minha vida.

*They asked me how well I understood theoretical physics. I said I had a theoretical degree in physics. They said welcome aboard.*

Fantastic

## RESUMO

A inibição da protease do HIV-1 é um marco da terapia anti-retroviral. Entretanto, a notória habilidade do HIV-1 de desenvolver resistência contra inibidores da protease (IPs) apresenta um desafio significativo ao inutilizar estratégias terapêuticas disponíveis. Entre os inibidores de protease disponíveis no mercado, o Darunavir (DRV) é uma das principais opções de tratamento para pacientes com infecções por variantes mutantes do HIV. No entanto, a emergência de cepas resistentes ao DRV é particularmente perigosa, pois reduz a eficácia deste medicamento essencial, limitando as opções terapêuticas disponíveis e aumentando o risco de falhas no tratamento. Sob esta ótica, este estudo busca representar as interações entre dois novos IPs, GRL-004 e GRL-063, com a protease (PR) de uma variante mutante (MUT) do HIV-1, utilizando ferramentas de ancoragem e simulações de dinâmica molecular (DM) para obter as conformações ideais entre ambos fármacos e a PR mutante do vírus. As simulações utilizaram cristais obtidos do Protein Data Bank (PDB). Após a seleção das estruturas correspondentes do GRL-004, GRL-063, e PR mutante com resistência adquirida ao DRV, uma série de otimizações foram feitas na proteína utilizando para identificar suas conformações de ligação ideal formando o complexo proteína-ligante apropriado. Estas conformações foram submetidas a simulações em DM de forma a selecionar o menor estado energético e a conformação mais estável entre todas as conformações obtidas. Em posse da conformação ideal, um destes complexos (PR<sub>MUT</sub>-GRL-063) foi submetido a cálculos quânticos para determinar suas energias de interação. De maneira notável, não foram identificadas nenhuma repulsão nas interações entre o ligante GRL-063 e a protease mutante do HIV. Aminoácidos VAL50A e VAL50B demonstram bons resultados energéticos ante sua importância funcional na protease e reforçam a eficácia do ligante. Outros aminoácidos de destaque pelas energias apresentadas foram PRO81B, ARG8A, PHE82A\*, e ASN25A\*. Este estudo aprimora nossa compreensão da dinâmica receptor-ligante e da adaptabilidade dos novos inibidores de protease, trazendo profundas implicações para a inovação de futuros medicamentos antirretrovirais.

Palavras-chave: Dinâmica molecular, Ab initio, DFT, MFCC, Interações proteína-ligante, Inibidor da protease

## ABSTRACT

The inhibition of HIV-1 protease is a cornerstone of antiretroviral therapy. However, the notorious ability of HIV-1 to develop resistance against protease inhibitors (PIs) presents a significant challenge by turning currently available therapeutic options ineffective. Among the protease inhibitors available on the market, Darunavir (DRV) is one of the primary treatment options for patients with infections caused by mutant variants of HIV. However, the emergence of strains resistant to Darunavir is particularly dangerous, as it reduces the efficacy of this essential drug, limits the available therapeutic options, and increases the risk of treatment failure. From this perspective, this study aims to represent the interactions between two new PIs, GRL-004 and GRL-063, with the protease (PR) of a mutant (MUT) variant of HIV-1, using docking tools and molecular dynamics (MD) simulations to obtain the ideal conformations between both drugs and the mutant PR of the virus. Our simulations utilized crystals obtained from the Protein Data Bank (PDB). After selecting the corresponding structures of GRL-004, GRL-063, and mutant PR with acquired resistance to DRV, a series of optimizations were performed on the protein to identify its ideal binding conformations, forming the appropriate protein-ligand complex. These conformations were subjected to MD simulations to select the lowest energy state and the most stable conformation among all the obtained replicas. With the ideal replica in hand, one of these complexes (PR<sub>MUT</sub>-GRL-063) was subjected to quantum calculations to determine its interaction energies. Notably, no repulsion was identified in the interactions between the GRL-063 ligand and the mutant HIV protease. Amino acids VAL50A\* and VAL50B\* demonstrated good energetic results given their functional importance in the protease, reinforcing the ligand's efficacy. Other amino acids highlighted by the presented energies were PRO81B, ARG8A, PHE82A\*, and ASN25A\*. This study enhances our understanding of receptor-ligand dynamics and the adaptability of new protease inhibitors, carrying profound implications for the innovation of future antiretroviral medications.

Keywords: Molecular dynamics, Ab initio, DFT, MFCC, Protein-ligand interactions, Protease inhibitor

## LISTA DE FIGURAS

|                    |  |    |
|--------------------|--|----|
| <b>Figura 1</b> -  | GRL-142 e modificações propostas para síntese do GRL-004 e GRL-063.....                              | 16 |
| <b>Figura 2</b> -  | Resultados da ancoragem molecular dos ligantes GRL-004 e GRL-063 na proteína PR <sub>MUT</sub> ..... | 29 |
| <b>Figura 3</b> -  | Análise das simulações de dinâmica molecular.....  | 31 |
| <b>Figura 4</b> -  | Análise MQ/MM dos últimos 500 modos conformacionais de cada complexo.....                            | 32 |
| <b>Figura 5</b> -  | GRL-063 e suas subdivisões.....  | 33 |
| <b>Figura 6</b> -  | GRL-063 interagindo com a protease.....  | 34 |
| <b>Figura 7</b> -  | Gráfico de convergência do GRL-063.....  | 35 |
| <b>Figura 8</b> -  | Gráfico BIRD do GRL-063.....   | 36 |
| <b>Figura 9</b> -  | Resíduos interagindo com a região I do GRL-063.....  | 38 |
| <b>Figura 10</b> - | Resíduos interagindo com a região II do GRL-063.....   | 40 |
| <b>Figura 11</b>   | Resíduos interagindo com a região III do GRL-063.....  | 41 |

## LISTA DE TABELAS

|  |    |
|--|----|
| <b>Tabela 1</b> – Resultados da ancoragem molecular..... | 29 |
| <b>Tabela 2</b> – Resultados dos cálculos MQ/MM.....     | 31 |

## LISTA DE ABREVIACOES

|      |                     |
|------|---------------------|
| MUT  | Mutante             |
| Vina | Autodock Vina 1.1.2 |

## LISTA DE SIGLAS

|        |  |
|--------|--|
| AIDS   | Nome por extenso da Sigla e em ordem alfabética  |
| ARG    | Nome por extenso                                 |
| ASN    | Nome por extenso                                 |
| ASP    | Nome por extenso                                 |
| AZT    | Zidovudine                                       |
| CHARMm | Chemistry at Harvard Molecular mechanics         |
| CPCM   | Modelo Contínuo Polarizável Condutor             |
| DFT    | Teoria do Funcional da Densidade                 |
| DM     | Dinâmica Molecular                               |
| DRV    | Darunavir  |
| EM     | Minimização Energética                           |
| FDA    | Food and Drug Administration                     |
| fs     | femtosegundos                                    |
| GBSA   | Generalized Born surface area                    |
| GLY    | Glicina  |
| HAART  | Terapia Anti-retroviral Altamente Eficaz         |
| HIV    | Síndrome da Imunodeficiência Adquirida           |
| ILE    | Isoleucina                                       |
| IP     | Inibidor da Protease                             |
| MFCC   | Fracionamento Molecular com Capas Conjugadas     |
| MM     | Mecânica Molecular                               |
| MQ     | Mecânica Quântica                                |
| nm     | nanômetro  |
| NRTI   | Inibidores Nucleosídeos da Transcriptase Reversa |
| ns     | nanosegundos                                     |
| PDB    | Protein Data Bank                                |
| PHE    | Fenilalanina                                     |
| PME    | Particle Mesh Ewald                              |
| PR     | Protease   |
| PRO    | Prolina  |
| ps     | picosegundos                                     |
| RMSD   | Raiz Quadrada do Erro Médio                      |

|      |                                  |
|------|----------------------------------|
| RMSF | Raiz Quadrada da Flutuação Média |
| SCRF | Campo de Reação Auto-Consistente |
| THR  | Treonina                         |
| VAL  | Valina                           |
| WT   | Cepa Selvagem                    |

## SUMÁRIO

|                                    |           |
|------------------------------------|-----------|
| <b>1 INTRODUÇÃO</b>                | <b>17</b> |
| 1.1 Panorama global do HIV         | 17        |
| 1.2 Métodos <i>ab initio</i>       | 19        |
| 1.3 Ancoragem e dinâmica molecular | 19        |
| 1.4 MFCC                           | 20        |
| 1.5 DFT                            | 21        |
| 1.6 Justificativa                  | 21        |
| <b>2 OBJETIVOS</b>                 | <b>23</b> |
| 2.1 Objetivo geral                 | 23        |
| 2.2 Objetivo específico            | 23        |
| <b>3 MATERIAL E MÉTODOS</b>        | <b>23</b> |
| 3.1 Preparo dos cristais           | 23        |
| 3.2 Ancoragem e dinâmica molecular | 24        |
| 3.4 MFCC e DFT                     | 26        |
| <b>4 RESULTADOS E DISCUSSÃO</b>    | <b>29</b> |
| 4.1 Ancoragem e dinâmica molecular | 29        |
| 4.2 Análise de convergência e BIRD | 34        |
| 4.3 Região I                       | 38        |
| 4.4 Região II                      | 39        |
| 4.5 Região III                     | 41        |
| <b>5 CONCLUSÃO</b>                 | <b>43</b> |
| <b>REFERÊNCIAS</b>                 | <b>45</b> |
| <b>APÊNDICE</b>                    | <b>51</b> |

## • 1 INTRODUÇÃO

### ○ 1.1 PANORAMA GLOBAL DO HIV

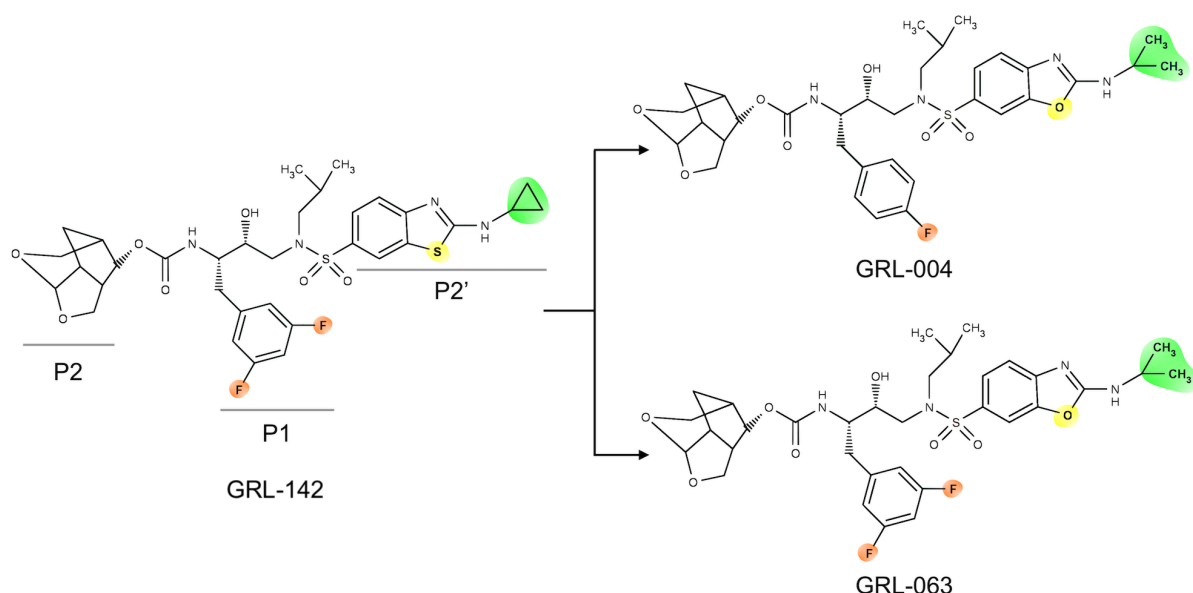
Apesar das terapias anti-retrovirais eficazes terem reduzido significativamente as mortes pelo Vírus da Imunodeficiência Humana (HIV, do inglês *Human Immunodeficiency Virus*), o mesmo ainda afeta 39 milhões de pessoas no mundo inteiro ([www.who.int/data/gho/data/themes/hiv-aids](http://www.who.int/data/gho/data/themes/hiv-aids)). Entre 2010 e 2019, regiões como a Europa Central, Europa Oriental e Ásia Central viram um aumento de 107.45% no número de pessoas infectadas (Fanales-Belasio *et al.*, 2010). O HIV se liga aos receptores CD4 e co-receptores de quimiocinas nas células do tecido linfático, resultando na diminuição das células T CD4+ e, eventualmente, progressão para a Síndrome da Imunodeficiência Adquirida (AIDS, do inglês *Acquired Immunodeficiency Syndrome*). Esta progressão abre caminho para infecções oportunistas e sintomas graves, aumentando a mortalidade (Fanales-Belasio *et al.*, 2010).

Nos anos 1980, o desenvolvimento de terapias anti-retrovirais, como o Zidovudine (AZT), foi uma marca significativa no combate à doença. O AZT, um fármaco da classe de inibidores nucleosídeos da transcriptase reversa (NRTI), revelou grande eficácia clínica ao inibir a síntese de DNA viral. Atualmente, a Terapia Anti-Retroviral Altamente Eficaz (HAART) é a principal abordagem de combate a doença, combinando diferentes classes de drogas anti-retrovirais para prevenir o surgimento de cepas resistentes a fármacos de primeira linha (Broder, 2010; Yarchoan *et al.*, 1989). Entretanto, a necessidade de tratamento contínuo é crucial devido à capacidade do vírus de permanecer latente nas células hospedeiras. Os IPs se apresentam, neste contexto, como alternativa para o combate da infecção pelo vírus HIV. Eles funcionam, como o nome sugere, pela inibição da atividade da protease viral, impedindo o processo de clivagem das poliproteínas em suas proteínas estruturais e não-estruturais subsequentes, assim levando a montagem de partículas virais inefetivas (Faray & Craik, 2010). O Duranavir (DRV), um dos últimos inibidores de protease (IPs) aprovados pela agência norteamericana Food and Drug Administration (FDA), se destaca pela sua alta eficácia aliada à grande barreira genética. Atualmente, o DRV é um dos principais fármacos de segunda linha de combate à AIDS.

Embora o DRV seja altamente eficaz, possíveis mutações que garantem ao vírus resistência a este IP já foram reportados na literatura (Aoki *et al.*, 2018). Recentemente, Bulut *et al.* (2020) propuseram uma série de novos IPs baseados na estrutura do DRV, introduzindo modificações pontuais que alterem características moleculares específicas. Estes novos IPs mostraram propriedades promissoras *in vitro*, incluindo permeabilidade celular e estabilidade térmica, e foram particularmente eficazes contra cepas resistentes ao DRV.

Nesta pesquisa, analisamos dois novos IPs como propostos por Bulut *et al.* (2020): GRL-004 e GRL-063, derivados do IP previamente descrito na literatura GRL-142 (Aoki *et al.*, 2017). O GRL-004 inclui a introdução de um átomo de flúor na posição para em P1, substituição de enxofre por oxigênio, e cisão do grupo ciclopropil em P2' (Figura 1); já o fármaco GRL-063 mantém as mudanças em P2 de GRL-004, mas reteve as estruturas de flúor originais em P1 (Figura 1).

Figura 1 - Estrutura do GRL-142 e modificações propostas para sintetizar as moléculas do GRL-004 e GRL-063



Fonte: Elaborado pelo autor, 2024

As interações entre proteínas, DNA/RNA e ligantes são fundamentais para muitos processos biológicos, como reconhecimento celular e ações enzimáticas. Métodos experimentais para obter estas interações são caros e limitados, tornando

métodos computacionais uma alternativa viável. A ancoragem molecular se apresenta como ferramenta computacional capaz de prever certos comportamentos moleculares por meio de simulações computacionais, utilizando sistemas de pontuação empíricas para prever a ligação de inibidores a enzimas. Com as coordenadas iniciais do sistema, a dinâmica molecular (DM) melhora a simulação em solução, considerando condições fisiológicas e analisando a dinâmica das proteínas e suas interações com os ligantes.

Neste trabalho, utilizamos metodologias como a ancoragem molecular e simulações de dinâmica molecular para encontrar a conformação ideal da ligação entre os dois fármacos propostos por Bulut *et al.* (2020) e a protease viral do HIV-1 obtida de uma cepa mutante que adquiriu resistência ao DRV.

## ○ 1.2 MÉTODOS *AB INITIO*

Métodos *ab initio* são métodos que buscam explicar fenômenos sem utilização de dados experimentais ou outras ferramentas empíricas, se baseando apenas na utilização de cálculos e modelos físicos para prever o comportamento de, por exemplo, proteínas partindo apenas de suas sequências de aminoácidos (CARLONI *et al.*, 2002; CHIVIAN *et al.*, 2003). Na área de bioquímica, esses métodos são particularmente valiosos porque permitem a análise de sistemas sem a necessidade de informações próprias do sistema em questão de maneira prévia. Sendo assim, abre-se a possibilidade de analisar estruturas bioquímicas cada vez mais complexas por meio de métodos computacionais. A abordagem *ab initio* tem com base a solução da equação de Schrödinger para sistemas eletrônicos, utilizando-se de metodologias como, no caso de sistemas biológicos, a teoria do funcional da densidade (DFT, do inglês *density functional theory*) (Bartlett, 2010).

## ○ 1.3 ANCORAGEM E DINÂMICA MOLECULAR

Para a análise das energias de interação entre os ligantes GRL-004 e GRL-063 e a protease mutante do HIV-1, fez-se necessária a utilização de ferramentas para estabelecer as conformações mais bem otimizadas entre cada

ligante. Dentre as ferramentas utilizadas, destacam-se a ancoragem e a dinâmica molecular.

A ancoragem molecular é uma técnica de modelagem computacional que utiliza simulações para prever a melhor conformação de interação tri-dimensional entre unidades biológicas como, a exemplo deste trabalho, uma molécula e sua proteína-alvo (Salmaso & Moro, 2018). Este método é especialmente útil para estudar a melhor conformação especial da interação entre moléculas e, neste contexto, disponibilizar uma estrutura de um complexo proteína-ligante para estudos de interação. Dessa forma, a ancoragem se demonstra ferramenta essencial na análise de potenciais compostos farmacêuticos na atualidade (Ahuja *et al.*, 2022)

A DM, por outro lado, analisa o comportamento de dada estrutura biológica ao longo do tempo a partir de uma série de parâmetros, a exemplo da temperatura, pressão, e interações com determinado solvente. Assim, a DM se dispõe como uma técnica para determinar, a partir da conformação obtida na ancoragem, o estado mais estável do complexo estudado, com possibilidade de uso nas mais diversas áreas do conhecimento (Schneider *et al.*, 2008). Dessa forma, este trabalho se utilizou da ancoragem como método de obtenção de uma estrutura molecular e da DM como ferramenta para análise de sua conformação mais estável.

#### ○ 1.4 MFCC

Para analisar com precisão as energias de interação do complexo proteína-ligante entre o sistema GRL-063-PR<sub>MUT</sub>, utilizou-se o método do MFCC (Fracionamento Molecular com Capas Conjugadas, do inglês *molecular fractionation with conjugated caps*) em conjunto com o DFT. Esse método envolve a fragmentação de cada monômero da protease em seus resíduos de aminoácidos correspondentes, calculando as interações energéticas separadamente para cada resíduo. Esta separação ocorre através da quebra das ligações peptídicas nas extremidades C- e N-terminais de cada aminoácido. Para fechar as valências abertas, outros aminoácidos (chamados de capas) são adicionados e, por sua vez, as valências das últimas capas utilizadas serão fechadas com átomos de hidrogênio. Este método permite o cálculo da energia total de interação do sistema, utilizando-se do DFT, pela soma das energias individuais de cada aminoácido fragmentado

(Zhang & Zhang, 2003). Essa metodologia permite uma abordagem com a aplicação de técnicas de MQ mantendo um custo computacional razoável, proporcionando resultados mais precisos em comparação com simulações baseadas apenas em MM. Trabalhos recentes demonstraram a eficácia desta metodologia no estudo e análise das interações energéticas existentes entre sistemas biológicos, assim justificando seu uso se comparado a metodologias baseadas apenas em MM (Ourique *et al.*, 2016; Diego *et al.*, 2015; Campos *et al.*, 2020).

- 1.5 DFT

Os cálculos propriamente ditos das interações energéticas do complexo proteína-ligante foram obtidos através do funcional do DFT. Este método leva em conta a distribuição da densidade eletrônica de um certo átomo para o cálculo de sua energia de afinidade com outras estruturas, assim permitindo analisar compostos biológicos a um custo computacional aceitável ante a complexidade de suas estruturas (Bartlett *et al.*, 2010; Argaman & Markov, 2000).

- 1.5 JUSTIFICATIVA

A pandemia da AIDS se destaca como um dos principais problemas de saúde pública da atualidade, demandando esforços contínuos com o objetivo de frear sua disseminação e garantir melhor qualidade de vida para indivíduos infectados pelo vírus do HIV-1, sua variante mais comum. Segundo a FDA, em 2020, 222 medicamentos antiretrovirais se encontram atualmente aprovados para uso clínico; entretanto, devido à natureza do vírus, várias cepas resistentes a diversos fármacos (e suas combinações) podem ser encontradas pelo mundo (Nastri *et al.*, 2023). Sob esta ótica, faz-se necessário a pesquisa constante por opções terapêuticas efetivas não somente contra cepas selvagens do vírus, mas também variantes mutantes que o paciente possa porventura adquirir ao longo do tratamento.

A análise experimental de novos fármacos se mostra a forma mais comum de avaliar seus parâmetros *in vitro* e obter detalhes sobre seu possível comportamento *in vivo* e interações com seus receptores desejados. Recentemente, porém, a utilização de ferramentas computacionais para determinar o comportamento de biomoléculas vem ganhando espaço na literatura devido a sua precisão, liberdade

de modificação de parâmetros, e rapidez se comparado a estudos experimentais (Hollingsworth & Dror, 2018). Assim, neste trabalho é realizada a análise das energias de interação entre dois novos fármacos – GRL-004 e GRL-063 – e duas proteases de duas cepas distintas do HIV-1: uma variante selvagem, e uma mutante com resistência adquirida ao DRV.

Com o objetivo de prover estruturas moleculares que melhor representam as interações existentes entre os dois fármacos e a protease mutante, uma série de métodos em dinâmica molecular foram aplicados. Em um primeiro momento, foi realizado a ancoragem molecular para predizer, por meio de pontuações empíricas, a forma como o inibidor se liga à proteína. Este método tem como principal vantagem a possibilidade de verificar a principal conformação estrutural do complexo formado a baixo custo computacional (Halperin *et al.*, 2002). Com os complexos iniciais em mãos, é possível realizar a otimização do sistema por meio da dinâmica molecular, permitindo não só adaptar o sistema a condições fisiológicas mais apropriadas mas também analisar como a dinâmica da proteína pode afetar as interações intermoleculares do complexo proteína-ligante (Karplus & McCammon, 2002). A partir desta otimização, podemos prosseguir com cálculos energéticos utilizando outros métodos *ab initio*, a exemplo da metodologia do MFCC aplicada em conjunto ao DFT para obter uma visão de como o ligante trabalhado interage com o receptor desejado.

Levando em conta o contexto da pandemia da AIDS, necessidade do desenvolvimento de novos fármacos e recentes avanços computacionais no campo da bioinformática, este trabalho busca demonstrar a utilização de ferramentas computacionais como método eficaz e preciso para o estudo das energias de interação de um complexo proteína-ligante formado por duas novas drogas inibidoras da protease viral e cepas selvagem e mutante do vírus, e descrevendo o comportamento e características de suas principais interações em um destes complexos.

## ● 2 OBJETIVOS

### ○ 2.1 OBJETIVO GERAL

Analisar a conformação ideal da interação entre dois inibidores da protease do HIV-1 e a protease viral de uma cepa mutante que adquiriu resistência ao fármaco Darunavir, e realizar a análise energética de um destes complexos.

### ○ 2.2 OBJETIVO ESPECÍFICO

- Avaliar a estrutura ideal de cada ligante para seu estado fisiológico;
- Analisar o estado de protonação e energético ideal da proteína;
- Analisar as melhores conformações de ligação através da ancoragem molecular entre os ligantes GRL-004 e GRL-063 com a protease viral;
- Verificar as réplicas com melhores resultados nas simulações de dinâmica molecular;
- Analisar as energias de interação do complexo GRL-063-PR<sub>MUT</sub> e os principais resíduos de aminoácidos.

## ● 3 MATERIAL E MÉTODOS

### ○ 3.1 PREPARO DOS CRISTAIS

Nossas simulações foram feitas com três estruturas cristalográficas de raio-X de uma protease WT do HIV-1 (PR<sub>WT</sub>) em complexo com os IPs GRL-004 (PDB ID 6OYD) e GRL-063 (PDB ID 6OGP), bem como uma protease resistente ao DRV (PR<sub>MUT</sub>) (PDB ID 5B18), todas obtidas através do PDB a uma resolução de 1.53, 1.46 e 1.80 Å, respectivamente. As mutações observadas na PR<sub>MUT</sub> foram: V11I, L23I, D25N, V32I, L33F, M46I, I47V, I50V, F53L, I54V, Q58E, D60E, L63P, H69R, A71V, G73S, V77I, V82F, L89V, L90M e I93L.

Por mais que a técnica de difração por raio-X seja capaz de prover um olhar preciso da estrutura do complexo proteína-ligante, ela é incapaz de representar corretamente átomos com baixo peso molecular, como o hidrogênio, e o cristal pode não estar no seu estado mais ideal para simulações em ambiente fisiológico. Dessa

forma, uma série de otimizações foram feitas com o intuito de fornecer uma representação mais fidedigna do comportamento da proteína no corpo humano.

Os átomos pesados nos complexos das proteases foram fixados, e seus respectivos átomos de hidrogênio foram adicionados, com as cadeias laterais dos resíduos corrigidas, se necessário. Logo após, a proteína foi ajustada de acordo com os resultados de protonação obtidos pelo PropKa Online 2.0. O ligante foi adaptado para sua conformação em pH fisiológico por uso do MarvinSketch, analisando o estado de protonação de seus respectivos IPs na corrente sanguínea (7.2 a 7.4); assim, cargas foram adicionadas e removidas no ligante quando necessário, enquanto interações não presentes na estrutura foram corrigidas. Com o complexo proteína-ligante corrigido para sua conformidade *in vivo* mais provável, suas cadeias laterais foram geometricamente otimizadas por uso do CHARMM (*Chemistry at Harvard Molecular Mechanics*), um campo de força especificamente modelado para aplicações em compostos bioquímicos, visando aumentar a precisão de nossos cálculos.

### ○ 3.2 ANCORAGEM E DINÂMICA MOLECULAR

Para avaliar as energias de interação dos ligantes GRL-063 e GRL-004 com a protease mutante, análises dos complexos PR<sub>MUT</sub>-GRL-063 e PR<sub>MUT</sub>-GRL-004 foram conduzidas. Para este propósito, o receptor e os ligantes GRL-063 e GRL-004 foram obtidos de seus respectivos cristais seguindo as otimizações previamente descritas. Os ligantes foram convertidos do formato PDB para o formato PDBQT usando Meeko (Meeko, 2022), e a proteína foi preparada pelo *software* MGLTools 1.5.7. Neste caso, cargas Gasteiger foram computadas, átomos do tipo AD4 distribuídos, e a protease foi salva no formato PDBQT.

A ancoragem foi feita utilizando AutoDock Vina 1.1.2 (referido aqui como "Vina") (Trott & Olson, 2009) 10 vezes para cada ligante, onde cada ancoragem gerou 10 modos conformacionais, resultando em um total de 100 conformações para cada ligante. Os sítios de ligação escolhidos foram os mesmos observados nos complexos PR<sub>WT</sub>-GRL-063 e PR<sub>WT</sub>-GRL-004. As conformações obtidas no Vina passaram por cálculos adicionais das energias de interação entre a proteína e o ligante usando o programa Uni-GBSA (Yang *et al.*, 2023). Este programa foi

recentemente desenvolvido para realizar cálculos MM/GB(PB)SA através de resultados de triagem virtual, e já foi descrito como tendo uma boa relação entre dados experimentais e superando os resultados do score Vina. Os melhores complexos energéticos para cada ligante obtidos através dos cálculos Uni-GBSA foram então selecionados para as simulações de DM.

Simulações de DM foram realizadas para ajustar a conformação dos ligantes em seus receptores e avaliar sua estabilidade em solução. Para este propósito, para cada sistema, três réplicas independentes de 200 nanossegundos (ns) foram feitas usando o *software* GROMACS 2023 (Abraham *et al.*, 2015). Inicialmente, os parâmetros dos ligantes devem ser gerados no formato GROMACS. Para este objetivo, o servidor ACPYPE (Sousa & Vranken, 2012) foi utilizado, aplicando carga Gasteiger como método e GAFF2 como campo de força. Para a proteína, o campo de força utilizado foi o Amberff99SB-ILDN, cujos parâmetros já se encontram disponíveis no GROMACS.

Para cada sistema (PR<sub>MUT</sub>-GRL-063 e PR<sub>MUT</sub>-GRL-004), uma caixa em formato cúbico foi construída, e a água modelo TIP3P foi adicionada de forma que a distância mínima entre a superfície da caixa e o soluto (complexo protease) foi de 12 Å. Adicionalmente, íons de Cl<sup>-</sup> foram adicionados para neutralizar o sistema. Duas rodadas de minimização de energia (EM, do inglês *energy minimization*) foram realizadas para remover contatos desfavoráveis com a estrutura inicial. A primeira EM foi limitada a um máximo de 20,000 etapas ou até que a força máxima agindo em cada átomo fosse menor que 50 kJ/mol/nm, usando o algoritmo de steepest descent com a proteína e ligante restringidos em suas posições de tal forma que a EM foque no relaxamento do solvente. A segunda EM, sem as restrições de posições da proteína e ligante, usou um solvente flexível e o mesmo algoritmo, sendo executado com um máximo de 10,000 etapas ou até a força aplicada a cada átomo ser menor que 250 kJ/mol/nm.

A temperatura do sistema foi ajustada para 298 K, a mesma observada nos experimentos *in vitro* dos respectivos ligantes e PR<sub>WT</sub>, ao longo de 100 picossegundos (ps) usando o algoritmo modificado de Berendsen (Berendsen *et al.*, 1984) para ajuste e controle da temperatura. Uma segunda etapa, que consistiu no ajuste da pressão, foi realizada ao longo de 100 picossegundos (ps) usando o

algoritmo de Parrinello-Rahman (Hutter, 2011) para controle da pressão. Durante ambos os processos, os átomos de hidrogênio foram restritos utilizando o algoritmo LINCS (Hess *et al.*), e as restrições posicionais dos complexos proteína-ligante foram aplicadas para estabilizar o solvente em volta do soluto.

Interações de longo alcance foram computadas usando o método de soma Particle Mesh Ewald (PME), aplicando um *cut-off* de 1 nm para interações não-ligadas. As equações de movimento foram integradas utilizando o algoritmo leap-frog (Hockney *et al.*, 1974) com um tempo por etapa de 0.2 femtosegundos (fs). Antes do estágio de produção da DM, uma pequena etapa de 1 ns foi conduzida com o complexo protéico sem restrições de posição, seguida por uma etapa de produção durante 200 ns. Cada simulação de DM produziu um total de 2000 frames conformacionais.

As trajetórias resultantes das simulações de DM foram visualizadas usando o programa UCSF Chimera (Pettersen *et al.*, 2004), e a raiz quadrada do erro médio (RMSD) e flutuação (RMSF) foram calculados usando o comando "gmx" no pacote GROMACS.

Para obter o complexo de menor energia dentre as simulações de DM, cálculos híbridos de MQ/MM-GBSA foram realizados utilizando o programa gmx\_MMPBSA (Valdés-Tresanco, M. S. *et al.*, 2021). Para isto, os últimos 50 ns (500 frames) de cada simulação foram utilizados. A região MQ foi limitada para resíduos até 5 Å de distância do ligante, e a função semi-empírica PM6-DH+ foi aplicada. O mesmo campo de força usado na DM foi usado para os cálculos dos resíduos remanescentes. Solventes moleculares explícitos e íons foram excluídos da análise. Os menores complexos de energia para cada sistema PR<sub>MUT</sub>-IP foram selecionados, com o complexo GRL-063-PR<sub>MUT</sub> sendo escolhido para as análises energéticas, enquanto o GRL-004-PR<sub>MUT</sub> se encontra disponível para análise em pesquisas futuras.

### ○ 3.3 MFCC E DFT

Com o objetivo de calcular de forma precisa as energias de interação do complexo proteína-ligante do sistema escolhido (GRL-063-PR<sub>MUT</sub>), utilizamos o método MFCC (Zhang & Zhang, 2003) com cálculos baseados no DFT para

obtenção de resultados completamente *ab initio* por uso da MQ. O método previamente citado consiste na fragmentação de cada monômero da PR em seus resíduos de aminoácidos através da clivagem de ligações peptídicas e calculando as interações energéticas separadamente para cada resíduo, uma vez que a soma das energias individuais equivale a energia total de interação do ligante com cada monômero. Considere  $R_i$  como o *i-ésimo* resíduo de um monômero de uma proteína  $P$  que interage com um ligante  $L$ . As capas conjugadas (Camila *et al.*, 2023)  $C_i$  e  $C_i^*$  são definidas como os resíduos ligados aos grupos amina e carboxila, respectivamente, de  $R_i$  com interações covalentes adicionais a átomos de hidrogênio em cada extremidade, mantendo as valências e além disso melhorando de maneira geral a simulação do ambiente eletrônico real envolvendo as interações entre  $R_i$  e  $L$ . Para o resíduo N(C)-terminal, a capa  $C_iC_i^*$  se encontra vazia. Com o objetivo de calcular a energia de interação ( $EI_i$ ) entre  $L$  e  $R_i$ , construímos quatro subsistemas distintos,  $A_i=(L - C_iR_iC_i^*)$ ,  $B_i=(C_iR_iC_i^*)$ ,  $C_i=(L - C_iC_i^*)$  e  $D_i=(C_iC_i^*)$ . Dessa forma,

$$EI_i = E(A_i) - E(B_i) - E(C_i) + E(D_i)$$

onde  $E(A_i)$ ,  $E(B_i)$ ,  $E(C_i)$  e  $E(D_i)$  são a energia total de cada subsistema.

Após aplicar o método de fragmentação do MFCC, cálculos energéticos foram realizados usando o *software* Gaussian 16 com o formalismo do DFT (Orio, Pantazis, Neese, 2009), uma abordagem baseada na MQ capaz de descrever com precisão as energias de interações intermoleculares a um custo computacional acessível (Mazurek, Łukasz Szeleszczuk, Dariusz Maciej Pisklak, 2020). Para este trabalho, utilizamos o funcional de gradiente de aproximação meta-generalizado (meta-GGA) B97D. Estudos comparativos envolvendo diferentes funcionais GGA (a exemplo de, mas não limitado a,  $\omega$ B97XD, B3LYP, B97D, B97D3, M06, e M062-X) apontou a eficiência do B97D em calcular interações e geometrias de estruturas orgânicas (Ovalle & Malardier-Jugroot, 2022; Chen *et al.*, 2020; Tsuzuki & Uchimar, 2020). Além disso, demonstra cálculos energéticos melhorados para complexos ligados de forma não-covalente e propriedades termoquímicas. De forma a expandir os orbitais Kohn-Sahm e melhor representar a função eletrônica de onda, selecionamos o conjunto de bases 6-311+G(d,p), um conjunto de valência dividida tripla (triplo- $\zeta$ ) empregando uma função difusa adicional (+) e funções de polarização (d,p).

Ao estudar propriedades biomoleculares por uma abordagem teórica, devido a presença do complexo em um ambiente fisiológico, onde ele necessariamente se encontraria envolvido por um solvente (seja o sangue, seja o citoplasma), faz-se necessária a análise do impacto do meio nas interações existentes por interferências eletrostáticas. De forma a simular este meio, a permissividade do meio é alterada por meio do solvente implícito Modelo Contínuo Polarizável Condutor (CPCM) aplicado ao método do Campo de Reação Auto-Consistente (SCRF). Este método consiste em envolver a molécula analisada em um solvente contínuo e, a partir disso, modificar a constante dielétrica do meio para valores apropriados para a análise desejada. Uma constante dielétrica menor está associada com uma menor permissividade do meio, resultando em uma superestimação das energias de interação ao permitir interações mais energéticas. Consequentemente, a energia total tende a diminuir à medida que aumentamos a permissividade. De acordo com dados da literatura, a constante dielétrica média para proteínas é igual a  $\epsilon = 3.23$ ; entretanto, estudos anteriores sugeriram que uma constante dielétrica maior pode ser utilizado como um bom meio termo entre simulações e resultados experimentais (Diego *et al.*, 2015; Ourique *et al.*, 2016; Amin & Küpper, 2020). Neste contexto, optamos por utilizar  $\epsilon = 10$  e  $\epsilon = 40$  para simular tanto um meio de menor permissividade que mais se aproxima ao valor médio da constante dielétrica de proteínas, quanto um valor  $\epsilon$  maior para prover resultados mais precisos em sistemas biológicos ao melhor representar a proteína em um estado de solvatação.

Considerando a distribuição espacial de resíduos de aminoácidos no sistema proteína-ligante, uma análise de convergência foi realizada para avaliar a energia total em função da distância à medida que o raio aumenta. Com o ligante no centro, o raio de interação foi definido como uma distância entre  $r = 2.0$  a  $10.0 \text{ \AA}$ , e a energia total de cada raio consecutivo foi observada. A convergência é alcançada quando a diferença total de energia entre um raio qualquer e seu raio seguinte for igual ou inferior a 10%; em outras palavras, assumimos que todas as interações intermoleculares essenciais ocorreram até o raio de convergência (Camila *et al.*, 2023; Vianna *et al.*, 2021).

#### ● 4 RESULTADOS E DISCUSSÃO

#### ○ 4.1 ANCORAGEM E DINÂMICA MOLECULAR

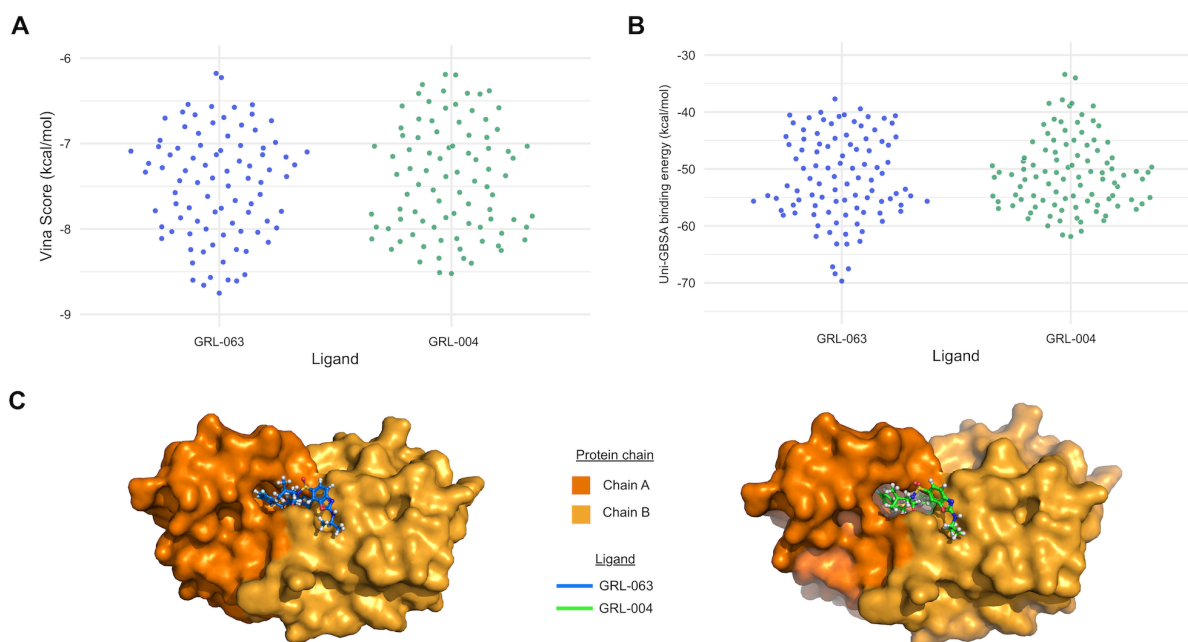
A ancoragem molecular realizado no programa Vina foi conduzido em 10 rodadas, cada uma fornecendo 10 conformações dos ligantes. Para o ligante GRL-063, o menor resultado foi de -8.751 kcal/mol, com uma média de -7.405 kcal/mol. Para o ligante GRL-004, o menor valor obtido foi de -8.552 kcal/mol, com uma média de -7.418 kcal/mol. A distribuição energética pode ser observada na Figura 2A. Todos esses modos conformacionais passaram por um novo estágio de cálculos de energia de afinidade no programa Uni-GBSA usando a metodologia MM/PBSA.

Nas análises Uni-GBSA, a melhor conformação energética para o GRL-063 foi de -69.69 kcal/mol, correspondendo com o valor de -7.801 kcal/mol no Vina. A média foi de -51.52 kcal/mol. Para o ligante GRL-004, a menor conformação energética foi de -61.87 kcal/mol, correspondendo ao valor de -7.072 kcal/mol no Vina, com uma média conformacional de -50.69 kcal/mol. A distribuição de energias no Uni-GBSA pode ser observada na Figura 2B, e os detalhes das energias na Tabela 1. Para ambos os *softwares*, foi sugerido que o ligante GRL-063 possui maior afinidade ao PR<sub>MUT</sub> se comparado ao GRL-004.

A melhor conformação de acordo com o Uni-GBSA para cada ligante em suas proteínas pode ser observada na Figura 2C. É possível notar que parte do ligante se mantém fora do sítio de ligação, o que pode ser explicado pelo fato de a proteína não estar em sua conformação ajustada ao ligante, devido a ter sido obtida do PDB sem nenhum inibidor ligado a ela, e a ancoragem foi realizada com um receptor rígido. Assim, a simulação de DM, onde ambos proteína e ligante estão flexíveis, é uma etapa crucial para o ajuste da conformação do complexo.

Figura 2 - Resultados da ancoragem molecular dos ligantes GRL-063 e GRL-004 na proteína PR<sub>MUT</sub>. A. Diagrama de Pontos da distribuição dos valores Vina para o ligante GRL-063 (pontos azuis) e ligante GRL-004 (pontos verdes). B. Diagrama de pontos da distribuição das energias de interação de acordo com o Uni-GBSA para o ligante GLR-063 (pontos azuis) e o ligante GRL-004 (pontos verdes). C. Melhores conformações dos ligantes GRL-063 (representado por um bastão azul) e GRL-004

(representado por um bastão verde). A proteína é representada por uma superfície laranja, com laranja escuro sendo a cadeia A e laranja claro, cadeia B.



Fonte: Elaborado pelo autor, 2024.

Tabela 1 - Resultados da ancoragem molecular pelos softwares Vina e Uni-GBSA.

| Valores / Energias de Interação           |         |        |                 |         |        |                 |        |                  |
|---|---------|--------|-----------------|---------|--------|-----------------|--------|------------------|
| Software                                  | Lig.    | Min.   | Q1 <sup>§</sup> | Mediana | Média  | Q3 <sup>#</sup> | Máx.   | Dp. <sup>‡</sup> |
| Vina                                      | GRL-063 | -8.75  | -7.98           | -7.40   | -7.48  | -7.04           | -6.18  | 0.6211           |
|   | GRL-004 | -8.52  | -7.98           | -7.41   | -7.42  | -6.91           | -6.19  | 0.6394           |
| Uni-GBSA                                  | GRL-063 | -69.69 | -56.54          | -52.28  | -51.52 | -45.07          | -37.71 | 7.46             |
|   | GRL-004 | -61.87 | -55.53          | -51.24  | -50.69 | -46.74          | -33.40 | 6.36             |
| #1° Quartil; §3° Quartil; ‡Desvio padrão. |         |        |                 |         |        |                 |        |                  |

Para analisar as conformações obtidas na ancoragem molecular em relação à estabilidade dos ligantes nas regiões de interação quando colocados em um

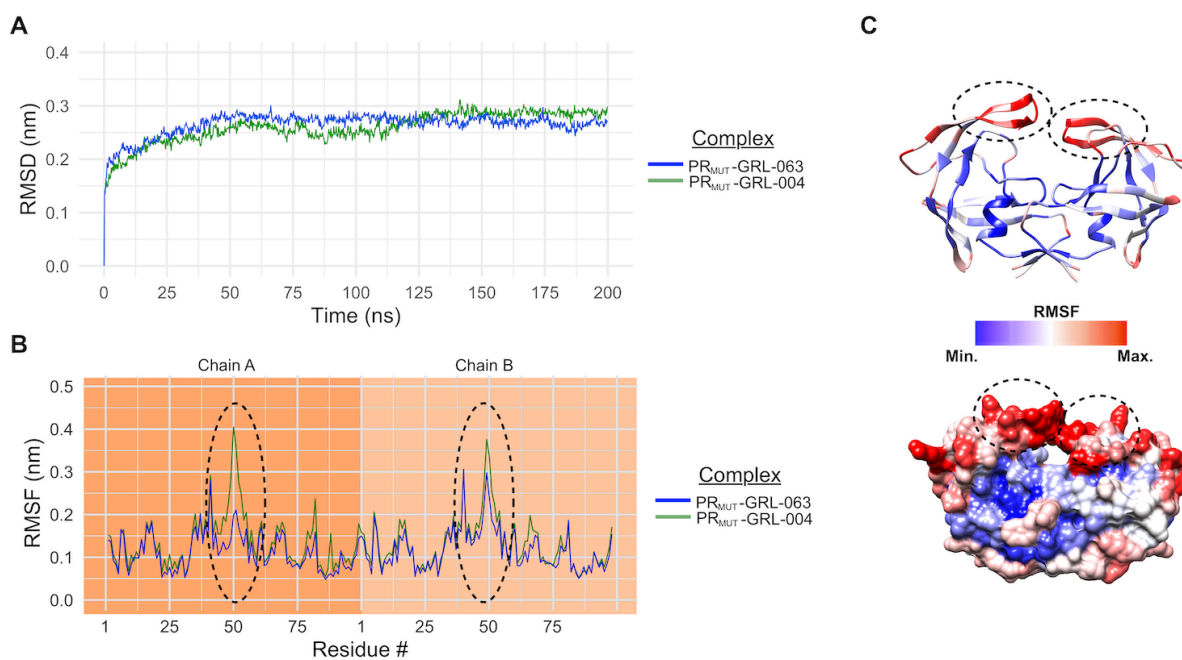
solvente, bem como os ajustes conformacionais do complexo proteína-ligante, simulações de DM foram realizadas. A Figura 3A mostra a evolução do RMSD médio ao longo da trajetória de 200 ns. Observa-se que ambos os complexos apresentaram similaridades em termos de comportamento na solução, com o complexo PR<sub>MUT</sub>-GRL-063 se estabilizando mais cedo (em torno de 50 ns), enquanto o complexo PR<sub>MUT</sub>-GRL-004 apenas após aproximadamente 112 ns, ambos com valores aproximados de 0.26 nm.

A média da análise RMSF (Figura 4B) revela uma região com grande flutuação em ambos os complexos das duas cadeias, destacada com uma linha tracejada preta. Esta região é composta por loops de folha-beta que englobam os sítios de ligação das moléculas (Figura 3C). É interessante notar que, após decorridos 150 ns da simulação, há uma redução da RMSF na região, indicando um aumento da estabilidade devido aos ajustes conformacionais dos ligantes nos sítios ativos (Apêndice A).

A análise MQ/MM usando a metodologia MQ-MM/GBSA indicou que as energias médias para ambos os complexos e todas as réplicas foram menores que 0, sugerindo a formação de complexos energeticamente favoráveis (Tabela 2). A conformação que obteve a menor energia do complexo GRL-063 foi da réplica 2, com um valor de -33.27 kcal/mol (Frame 1870). Para o complexo GRL-004, a melhor conformação foi obtida da simulação da réplica 1 com energia de -35.12 kcal/mol (Frame 1574) (Figura 4). É importante notar que, em ambos os complexos, houve um ajuste dos ligantes em seus respectivos sítios ativos, movendo-se para uma região mais interna do sítio (Figura 4), se comparado às posições da ancoragem (Figura 2). Estes resultados mostram que as simulações de DM alcançaram com sucesso seus objetivos, com as conformações estando apropriadas para uso em pesquisas que busquem analisar suas interações energéticas sob a ótica da MQ.

Figura 3 - Análise das simulações de dinâmica molecular. A. RMSD médio das três réplicas dos complexos PR<sub>MUT</sub>-GRL-063 e PR<sub>MUT</sub>-GRL-004 ao longo da trajetória de 200 ns. O gráfico RMSD de cada réplica pode ser visto nos Apêndices B-C. B. RMSF médio dos resíduos das três réplicas dos complexos PR<sub>MUT</sub>-GRL-063 e PR<sub>MUT</sub>-GRL-004. As cores de fundo laranja escuro e laranja claro separam os resíduos das cadeias A e B, respectivamente. O gráfico RMSF de cada réplica pode

ser visto nos Apêndices D-E. C. Representação da fita e superfície da análise estrutural RMSF do complexo PR<sub>MUT</sub>-GRL-004 pela réplica 1. Ambas representações estão coloridas de acordo com o RMSF, onde valores maiores estão em vermelho e valores menores em azul.



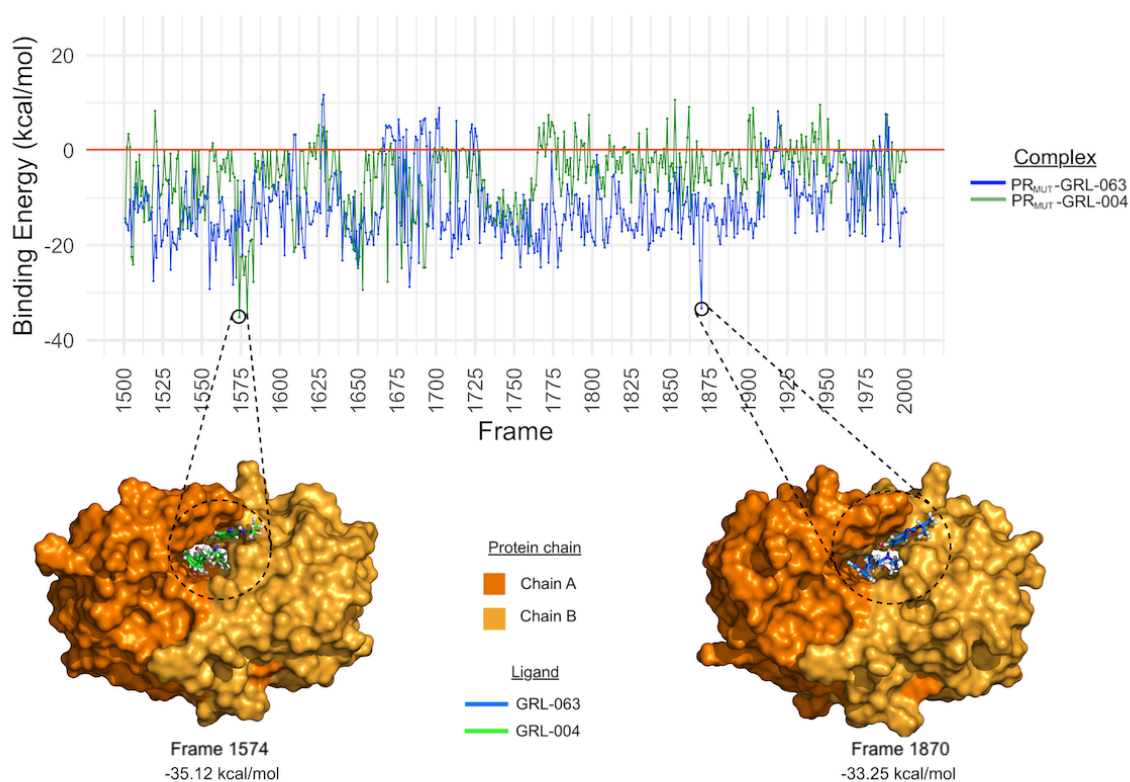
Fonte: Elaborado pelo autor, 2024.

Tabela 2 - Sumário dos resultados dos cálculos MQ/MM dos últimos 50 ns de cada réplica das simulações da DM.

| Lig.    | Réplica | Min.          | Q1 <sup>&amp;</sup> | Mediana | Média  | Q3 <sup>#</sup> | Máx.  | Dp. <sup>±</sup> |
|---------|---------|---------------|---------------------|---------|--------|-----------------|-------|------------------|
| GRL-063 | 1       | -29.20        | -15.33              | -11.73  | -10.68 | -6.77           | 15.22 | 7.58             |
|         | 2       | <b>-33.27</b> | -16.86              | -12.93  | -11.57 | -8.07           | 11.69 | 7.50             |
|         | 3       | -28.87        | -11.04              | -7.22   | -7.05  | -2.38           | 8.50  | 6.65             |
| GRL-004 | 1       | <b>-35.12</b> | -9.08               | -4.09   | -5.39  | -0.05           | 10.63 | 7.09             |
|         | 2       | -29.15        | -12.33              | -8.23   | -8.05  | -4.08           | 11.14 | 6.95             |
|         | 3       | -22.54        | -7.81               | -3.58   | -2.81  | 1.20            | 19.69 | 7.59             |

#1º Quartil; &3º Quartil; ±Desvio padrão.

Figura 4 - Análise MQ/MM dos últimos 500 modos conformacionais (50 ns da simulação da DM) dos complexos PR<sub>MUT</sub>-GRL-063 (réplica 2 da simulação de DM) e PR<sub>MUT</sub>-GRL-004 (réplica 1 da simulação de DM). O PR<sub>MUT</sub> está representado em uma superfície laranja, com seus respectivos ligantes em bastão azul (GRL-063) e bastão verde (GRL-004). A proteína está representada em uma superfície laranja, com laranja escuro representando a cadeia A e laranja claro representando a cadeia B. O gráfico MQ/MM para cada réplica individual pode ser visto nos Apêndices F-G.



Fonte: Elaborado pelo autor, 2024.

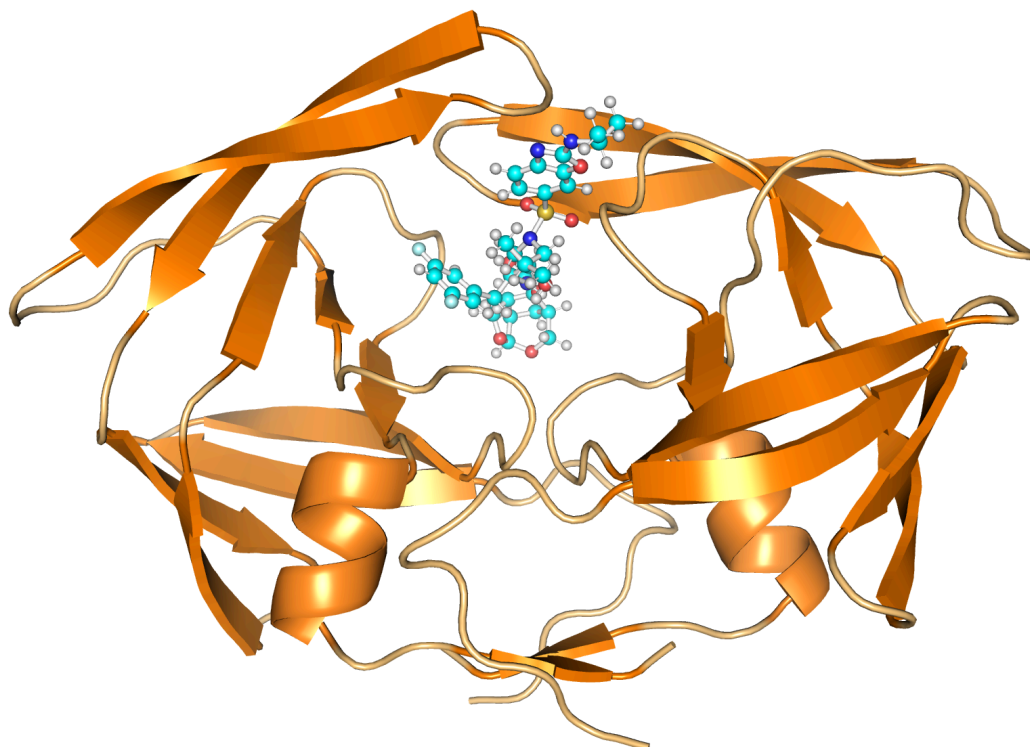
#### ○ 4.2 ANÁLISE DE CONVERGÊNCIA E BIRD

O ligante GRL-063 foi esquematicamente subdividido em três regiões distintas para facilitar a análise das interações proteína-ligante, como pode ser visto na Figura 5.



Figura 6.

Figura 6 - Ligante GRL-063 interagindo com a protease da cepa MUT do HIV-1.



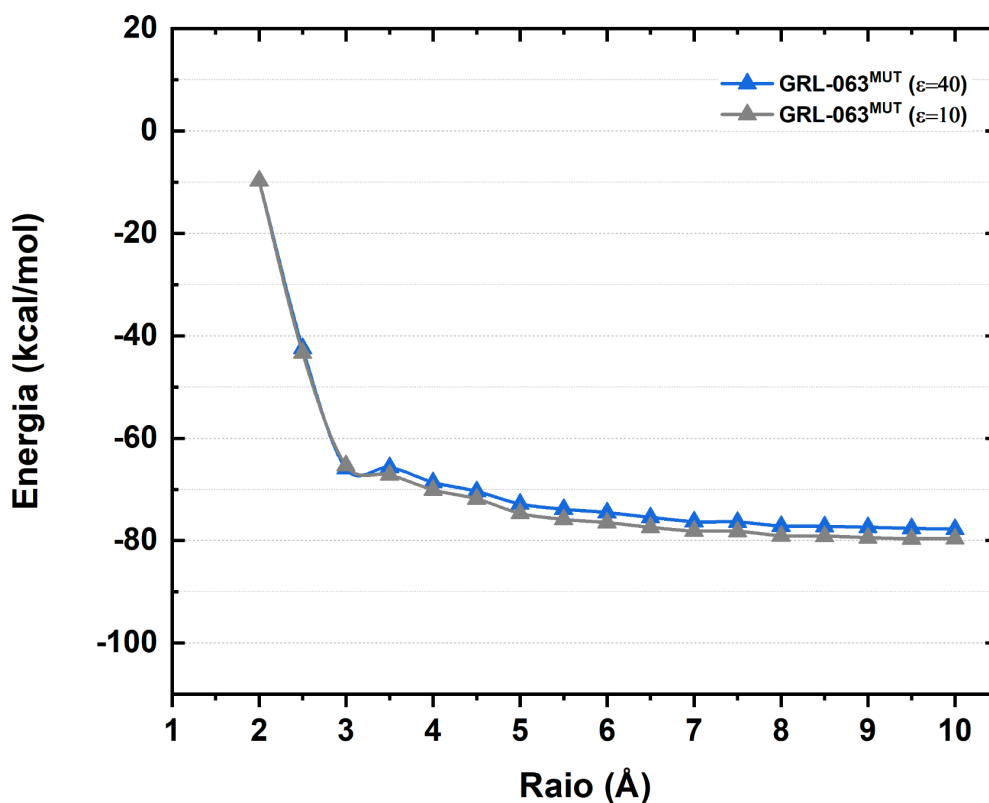
Fonte: Elaborado pelo autor, 2024.

De acordo com nosso critério de convergência, mostrado na Figura 7, começando em um raio de  $r = 2.0 \text{ \AA}$ , a convergência foi obtida a um raio de  $r = 3.5 \text{ \AA}$  no GRL-063 para a variante  $PR_{MUT}$  ( $-65.67 \text{ kcal mol}^{-1}$ ), representando uma variância menor que 10% comparado aos seus valores no raio  $r = 4.0 \text{ \AA}$ . Além disso, como mencionado previamente, adotamos duas constantes dielétricas distintas para representar o ambiente de solvatação e seus efeitos em ambos ligantes. Entretanto, de agora em diante, iremos limitar a discussão de nossos resultados para o valor da constante dielétrica  $\epsilon = 40$ , um melhor valor teórico para simulações de um estado fisiológico e bom meio termo entre o esperado experimentalmente e resultados computacionais (Dantas *et al.*, 2015; Ourique *et al.*, 2016; Amin *et al.*, 2020).

Após realizarmos os cálculos de energia, focamos nossa atenção para resíduos específicos que demonstraram as interações energéticas mais notáveis, seja atrativa ou repulsiva. Os principais resíduos do complexo  $PR_{MUT}$ -GRL-063, em

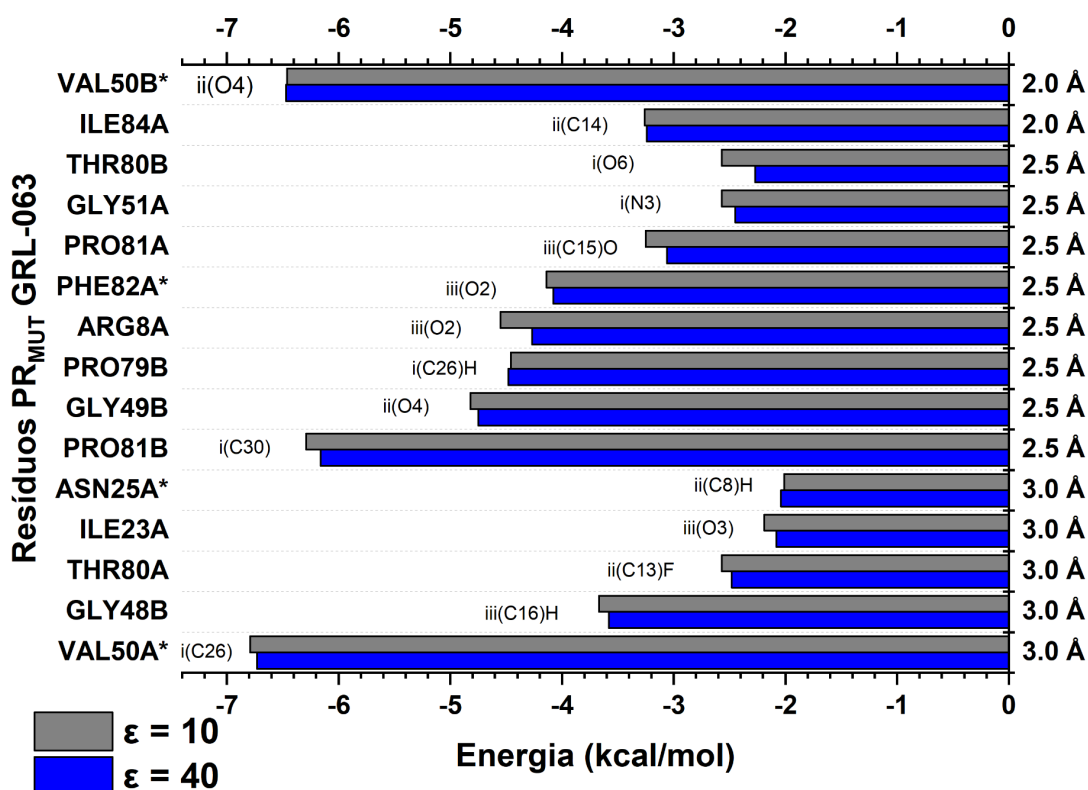
ordem de interações atrativas em que A denota o resíduo pertencente à cadeia A da protease do HIV-1 e B, para a cadeia B da protease (em kcal/mol): VAL50A\* (-6.73), VAL50B\* (-6.47), PRO81B (-6.16), GLY49B (-4.75), PRO79B (-4.48), ARG8A (-4.27), PHE82A\* (-4.08), GLY48' (-3.58), ILE84A (-3.24), PRO81A (-3.06), THR80A (-2.48), GLY51A (-2.45), THR80B (-2.27), ILE23A\* (-2.08), e ASN25A\* (-2.04). Resíduos denotados com \* indicam aminoácidos mutados se comparados a PR<sub>WT</sub> e apontam mutações adquiridas pela cepa resistente ao DRV. Não houve nenhuma repulsão notável nos resíduos da PR<sub>MUT</sub> no complexo proteína-ligante. Todos estes resíduos podem ser vistos na Figura 8.

Figura 7 - Gráfico de convergência do complexo PR<sub>MUT</sub>-GRL-063 com suas respectivas energias de interação (em kcal/mol) cumulativas para cada raio, nas constantes  $\epsilon = 10$  e  $\epsilon = 40$ . Cada curva representa uma constante dielétrica diferente. A linha cinza com triângulos indica valores para a constante  $\epsilon = 10$ , enquanto a linha azul com triângulos indica valores para a constante  $\epsilon = 40$ .



Fonte: Elaborado pelo autor, 2024.

Figura 8 - Gráfico BIRD (acrônimo em inglês para sítios de ligação, energias de interação, e domínio dos resíduos) mostrando em ordem crescente de raio (em Å) a os resíduos mais relevantes do complexo PR<sub>MUT</sub>-GRL-063 e suas energias em kcal/mol nas constantes dielétricas  $\epsilon = 10$  (em cinza) e  $\epsilon = 40$  (em azul escuro).



Fonte: Elaborado pelo autor, 2024.

Alguns destes resíduos merecem atenção especial devido a sua importância no funcionamento da protease bem como propriedades bem descritas na literatura. De acordo com Weber *et al.* (1998), alguns resíduos da protease desempenham papel fundamental na formação de interações de hidrogênio com os substratos. ARG8, GLY49, e PRO81 são alguns destes resíduos importantes que demonstraram energias de interação relevantes em nossos estudos. Além disso, Stebbins *et al.* (1997) enfatizou a importância do *loop* dos 80 na protease do HIV-1, uma região composta pelos resíduos 78 a 85 que agem na região S<sub>1</sub> do sítio ativo de ligação e modificam sua especificidade com os substratos, permitindo melhor acomodação de estruturas importantes (Stebbins *et al.*, 1997; Kuiper *et al.*, 2015). Os principais

resíduos da PR<sub>MUT</sub> interagindo com o GRL-063 podem ser vistos na Figura 8.

#### ○ 4.3 REGIÃO I DO COMPLEXO PR<sub>MUT</sub>-GRL-063

Região I é notável devido a abundância de interações hidrofóbicas, com os resíduos VAL50A\* (i(C26)) e PRO81B (i(C30)), mostrados na Figura 9A e Figura 9B respectivamente, formando diferentes tipos desta interação. Esta região engloba um grupo amina e o composto heterocíclico isobenzofurano, compondo a totalidade da região P2'. A região I se destaca pelo seu número de interações altamente energéticas, a exemplo dos resíduos mencionados e outros pertencentes ao *loop* dos 80 descrito anteriormente, como a PRO79B (Figura 9A), THR80B e PRO81B (Figura 9B).

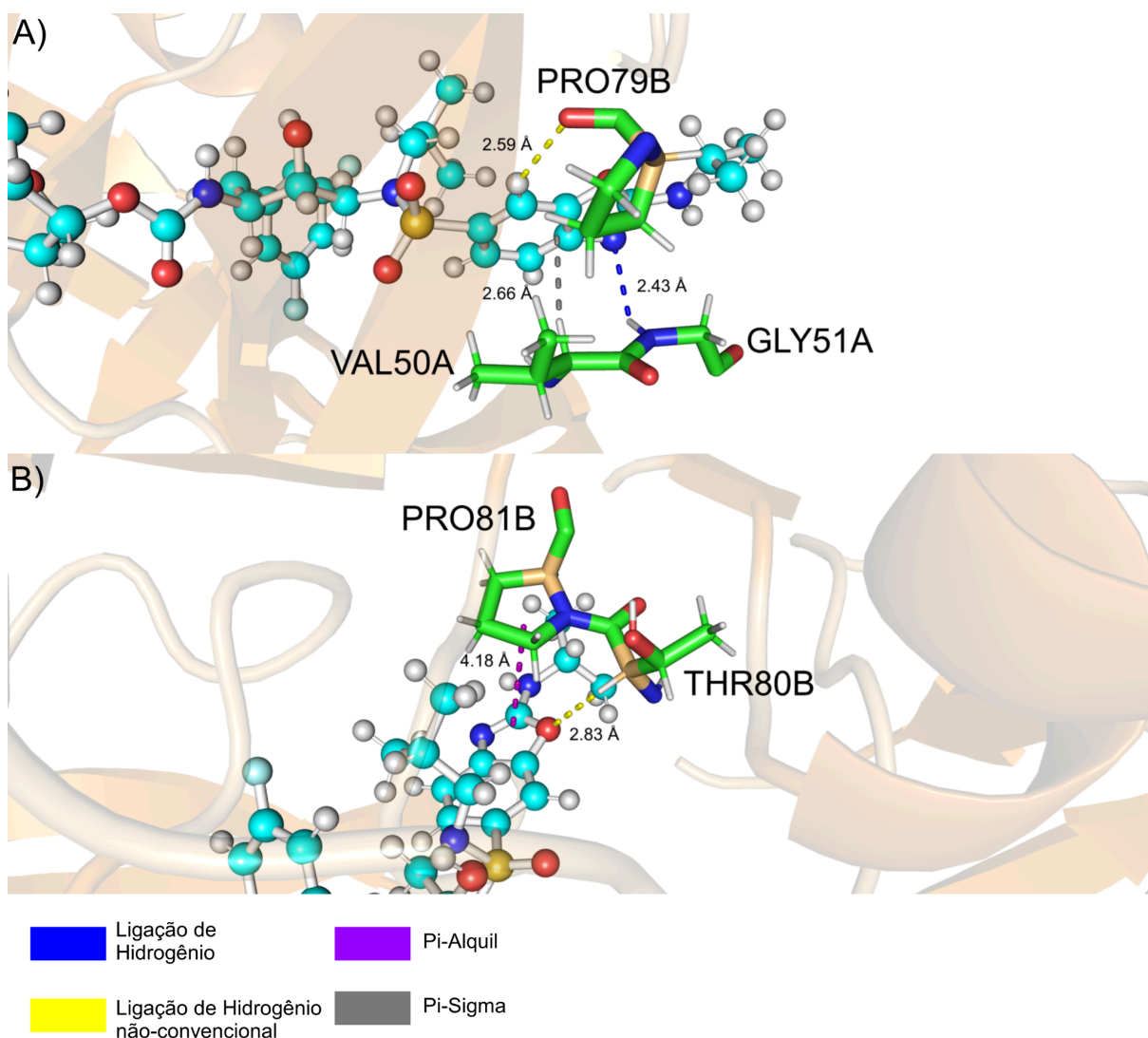
Destacamos na PR<sub>MUT</sub> os resíduos VAL50\*, com a maior energia atrativa na análise do complexo proteína-ligante do GRL-063, PRO81B (i(C30)) com a terceira maior, e THR80B (i(O6)), cuja interação com a ILE50 na cepa selvagem (e VAL50 na mutante, a exemplo do nosso trabalho) permite um maior grau de mobilidade nas extremidades da alça da protease (Liu *et al.*, 2008). Estes resíduos interagem com o GRL-063 a distâncias de 2.66, 4.18, e 2.83 Å, respectivamente.

Como mencionado acima, o resíduo VAL50A\* foi um dos resíduos mais energéticos do ligante GRL-063, apresentando energia de -6.73 kcal/mol. Este é um resíduo chave na ligação com o substrato ao interagir com THR80, outro aminoácido abordado em nossa análise, e modificando a mobilidade da alça da protease. Ele interage com o GRL-063 através de uma ligação hidrofóbica do tipo pi-sigma no segundo anel aromático da região (Figura 9A). Já em relação a PRO81B, este resíduo foi um dos mais atrativos de toda a estrutura, com uma energia de -6.61 kcal/mol. Ele se liga com o GRL-063 através de outra interação hidrofóbica pi-alquil entre o resíduo e o anel aromático do ligante (Figura 9B). Por fim, THR80B forma uma ligação de hidrogênio não-convencional (Figura 9B) com uma energia de -2.27 kcal/mol. Isto é especialmente notável devido a importância da THR80 no controle da mobilidade da alça da protease, uma região em contato direto com o sítio ativo da proteína e apresentando papel crucial nas interações com substratos (Liu *et al.*, 2008; Yu *et al.*, 2017).

Importante notar a presença de interações de hidrogênio nesta região, a exemplo da GLY51A (-2.45 kcal/mol) (Figura 9A), bem como outras interações

hidrofóbicas do tipo alquil, como a PRO81B (-6.16 kcal/mol) (Figura 9B). Estas descobertas estão de acordo com os objetivos da pesquisa original e achados da literatura (Bulut *et al.*, 2020; Nalam *et al.*, 2010; Shen *et al.*, 2013; Prabu-Jeyabalan, Nalivaika, Schiffer, 2002).

Figura 10 - Alguns dos resíduos mais importantes da PR<sub>MUT</sub> interagindo com o ligante GRL-063 na região I.



Fonte: Elaborado pelo autor, 2024.

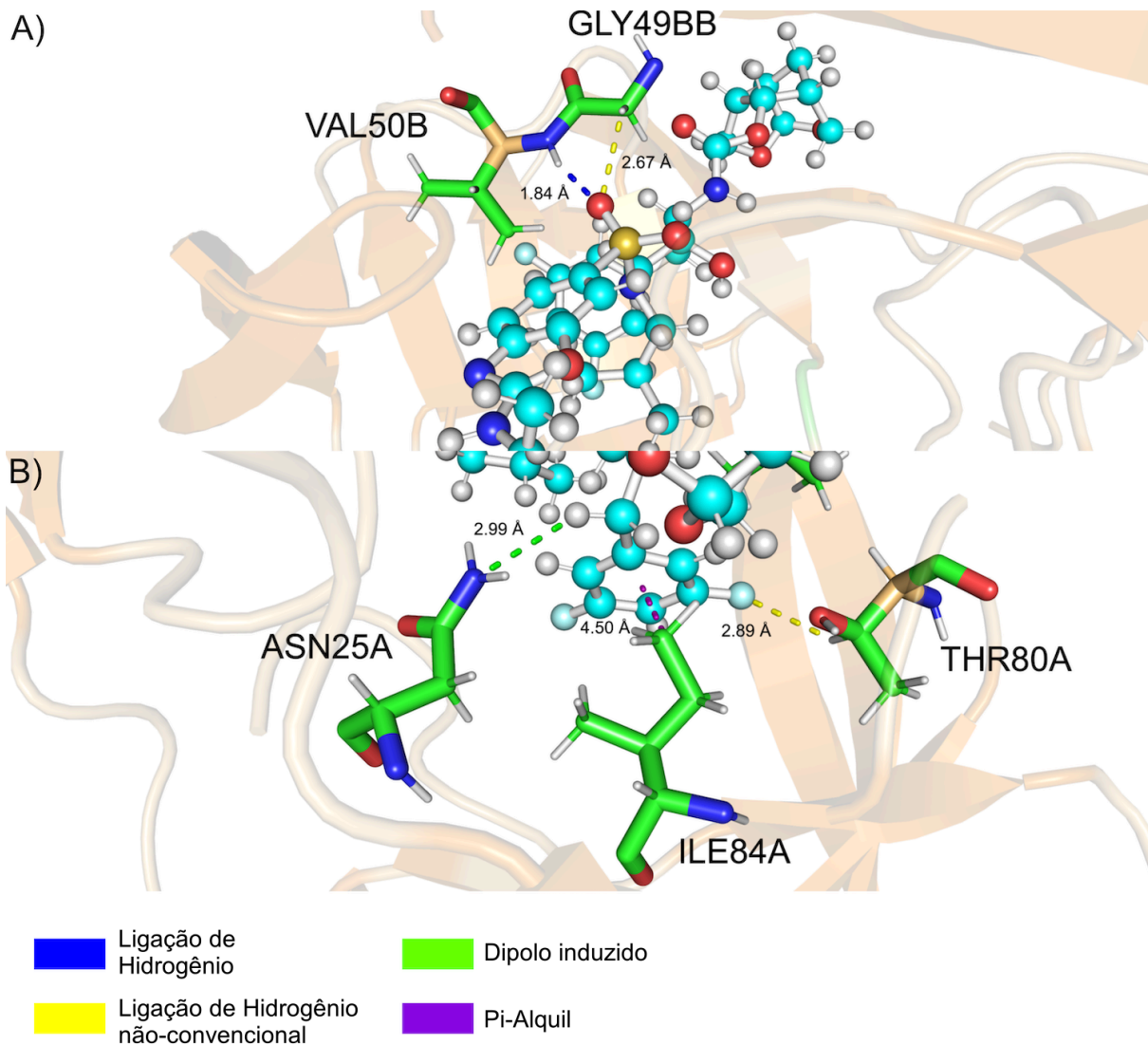
#### ○ 4.4 REGIÃO II DO COMPLEXO PR<sub>MUT</sub>-GRL-063

A região II se destaca pela presença de dois resíduos mutados com grande importância no funcionamento da cepa selvagem da protease. O principal resíduo

desta região no complexo PR<sub>MUT</sub>-GRL-063 é a VAL50B\* (ii(O4)) (Figura 10A). Em sua cepa não-mutada, este resíduo se dispõe sob forma de uma isoleucina (ILE50), desempenhando um importante papel na modelagem das extremidades da alça da protease e permite interações dos substratos com os principais sítios ativos da proteína. Este resíduo é um dos principais alvos para IPs alternativos que alvejam não os sítios ativos mas outras regiões na proteína (Weber, Kneller, Wong-Sam, 2015; Zhang Y. *et al.*, 2014). Entretanto, a nossa análise identificou uma mutação I50V neste aminoácido, que apresentou uma grande energia atrativa de -6.47 kcal/mol a uma distância de 1.84 Å.

Outro aminoácido de destaque na região II é o ASN25A\* (ii(C8)H) (Figura 10B). Em sua forma selvagem, este resíduo se encontra como uma ASP e faz parte do principal sítio ativo da protease, a tríade catalítica formada pelos resíduos ASP25-THR26-GLY27, uma sequência de aminoácidos presentes em diversas proteases aspárticas (a exemplo do HIV-1) e desempenham importante papel na clivagem de substratos (Short *et al.*, 2000; Chatterjee *et al.*, 2005; Johnson *et al.*, 2007). Na nossa análise, este aminoácido apresentou uma energia atrativa de -2.04 kcal/mol, formando uma ligação não-convencional de hidrogênio (Figura 10B) a distância de 2.99 Å.

Figura 10 - Alguns dos resíduos mais importantes da PR<sub>MUT</sub> interagindo com o ligante GRL-063 na região II.



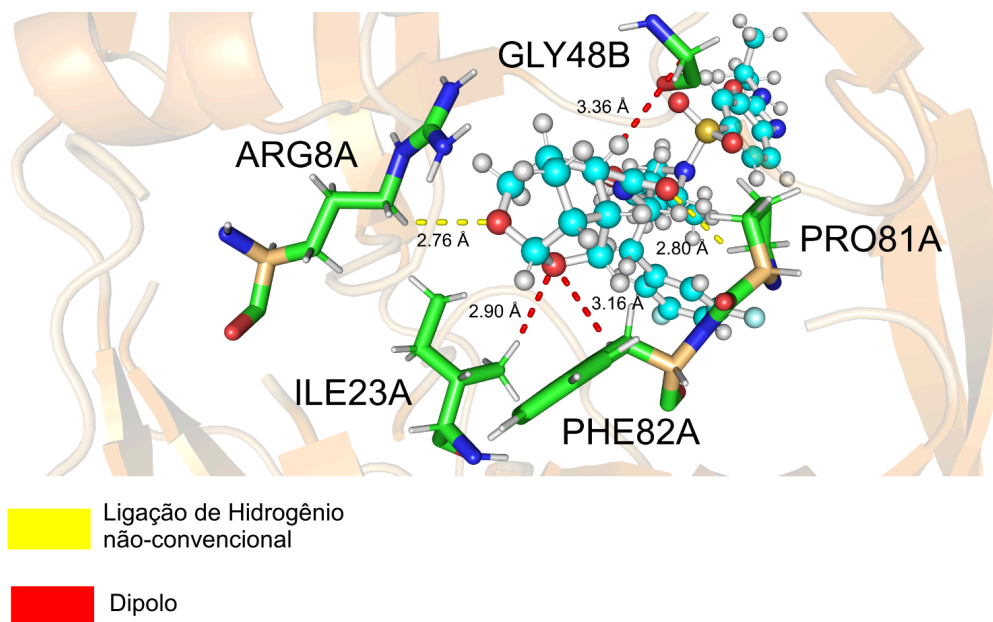
Fonte: Elaborado pelo autor, 2024.

#### ○ 4.5 REGIÃO III DO COMPLEXO PR<sub>MUT</sub>-GRL-063

A região III não apresenta resíduos muito energeticamente relevantes, dada a predominância de interações dipolo e ligações de hidrogênio não-convencionais, com grande parte de suas interações se mantendo dentro do *loop* dos 80: PHE82A\* (-4.08) e PRO81A (-3.06) (Figura 11) com o primeiro sendo resultado de uma mutação V82F; adicionalmente, ARG8A (-4.27) se apresenta como um resíduo altamente conservado entre proteases do HIV-1 (Figura 11B) (Ho *et al.*, 1998).

GLY48B (-3.58) e ILE23A\* (-2.08) são outros dois resíduos que interagem com essa região (Figura 11).

Figura 11 - Alguns dos resíduos mais importantes da PR<sub>MUT</sub> interagindo com o ligante GRL-063 na região III.



Fonte: Elaborado pelo autor, 2024.

## ● 5 CONCLUSÃO

O panorama global da infecção por HIV e a necessidade de considerar não apenas as cepas mais comuns, mas também as inúmeras cepas mutadas, justificam a busca constante por opções terapêuticas novas e eficientes. Uma dessas alternativas é o DRV, um dos mais recentes medicamentos inibidores da protease aprovados pela FDA capaz de tratar cepas mutadas altamente resistentes. No entanto, casos de HIV resistente ao DRV já foram relatados, demonstrando a capacidade do vírus de adquirir resistência mesmo às opções mais eficazes atualmente disponíveis. Neste contexto, os IPs experimentais GRL-004 e GRL-063 foram projetados como alternativas para tratar não apenas variantes PR<sub>WT</sub> do HIV-1, mas também mutadas que adquiriram resistência ao DRV. Portanto, entender seus mecanismos de ação e suas energias de ligação à protease viral é fundamental para descrever como tais medicamentos podem tratar adequadamente a doença.

Nossas descrições energéticas e parâmetros de simulação usados para descrever as energias de interação entre cada complexo proteína-ligante tem como base o uso de ferramentas *ab initio* cujo uso se encontra extensivamente descrito na literatura e estão de acordo com estudos experimentais anteriores e análises computacionais similares com outras estruturas cristalográficas. Através de simulações computacionais e cálculos *ab initio*, conseguimos montar duas estruturas representando cada ligante interagindo com a PR<sub>MUT</sub> do HIV-1 em sua forma mais otimizada, e prosseguimos com a descrição energética de uma destas estruturas (complexo PR<sub>MUT</sub>-GRL-063) para apresentar não só a utilização da metodologia do MFCC aliado ao DFT, mas também demonstrar que ambos complexos se encontram prontos para análises energéticas futuras mais robustas e aprofundadas. Nossas descrições energéticas revelaram a presença de aminoácidos cruciais interagindo dentro do complexo PR<sub>MUT</sub>-GRL-063, em especial VAL50A\*, VAL50B\*, e PRO81B como os resíduos mais energéticos de nossa análise; além de resíduos que, em sua forma selvagem, desempenham importante papel no funcionamento da protease viral e aos quais o ligante GRL-063 obteve sucesso em manter interações relevantes, a exemplo do ASN25A\* e PHE82A\*. A região I se destacou como a mais energeticamente favorável em nosso estudo, seguido pela região II; entretanto, a região III se demonstrou como a área de menor afinidade a proteína, muito pelo

predomínio de interações dipolo em contraste com ligações de hidrogênio mais fortes encontradas nas outras regiões.

A metodologia adotada apresenta resultados promissores, enfrentando como principal limitação o custo computacional elevado das ferramentas envolvidas, além da dificuldade inerente de se tratar áreas maiores sob a ótica da MQ. Contudo, o êxito obtido nas simulações demonstra a viabilidade de sua aplicação para simplificar a análise de uma ampla variedade de compostos farmacológicos, permitindo a complementação de análises *in vitro* com ferramentas computacionais precisas. Os complexos gerados pela nossa ancoragem e DM mostram-se otimizados para trabalhos futuros, especialmente através da utilização da metodologia do MFCC em conjunto com o funcional do DFT, como evidenciado pela análise do PR<sub>MUT</sub>-GRL-063, que elucidou os principais resíduos atrativos. Dessa forma, abre-se espaço para o desenvolvimento de pesquisas futuras utilizando esses dois complexos, não apenas em suas cepas mutantes, mas também em estudos comparativos com suas versões selvagens. Além disso, esperamos que nossa análise contribua para o entendimento do funcionamento dos novos inibidores de protease e sua interação com a protease do HIV-1. Também destacamos o potencial da MQ e das simulações computacionais como uma alternativa valiosa para o estudo e o design de novos fármacos, permitindo análises aprofundadas que considerem não só a farmacodinâmica, mas também, por meio de outras ferramentas *ab initio*, um estudo abrangente das características farmacocinéticas dos compostos.

## REFERÊNCIAS

- ABRAHAM, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, set. 2015. v. 1-2, p. 19–25.
- AHUJA, S. *et al.* Molecular Docking; future of Medicinal Research. **Ecology, Environment and Conservation**, 31 jan. 2022. v. 28, n. 01s, p. 18–18. Acesso em: 16 out. 2022.
- ALBUQUERQUE, E. L. *et al.* **Quantum Chemistry Simulation of Biological Molecules**. [S.l.]: Cambridge University Press , 2020.
- AMIN, M.; JOCHEN KÜPPER. Variations in Proteins Dielectric Constants. **ChemistryOpen**, 1 jun. 2020. v. 9, n. 6, p. 691–694.
- AOKI, M. *et al.* A novel central nervous system-penetrating protease inhibitor overcomes human immunodeficiency virus 1 resistance with unprecedented aM to pM potency. **eLife**, 17 out. 2017. v. 6.
- \_\_\_\_\_ *et al.* Mechanism of Darunavir (DRV)'s High Genetic Barrier to HIV-1 Resistance: A Key V32I Substitution in Protease Rarely Occurs, but Once It Occurs, It Predisposes HIV-1 To Develop DRV Resistance. **mBio**, 2 maio. 2018. v. 9, n. 2.
- ARGAMAN, N.; MAKOV, G. Density functional theory: An introduction. **American Journal of Physics**, jan. 2000. v. 68, n. 1, p. 69–79. Acesso em: 16 set. 2020.
- BARTLETT, R. J. *Ab initio*DFT and its role in electronic structure theory. **Molecular Physics**, 10 nov. 2010. v. 108, n. 21-23, p. 3299–3311.
- BERENDSEN, H. J. C. *et al.* Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, 15 out. 1984. v. 81, n. 8, p. 3684–3690.
- BRODER, S. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. **Antiviral Research**, jan. 2010. v. 85, n. 1, p. 1–18.

BULUT, H. *et al.* Single atom changes in newly synthesized HIV protease inhibitors reveal structural basis for extreme affinity, high genetic barrier, and adaptation to the HIV protease plasticity. **Scientific Reports**, 30 jun. 2020. v. 10, n. 1.

CAMILA, A. *et al.* In Silico Evaluation of the Binding Energies of Androgen Receptor Agonists in Wild-Type and Mutational Models. **The Journal of physical chemistry. B**, 29 maio. 2023. v. 127, n. 22, p. 5005–5017.

CAMPOS, D. M. O. *et al.* Intermolecular interactions of cn-716 and acyl-KR-aldehyde dipeptide inhibitors against Zika virus. **Physical Chemistry Chemical Physics**, 2020. v. 22, n. 27, p. 15683–15695.

CARLONI, P.; ROTH LISBERGER, U.; PARRINELLO, M. The Role and Perspective of Ab Initio Molecular Dynamics in the Study of Biological Systems. **Accounts of Chemical Research**, 15 maio. 2002. v. 35, n. 6, p. 455–464.

CHATTERJEE, A. *et al.* Folding Regulates Autoprocessing of HIV-1 Protease Precursor. **Journal of Biological Chemistry**, 4 jan. 2005. v. 280, n. 12, p. 11369–11378.

CHEN, L. *et al.* Performance of DFT functionals for properties of small molecules containing beryllium, tungsten and hydrogen. **Nuclear materials and energy**, 1 jan. 2020. v. 22, p. 100731–100731.

CHIVIAN, D. *et al.* Ab Initio Methods. **Methods of biochemical analysis**, 10 fev. 2003. p. 547–557.

DIEGO *et al.* Quantum molecular modelling of ibuprofen bound to human serum albumin. **RSC Advances**, 1 jan. 2015. v. 5, n. 61, p. 49439–49450.

FANALES-BELASIO, E. *et al.* HIV virology and pathogenetic mechanisms of infection: a brief overview. **Annali dell'Istituto Superiore di Sanità**, 2009. v. 46, n. 1, p. 5–14.

FARADY, C. J.; CRAIK, C. S. Mechanisms of Macromolecular Protease Inhibitors. **ChemBioChem**, 4 nov. 2010. v. 11, n. 17, p. 2341–2346.

HALPERIN, I. *et al.* Principles of docking: An overview of search algorithms and a guide to scoring functions. **Proteins: Structure, Function, and Genetics**, 2 maio. 2002. v. 47, n. 4, p. 409–443.

HARDIN, Corey; POGORELOV, Taras V.; LUTHEY-SCHULTEN, Zaida. Ab initio protein structure prediction. *Current opinion in structural biology*, v. 12, n. 2, p. 176-181, 2002.

HESS, B. *et al.* LINCS: A linear constraint solver for molecular simulations. **Journal of Computational Chemistry**, set. 1997. v. 18, n. 12, p. 1463–1472.

HO, D. D. *et al.* Characterization of human immunodeficiency virus type 1 variants with increased resistance to a C2-symmetric protease inhibitor. **Journal of Virology**, 1 mar. 1994. v. 68, n. 3, p. 2016–2020. Acesso em: 13 nov. 2023.

HOCKNEY, R. W.; GOEL, S. P.; EASTWOOD, J. W. Quiet High-Resolution Computer Models of a Plasma. **JCoPh**, 1 fev. 1974. v. 14, n. 2, p. 148–158.

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

HUTTER, J. Car-Parrinello molecular dynamics. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, 15 set. 2011. v. 2, n. 4, p. 604–612.

ISOM, D. G. *et al.* Charges in the hydrophobic interior of proteins. **Proceedings of the National Academy of Sciences**, 26 ago. 2010. v. 107, n. 37, p. 16096–16100.

JOHNSON, E. C. B. *et al.* Insights from Atomic-Resolution X-Ray Structures of Chemically Synthesized HIV-1 Protease in Complex with Inhibitors. **Journal of Molecular Biology**, out. 2007. v. 373, n. 3, p. 573–586.

KARPLUS, M.; MCCAMMON, J. A. Molecular dynamics simulations of biomolecules. **Nature Structural Biology**, set. 2002. v. 9, n. 9, p. 646–652.

KUIPER, B. D. *et al.* The L33F darunavir resistance mutation acts as a molecular anchor reducing the flexibility of the HIV-1 protease 30s and 80s loops. **Biochemistry and Biophysics Reports**, 1 jul. 2015. v. 2, p. 160–165.

LIU, F. *et al.* Effect of Flap Mutations on Structure of HIV-1 Protease and Inhibition by Saquinavir and Darunavir. **Journal of molecular biology**, 1 ago. 2008. v. 381, n. 1, p. 102–115.

MAZUREK, A.; ŁUKASZ SZELESZCZUK; DARIUSZ MACIEJ PISKLAK. Periodic DFT Calculations—Review of Applications in the Pharmaceutical Sciences. **Pharmaceutics**, 1 maio. 2020. v. 12, n. 5, p. 415–415.

MEEKO. Meeko: preparation of small molecules for AutoDock. **GitHub**, maio. 2022. Disponível em: <<https://github.com/forlilab/Meeko>>. Acesso em: 26 jul. 2024.

MENEZES, G. De L. **Prospecção de inibidores da formação do dímero da NS1 de três principais Flavivirus de importância médica : uma abordagem por dinâmica molecular e triagem virtual de compostos**. Universidade Federal de Goiás: [s.n.], 2021. Dissertação.

MGLTools 1.5.7. **MGLTools**, 2022. Disponível em: <<http://mgltools.scripps.edu/>>. Acesso em: 26 jul. 2024.

NALAM, M. N. L. *et al.* Evaluating the Substrate-Envelope Hypothesis: Structural Analysis of Novel HIV-1 Protease Inhibitors Designed To Be Robust against Drug Resistance. **Journal of Virology**, 15 maio. 2010. v. 84, n. 10, p. 5368–5378.

NASTRI, B. M. *et al.* HIV and Drug-Resistant Subtypes. **Microorganisms**, 15 jan. 2023. v. 11, n. 1, p. 221.

ORIO, M.; PANTAZIS, D. A.; NEESE, F. Density functional theory. **Photosynthesis Research**, 24 fev. 2009. v. 102, n. 2-3, p. 443–453.

OURIQUE, G. S. *et al.* A quantum chemistry investigation of a potential inhibitory drug against the dengue virus. **RSC Advances**, 2016. v. 6, n. 61, p. 56562–56570.

OVALLE, S.; MALARDIER-JUGROOT, C. Choice of functional for iron porphyrin density functional theory studies: Geometry, spin-state, and binding energy analysis. **Computational and Theoretical Chemistry**, jul. 2022. v. 1213, p. 113726.

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

PINZI, L.; RASTELLI, G. Molecular Docking: Shifting Paradigms in Drug Discovery. **International Journal of Molecular Sciences**, 4 set. 2019. v. 20, n. 18, p. 4331.

PRABU-JEYABALAN, M.; NALIVAICA, E.; SCHIFFER, C. A. Substrate Shape Determines Specificity of Recognition for HIV-1 Protease. **Structure**, mar. 2002. v. 10, n. 3, p. 369–381.

SALMASO, V.; MORO, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. **Frontiers in Pharmacology**, 22 ago. 2018. v. 9.

SCHNEIDER, R.; SHARMA, A. R.; RAI, A. Introduction to Molecular Dynamics. **Computational Many-Particle Physics**, 2008. p. 3–40.

SHEN, Y. *et al.* Testing the Substrate-Envelope Hypothesis with Designed Pairs of Compounds. **ACS Chemical Biology**, 26 set. 2013. v. 8, n. 11, p. 2433–2441.

SHORT, G. F. *et al.* Probing the S1/S1' Substrate Binding Pocket Geometry of HIV-1 Protease with Modified Aspartic Acid Analogues. **Biochemistry**, 6 jul. 2000. v. 39, n. 30, p. 8768–8781.

SOUSA DA SILVA, A. W.; VRANKEN, W. F. ACPYPE - AnteChamber PYthon Parser interface. **BMC Research Notes**, 2012. v. 5, n. 1, p. 367.

STEBBINS, J. *et al.* The 80's loop (residues 78 to 85) is important for the differential activity of retroviral proteases 1 Edited by A. R. Fresht. **Journal of Molecular Biology**, 1 abr. 1997. v. 267, n. 3, p. 467–475.

TROTT, O.; OLSON, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring function, Efficient optimization, and Multithreading. **Journal of Computational Chemistry**, 2009. v. 31, n. 2.

TSUZUKI, S.; UCHIMARU, T. Accuracy of intermolecular interaction energies, particularly those of hetero-atom containing molecules obtained by DFT calculations

with Grimme's D2, D3 and D3BJ dispersion corrections. **Physical Chemistry Chemical Physics**, 2020. v. 22, n. 39, p. 22508–22519.

VALDÉS-TRESANCO, M. S. *et al.* gmx\_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. **Journal of Chemical Theory and Computation**, 29 set. 2021. v. 17, n. 10, p. 6281–6291.

VIANNA, J. F. *et al.* New ethionamide boosters and EthR2: structural and energetic analysis. **Physical Chemistry Chemical Physics**, 20 out. 2021. v. 23, n. 40, p. 23233–23241.

WEBER, I. T.; KNELLER, D. W.; WONG-SAM, A. Highly resistant HIV-1 proteases and strategies for their inhibition. **Future medicinal chemistry**, 1 jun. 2015. v. 7, n. 8, p. 1023–1038.

YANG, M. *et al.* Uni-GBSA: an open-source and web-based automatic workflow to perform MM/GB(PB)SA calculations for virtual screening. **Briefings in bioinformatics**, 16 jun. 2023. v. 24, n. 4.

YARCHOAN, R. *et al.* In Vivo Activity Against HIV and Favorable Toxicity Profile of 2',3'-Dideoxyinosine. **Science**, 28 jul. 1989. v. 245, n. 4916, p. 412–415.

YU, Y. *et al.* Structural insights into HIV-1 protease flap opening processes and key intermediates. **RSC Advances**, 2017. v. 7, n. 71, p. 45121–45128.

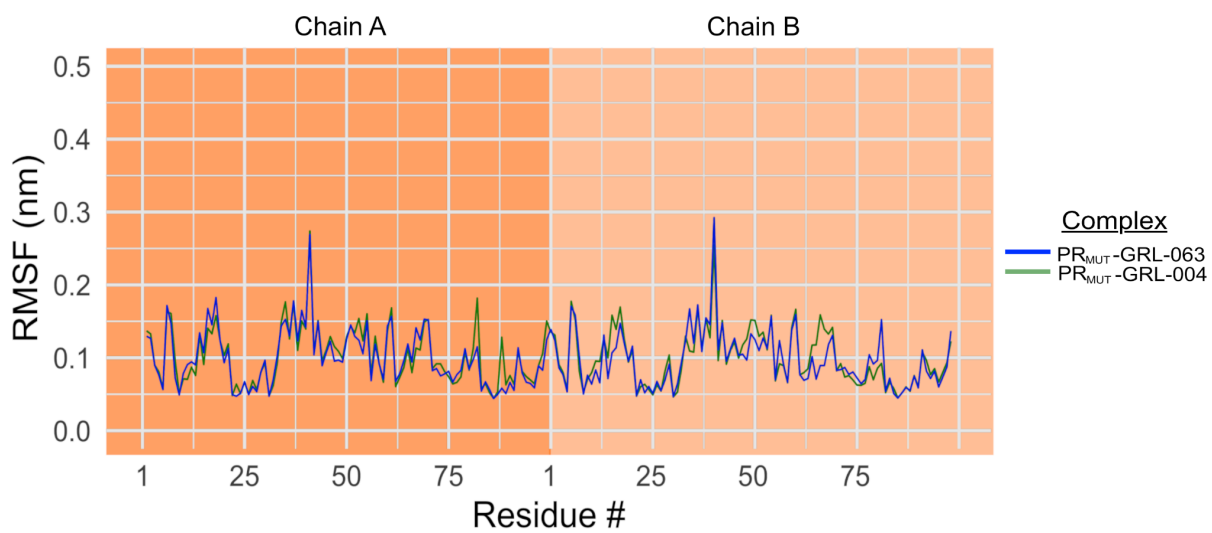
ZEPHYR, J.; KURT YILMAZ, N.; SCHIFFER, C. A. Viral proteases: Structure, mechanism and inhibition. **The Enzymes**, 2021. v. 50, p. 301–333.

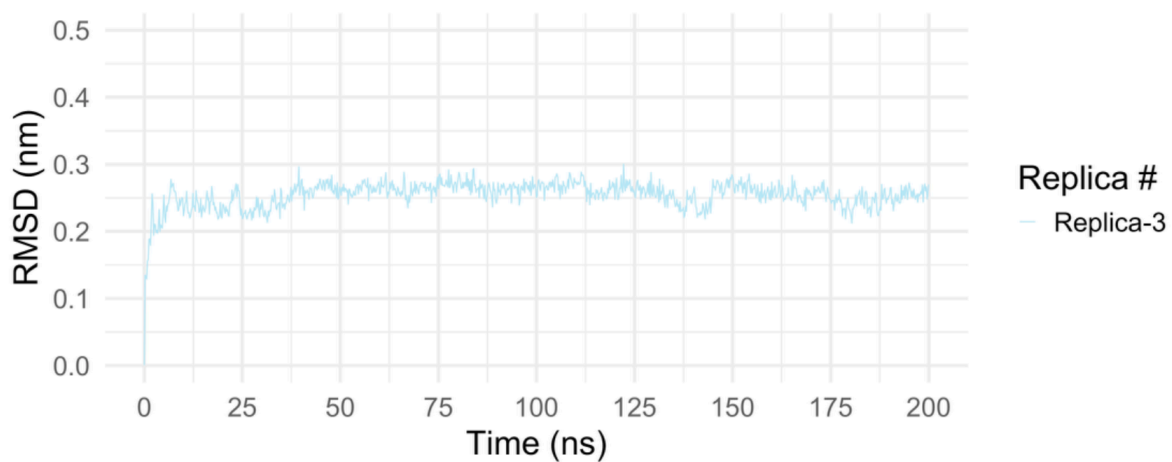
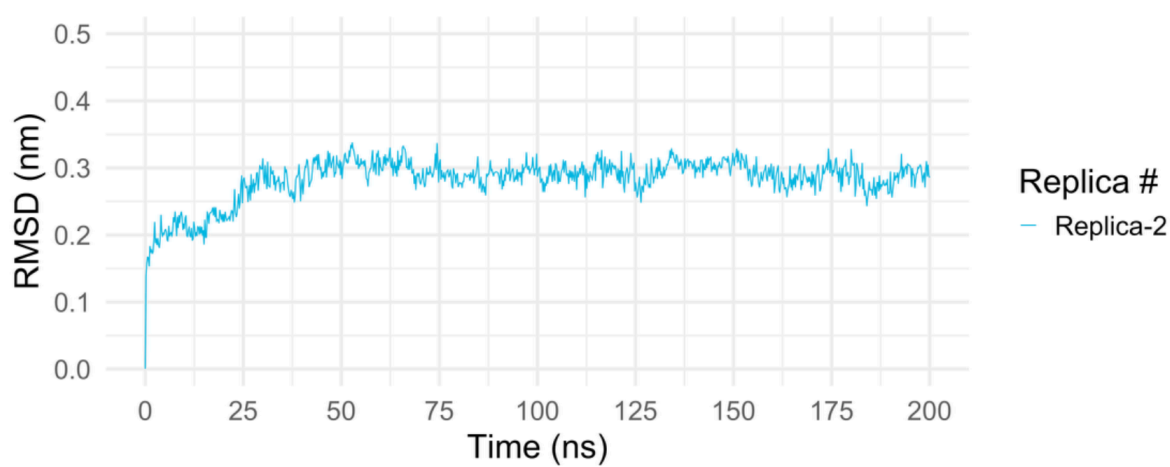
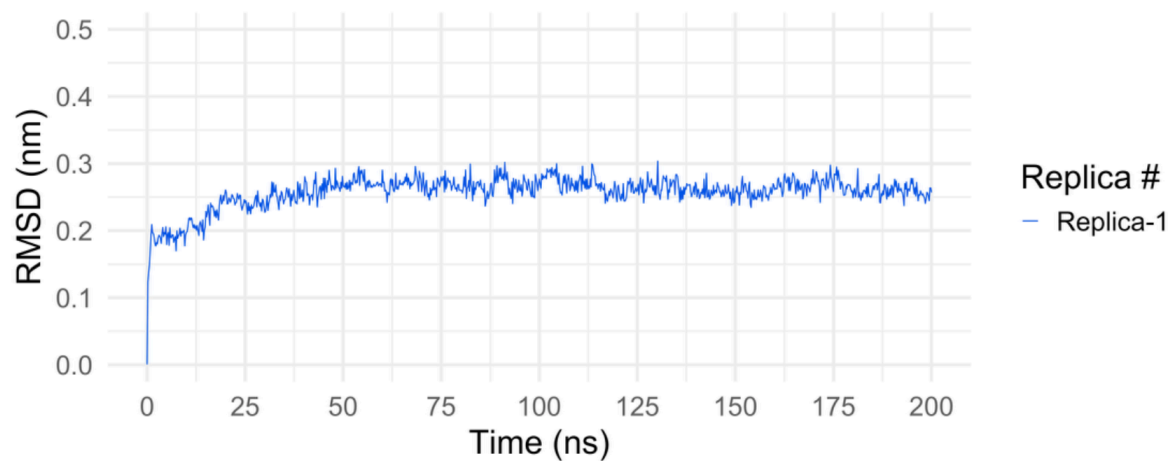
ZHANG, D. W.; ZHANG, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy. **The Journal of Chemical Physics**, 15 ago. 2003. v. 119, n. 7, p. 3599–3605.

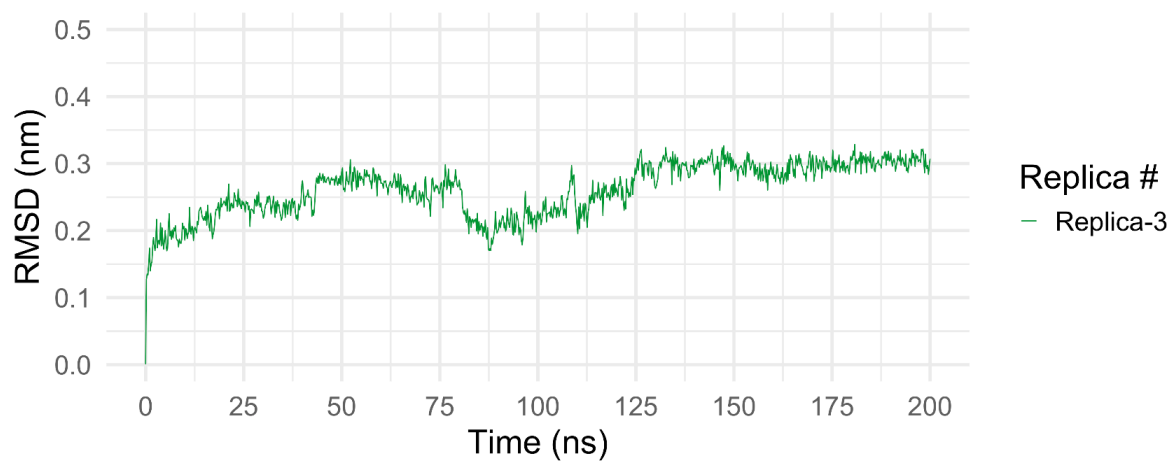
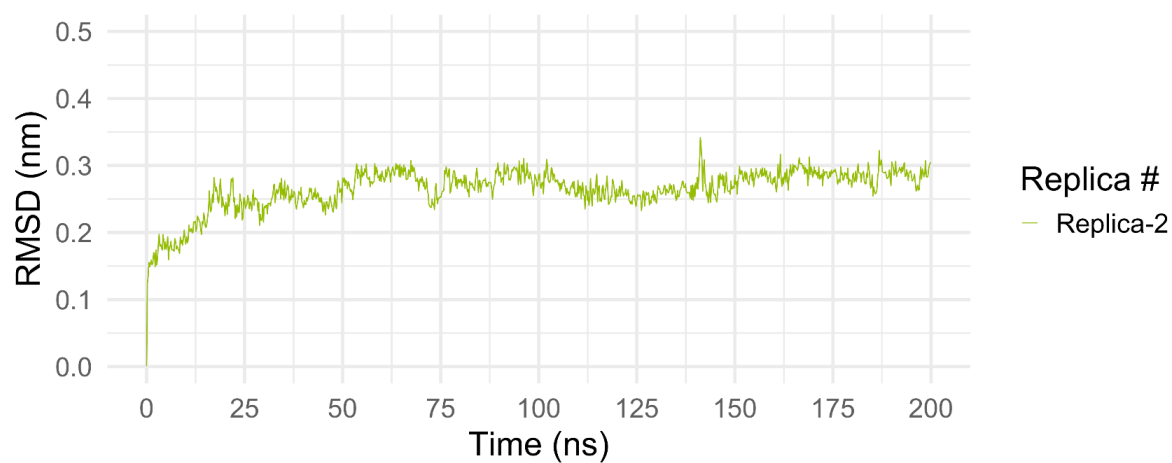
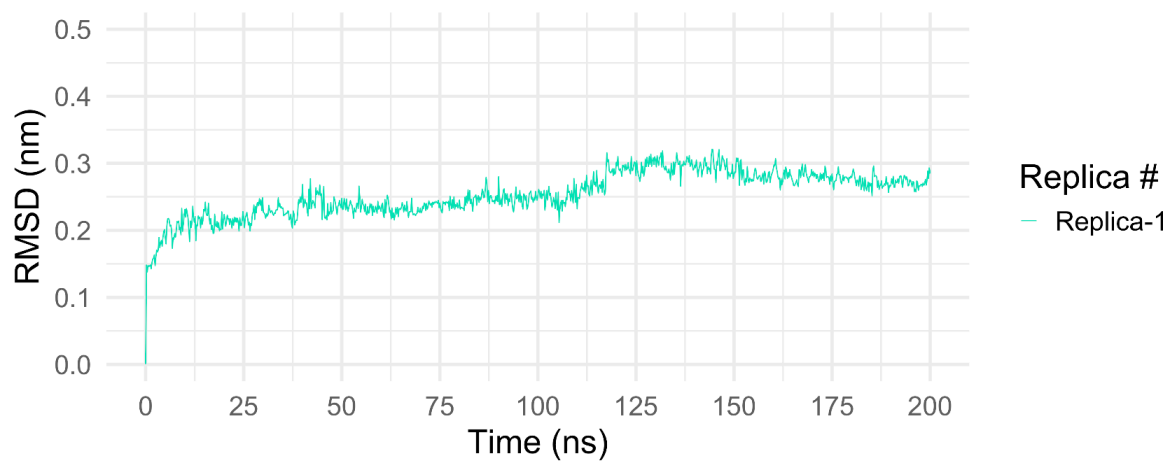
ZHANG, Y. *et al.* Structures of Darunavir-Resistant HIV-1 Protease Mutant Reveal Atypical Binding of Darunavir to Wide Open Flaps. **ACS Chemical Biology**, 28 abr. 2014. v. 9, n. 6, p. 1351–1358.

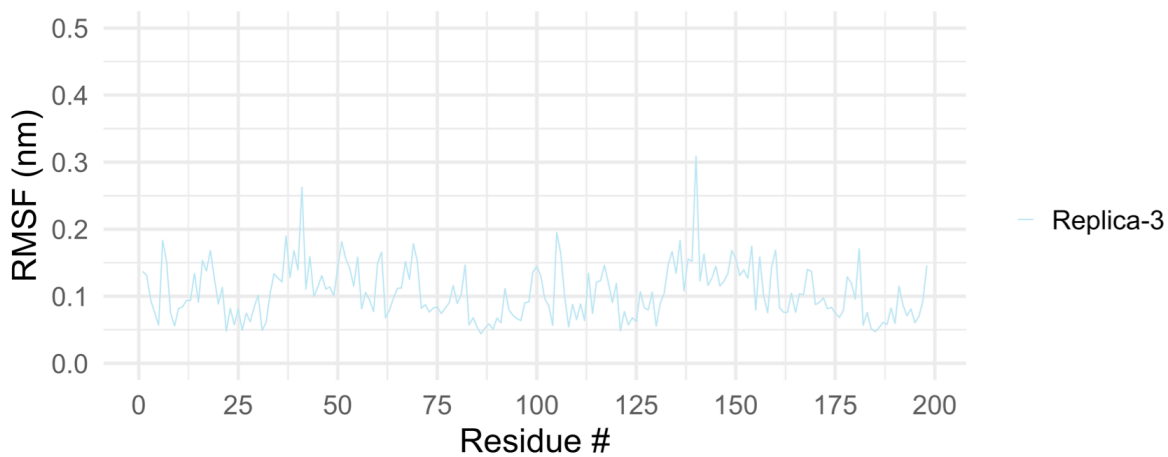
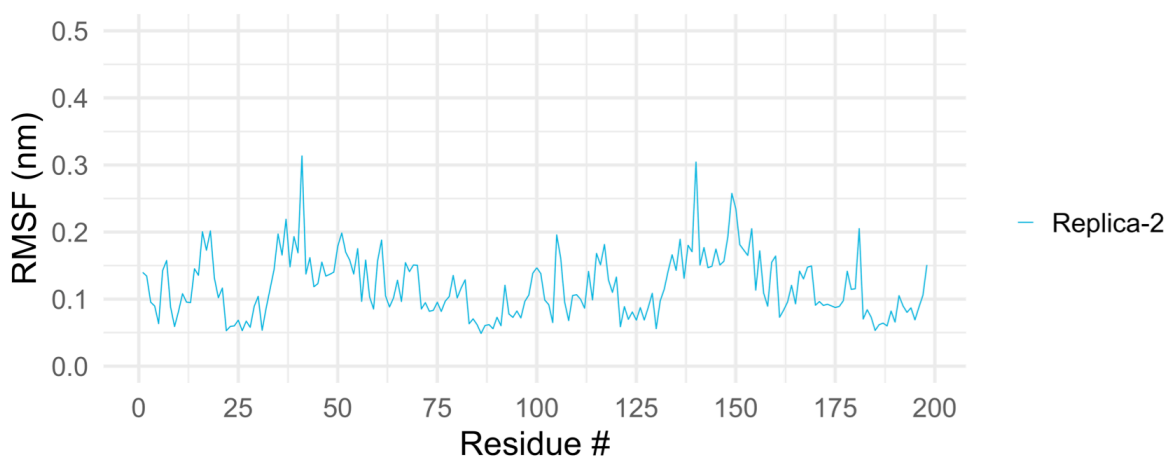
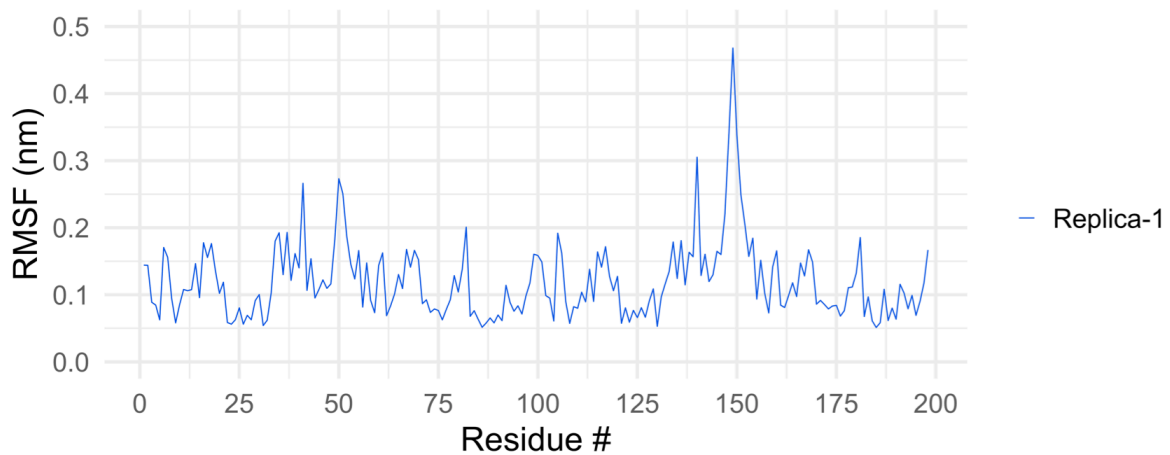
## APÊNDICE

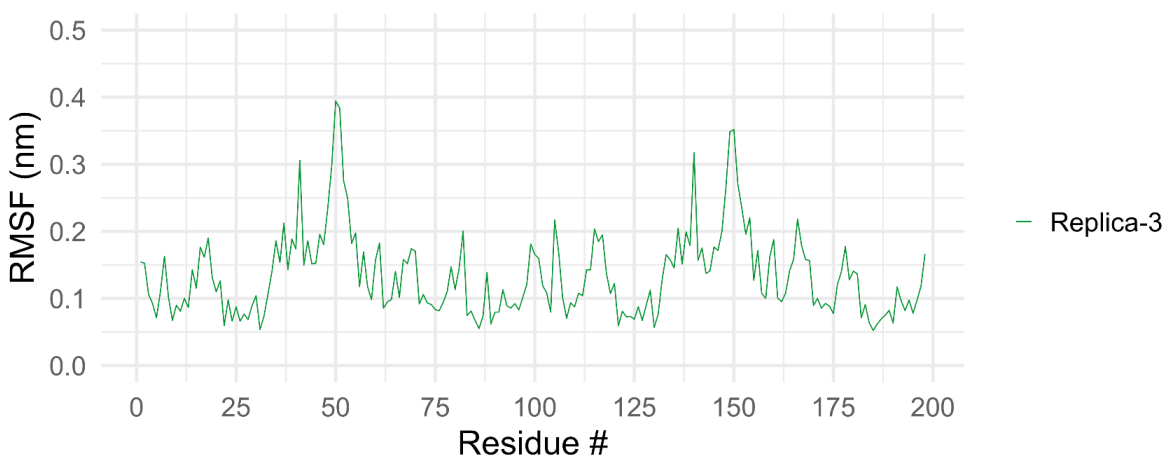
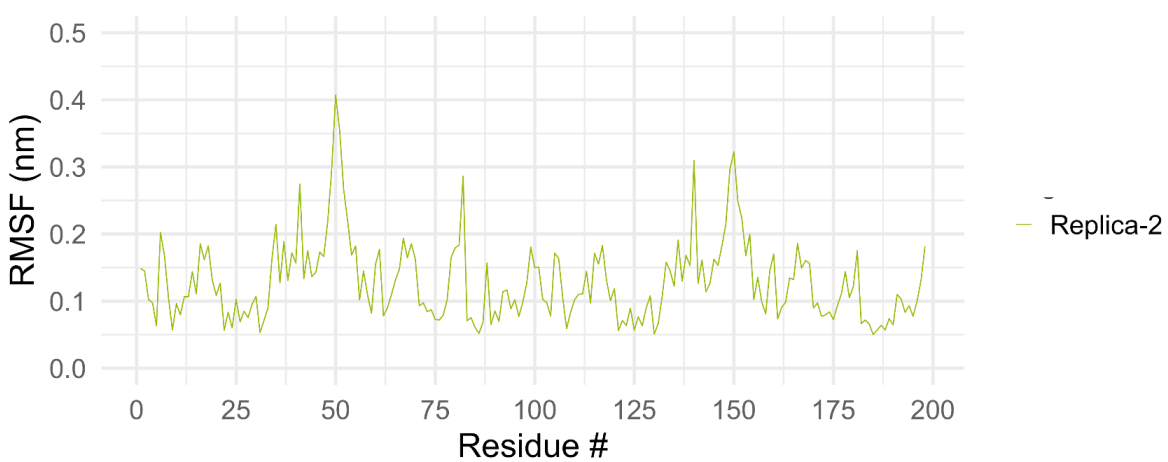
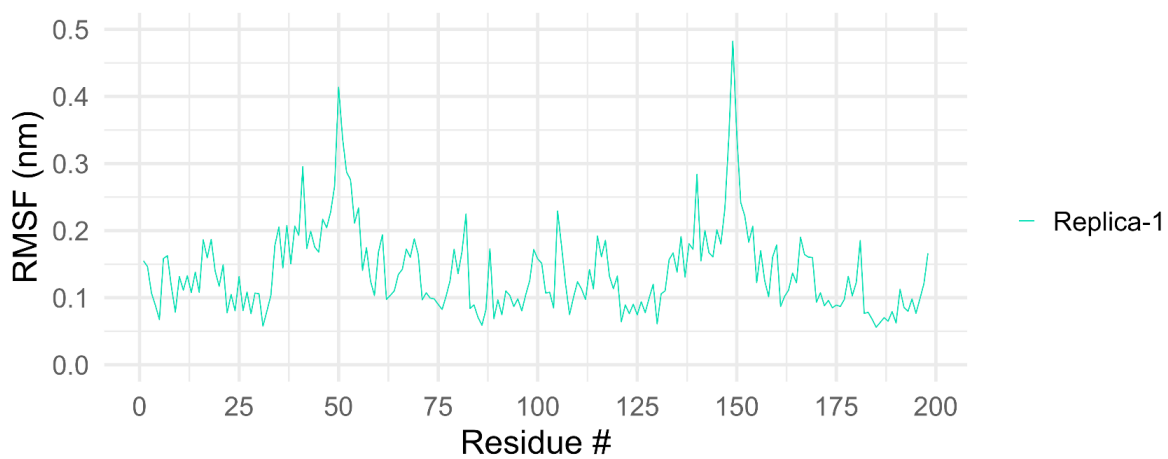
### Apêndice A - Análise RMSF entre 150 - 200 ns de simulação



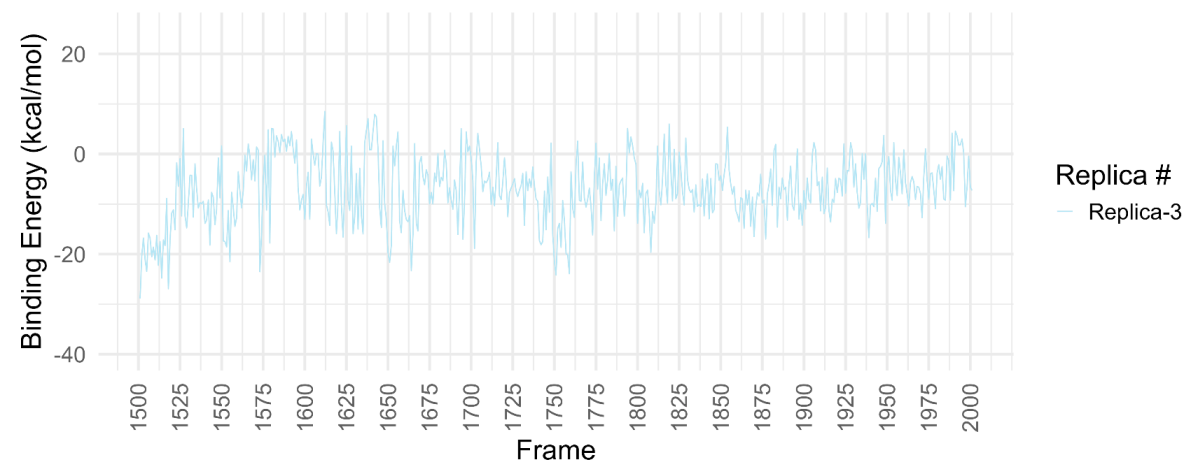
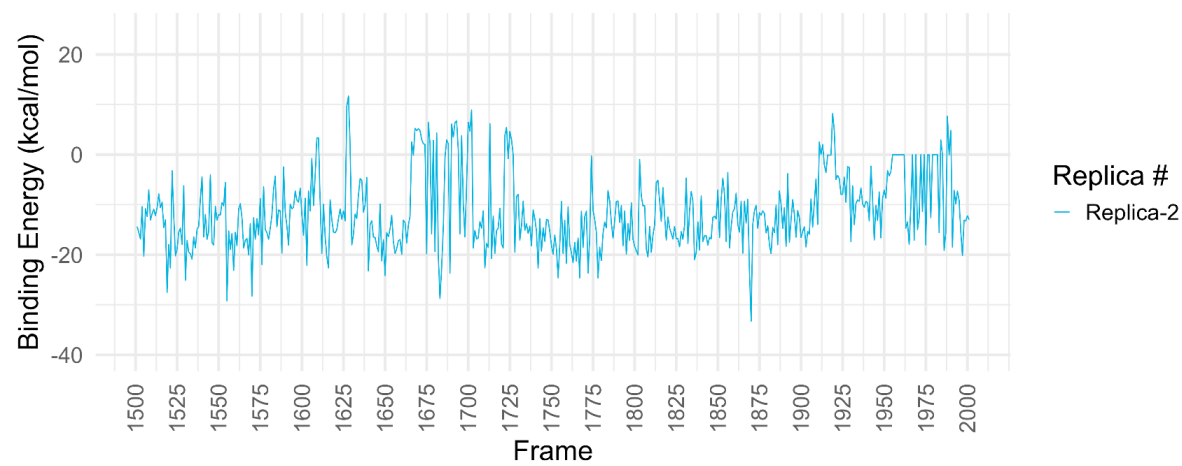
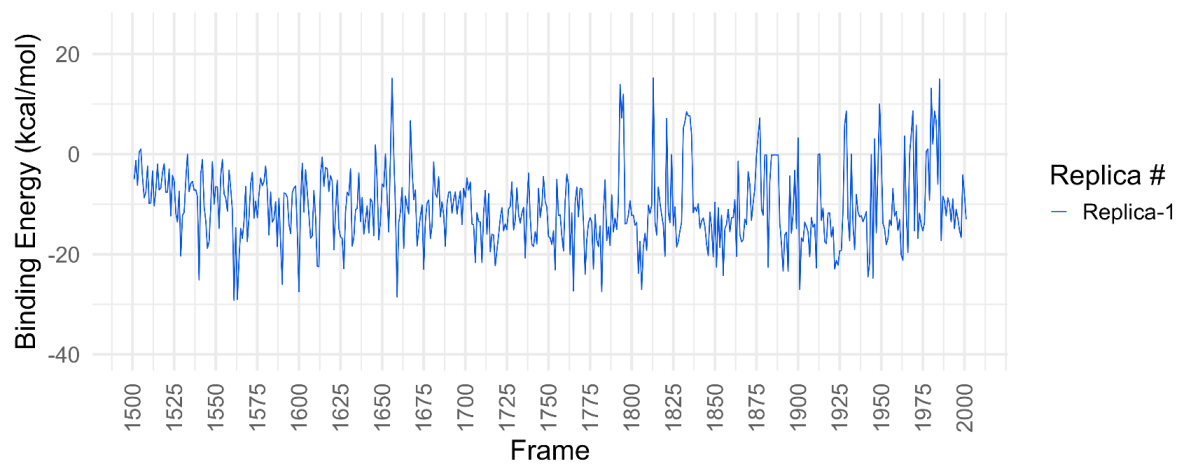
Apêndice B - Análise RMSD das réplicas individuais do complexo PR<sub>MUT</sub>-GRL-063

Apêndice C - Análise RMSD das réplicas individuais do complexo PR<sub>MUT</sub>-GRL-004

Apêndice D - Análise RMSF das réplicas individuais do complexo PR<sub>MUT</sub>-GRL-063

Apêndice E - Análise RMSF das réplicas individuais do complexo PR<sub>MUT</sub>-GRL-004

Apêndice F - Análise MQ/MM das réplicas individuais do complexo PR<sub>MUT</sub>-GRL-004 nos últimos 500 frames



### Apêndice G - Análise MQ/MM das réplicas individuais do complexo PR<sub>MUT</sub>-GRL-004 nos últimos 500 frames

