



**Rudielson dos Santos Silva**

**Energias livres de ligação absoluta do peptídeo antiviral  
ATN-161 com alvos proteicos de SARS-CoV-2**

**Dissertação de Mestrado**

Dissertação apresentada ao Programa de Pós Graduação em Química da PUC-Rio como requisito parcial para obtenção do grau de Mestre em Química.

Orientador: Prof. Dr. André Silva Pimentel

Rio de Janeiro, fevereiro de 2023



**Rudielson dos Santos Silva**

**Energias livres de ligação absoluta do peptídeo antiviral  
ATN-161 com alvos proteicos de SARS-CoV-2**

Dissertação apresentada ao Programa de Pós Graduação em **Química da PUC-Rio** como requisito parcial para obtenção do grau de Mestre em Química.

**Prof. André Silva Pimentel**

Orientador

Departamento de Química – PUC-Rio

**Prof. Luis Gustavo Dias**

Departamento de Química – FFCLRP – USP

**Prof. Roberto Dias Lins Neto**

Instituto Aggeu Magalhães – FIOCRUZ

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Rudielson dos Santos Silva**

Graduou-se em Licenciatura plena em Química na Universidade Estadual do Piauí em 2021. Atua na área de Físico-química, nas linhas de pesquisa de Interação de Fármacos em Sistemas Biológicos e Química Computacional.

#### Ficha Catalográfica

Silva, Rudielson dos Santos

Energias livres de ligação absoluta do peptídeo antiviral ATN-161 com alvos proteicos de SARS-CoV-2 / Rudielson dos Santos Silva ; orientador: André Silva Pimentel. – 2023.

117 f. : il. color. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Química, 2023.

Inclui bibliografia

1. Química – Teses. 2. Covid-19. 3. Dinâmica molecular. 4. Amostragem Guarda-chuva. 5. Integrina. 6. Complexo proteína-ligante. I. Pimentel, André Silva. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Química. III. Título.

CDD: 540

## Dedicatória

*Dedico esta dissertação à minha família, em nome dos meus pais Francisco (in memoriam) e Francisca e aos meus avós Francisco e Cerciana, que nunca mediram esforços para me dar uma educação de qualidade e sempre me ensinarem a ser uma pessoa melhor. Obrigado pai por todo apoio e inspiração que me destes em vida, sei que deve estar muito orgulhoso.*

## **Agradecimentos**

Primeiramente agradeço a Deus por me segurar em todos os momentos difíceis, por sempre me mostrar os caminhos certos e dar-me a oportunidade de me tornar um ser humano melhor a cada dia.

A minha família por sempre acreditarem no meu potencial de crescimento e sempre me apoiarem incondicionalmente. A minha mãe Francisca, por todo carinho e compreensão, aos meus avós Francisco e Cerciana que sempre acreditaram e apostaram em mim e por todo o carinho e amor que tem por mim.

Ao meu pai Francisco (in memoriam) que sempre foi minha principal inspiração e o maior exemplo de um homem que eu poderia ter. Pai, sei que se estivesse aqui com certeza estaria muito feliz e mostrando a todos mais uma conquista que alcancei. Obrigado por tudo.

As minhas irmãs Rutiely e Rithyele, por todo carinho e irmandade e por sempre estarem ao meu lado, ao meu cunhado Gabriel pela amizade e parceria. Aos meus amigos de sempre, Willian e Eduardo, por sempre dividirem os melhores momentos de descontração e por toda a amizade que construímos.

A minha namorada Cilene por sempre me apoiar e buscar compreender o motivo que nos distanciou fisicamente por um tempo, além de todo carinho e amor que cultivava em nossa relação.

Aos meus amigos da igreja e vida Bruno, Wesley, Marília, Tainey, Elane, Helderson, Natalia, Mairis, Karen, Rebeca, Waleska, Lya, Conceição, Celida, Aristóteles, Leonardo, Jonas, Alexandre Diógenes, Jaqueline Fernanda, Davi Libânio e Isnaider. Obrigado por todos os bons momentos compartilhados.

Aos meus amigos do LFCx, Lucas Miguel, Felipe e João Paulo por toda a paciência e troca de conhecimento durante este período que me dediquei ao mestrado. A minha amiga Rayla pela parceria que vem desde a graduação.

Ao meu orientador, André Pimentel, pela orientação super competente durante esses 2 anos de parceria. Por todas as oportunidades ofertadas, pela recepção, cuidado e por todos os conselhos e conhecimentos repassados durante esta fase.

A PUC-RIO, CNPQ e INCT-FCx pelo apoio financeiro e institucional.

Ao Rio de Janeiro por ter me acolhido tão bem, ao Flamengo, meu time de coração por me conceder momentos felizes durante dias difíceis e de saudade.

A todos aqueles que, de certa forma, contribuíram para que eu alcançasse voos maiores, tanto como profissional, quanto como ser humano.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## Resumo

Silva, Rudielson dos Santos; Pimentel, André Silva. **Energias livres de ligação absoluta do peptídeo antiviral ATN-161 com alvos proteicos de SARS-CoV-2.** Rio de Janeiro, 2022. Dissertação de mestrado – Departamento de Química, Pontifícia Universidade Católica do Rio de Janeiro.

A afinidade de ligação do pentapeptídeo antiviral ATN-161 com os alvos da integrina  $\alpha_5\beta_1$  conformações fechada e aberta, a protease principal SARS-CoV-2 e a proteína S variante omicron (B.1.1.529) complexada com hACE2 foram estudados usando ancoramento molecular e simulação de dinâmica molecular. Os complexos foram formados com sucesso através do uso de simulações de ancoragem molecular, obtendo as poses de ligação ATN-161 com os alvos estudados. Subsequentemente, simulações de dinâmica molecular clássica foram executadas para investigar a estabilidade do complexo formado. Simulações de dinâmica molecular direcionada e amostragem guarda-chuva usando análise de histograma ponderado foram realizadas para obter o potencial de força média de cada sistema, e calcular a energia livre de ligação para o ATN-161 com o sítio de ligação dos respectivos alvos. Os resultados mostram que o ATN-161 se liga à integrina  $\alpha_5\beta_1$  em sua forma ativa e inativa, com energias livres de ligação de  $-14.4 \pm 0.5$  e  $-12.6 \pm 0.2$  kcal mol<sup>-1</sup>, respectivamente. Além disso, o ATN-161 se liga fracamente à proteína S variante omicron complexada com hACE2 com uma energia de apenas  $-3.2 \pm 0.3$  kcal mol<sup>-1</sup>. Entretanto, o ATN-161 apresenta maior afinidade de ligação na M<sup>PRO</sup>, computando uma energia de ligação de  $-17.7 \pm 0.2$  kcal mol<sup>-1</sup>, indicando que este peptídeo se liga fortemente ao M<sup>PRO</sup>. Por fim, o estudo da cinética de ligação com o método  $\tau$ -RAMD indica que o ATN-161 apresenta um tempo de residência relativo similar ao do ligante de referência GC376. Os tempos de residências relativos são  $2,8 \pm 0,4$  e  $2,5 \pm 0,5$  ns para o ATN-161 e para o GC376, respectivamente. Os resultados obtidos neste trabalho indicam um potencial candidato a fármaco contra o SARS-CoV-2, em especial para o surgimento de novas variantes do SARS-CoV-2.

## Palavras-chave

Covid-19; Dinâmica molecular; Amostragem Guarda-chuva; Integrina;  
Complexo proteína-ligante; Ancoragem molecular.



## Abstract

Silva, Rudielson dos Santos; Pimentel, André Silva. **Absolute binding free energies of the antiviral peptide ATN-161 with protein targets of SARS-CoV-2.** Rio de Janeiro, 2022. Dissertação de mestrado – Departamento de Química, Pontifícia Universidade Católica do Rio de Janeiro.

The binding affinity of the antiviral pentapeptide ATN-161 with the targets of the  $\alpha_5\beta_1$  integrin closed and open conformations, the main protease and the omicron variant S (B.1.1.529) protein complexed with hACE2 of SARS-CoV-2 was studied using molecular docking and molecular dynamics simulation. The complexes were successfully formed through the use of molecular docking simulations, in which binding poses of ATN-161 with the studied targets were obtained. Subsequently, classical molecular dynamics simulations were performed to investigate the stability of the formed complex. Steered molecular dynamics simulations, umbrella sampling and weighted histogram analysis method were used to obtain the mean force potential of each system, and thus calculate the binding energy for ATN-161 in the binding site of the respective protein targets. The results shows that ATN-161 binds to  $\alpha_5\beta_1$  integrin in its active and inactive form, with binding energies of  $-14.4 \pm 0.5$  and  $-12.6 \pm 0.2$  kcal mol<sup>-1</sup>, respectively. ATN-161 binds weakly to omicron variant protein S complexed with hACE2 with an energy of only  $-3.2 \pm 0.3$  kcal mol<sup>-1</sup>. However, ATN-161 binds to M<sup>PRO</sup> with the binding energy of  $-17.7 \pm 0.2$  kcal mol<sup>-1</sup>, indicating that this peptide binds strongly to M<sup>PRO</sup>. Finally, the study of the binding kinetics with the  $\tau$ -RAMD method indicates that ATN-161 presents the relative residence time of  $2.8 \pm 0.4$  and  $2.5 \pm 0.5$  ns for ATN-161 and GC376, respectively. The results obtained in this work indicate a potential drug candidate against SARS-CoV-2, especially with the emergence of new variants of SARS-CoV-2.

## Keywords

Covid-19; Molecular Dynamic; Umbrella Sampling; Integrin; Protein-ligand complex; molecular docking.

## Sumário

1. Introdução .....	17
2. Justificativa e motivação.....	20
3. Objetivos .....	23
3.1. Objetivo geral .....	23
3.2. Objetivos específicos .....	23
4. Referencial Teórico .....	24
4.1. Peptídeo antimicrobiano.....	24
4.1.1. ATN-161 .....	24
4.2. Moléculas de adesão celular .....	26
4.2.1. Integrinas.....	26
4.3. SARS-CoV-2 .....	29
4.3.1. Mecanismo de entrada.....	31
4.3.1.1. Principais receptores.....	32
4.3.1.2. Tratamentos disponíveis .....	34
4.4. Simulação molecular .....	34
4.4.1. Ancoramento molecular.....	35
4.4.1.1. Autodock .....	36
4.4.2. Dinâmica molecular .....	38
4.4.2.1. Campo de Força.....	39
4.4.2.1.1. Estiramento de ligação .....	42
4.4.2.1.2. Deformação angular de ligação.....	42
4.4.2.1.3. Torsões de ligação .....	43
4.4.2.1.4. Potencial de van der Waals.....	43
4.4.2.1.5. Potencial eletrostático .....	44
4.4.2.2. Condições iniciais do sistema em dinâmica molecular.....	44
4.4.2.2.1. Algoritmo de descida mais íngreme .....	45
4.4.2.2.2. Método de gradientes conjugados .....	46
4.4.2.3. Ensembles.....	46
4.4.2.4. Termostatos .....	47
4.4.2.4.1. Termostato de Berendsen .....	47
4.4.2.4.2. Termostato de Reescalamento de Velocidade .....	48
4.4.2.4.3. Termostato de Nosé-Hoover .....	48
4.4.2.5. Barostatos .....	49
4.4.2.5.1. Barostato de Berendsen.....	49
4.4.2.5.2. Barostato de Parrinello-Rahman .....	50

4.4.2.6. Condições periódicas de contorno .....	51
4.4.2.7. Algoritmos de integração em dinâmica molecular .....	52
4.4.2.7.1. Algoritmo de Verlet .....	53
4.4.2.7.2. Algoritmo de Leap-Frog .....	54
4.4.2.8. Soma de Ewald (PME) .....	55
4.4.2.9. Algoritmo de Restrição .....	56
4.4.2.9.1. Algoritmo de Restrição LINCS .....	56
4.4.2.9.2. Algoritmo de Restrição SHAKE .....	57
4.4.2.10. Amostragem guarda-chuva .....	57
4.4.2.11. Método da análise de histograma ponderado (WHAM) .....	58
4.4.2.12. Dinâmica molecular de aceleração randômica ( $\tau$ -RAMD) .....	59
4.4.2.13. GROMACS .....	60
 5. Metodologia .....	 62
5.1. Preparando as estruturas tridimensionais .....	62
5.2. Ancoramento Molecular .....	63
5.3. Dinâmica molecular clássica .....	64
5.4. Amostragem guarda-chuva .....	66
5.5. Dinâmica molecular de aceleração randômica ( $\tau$ -RAMD) .....	67
 6. Resultados e Discussão .....	 69
6.1. Ancoramento molecular .....	69
6.2. Dinâmica molecular .....	73
6.3. Método de Amostragem Guarda-Chuva .....	81
6.4. Dinâmica molecular de aceleração randômica ( $\tau$ -RAMD) .....	86
 7. Implicações .....	 90
 8. Conclusões .....	 91
 9. Perspectivas Futuras .....	 93
 10. Referências Bibliográficas .....	 94
 11. Anexos .....	 109

## Lista de figuras

<b>Figura 1.</b> Formula Estrutural do ATN-161 (a) 2D e (b) 3D. Figura extraída de (SILVA et al., 2022). ....	25
<b>Figura 2.</b> Estados conformacionais das integrinas. A. Conformação inativa, B. Conformação intermediária, C. conformação Ativa. Figura adaptada de (PETERSON; KOVAL, 2021) .....	27
<b>Figura 3.</b> Representação da estrutura tridimensional da integrina $\alpha_5\beta_1$ em modelo Cartoon. (a) conformação fechada (código PDB: 4WK4) e (b) conformação aberta (código PDB: 7NWL). As cadeias da subunidade $\alpha_5$ estão representadas na cor azul e $\beta_1$ na cor verde. ....	28
<b>Figura 4.</b> Representação da estrutura viral do SARS-CoV-2. Onde as proteínas estruturais S (spike), E (envelope), M (membrana) e N (nucleocapsídeo) são mostradas. Figura extraída de (LIMA et al., 2020) .....	30
<b>Figura 5.</b> Representação tridimensional em modelo Cartoon da Estrutura Cristalográfica da (a) proteína S, na cor bege complexada a Enzima Conversora de Angiotensina Humana (hACE2) na cor marrom e (b) protease principal ( $M^{PRO}$ ) na cor verde. ....	30
<b>Figura 6.</b> Representação do Mecanismo de Entrada do Virus do SARS-CoV-2 na Célula hospedeira. Figura extraída do Glossário Fiocruz. ....	31
<b>Figura 7.</b> Representação da estrutura da Glicoproteína S do SARS-CoV-2, domínio de ligação ao receptor (RBD), e regiões das subunidades S1 e S2. Figura adaptada de Lubio Science.....	32
<b>Figura 8.</b> Esquema das interações em um campo de força. As ligações covalentes são representadas por linhas sólidas e a interações intermoleculares são indicadas como linhas tracejadas. b representa o estiramento da ligação entre as partículas 1 e 2, $\Phi$ são as torções das ligações que envolve as partículas 1, 2, 3 e 4, $\theta$ representa as deformações angulares entre as partículas 2, 3 e 4, e $r_{ij}$ representa as interações intermoleculares de van der Waals e Coulomb entre a partícula externa 5 e partícula 3. Adaptada de (PONDER; CASE, 2003). ....	41
<b>Figura 9.</b> Representação de um sistema com condições periódicas de contorno. A figura central é o sistema estudado e nas adjacências os sistemas especulares, a seta indica que o movimento executado na figura central é repetida nas imagens especulares. Figura extraída da literatura (ALLEN.P, 2004) .....	52
<b>Figura 10.</b> As primeiras 10 poses geradas pelo acoplamento molecular do ATN-161 na (a) integrina conformação fechada (código PDB: 4WK4) e na (b) conformação aberta (código PDB: 7NWL), na (c) proteína S da variante ômicron/hACE2 (código PDB: 7WPB) e na (d) $M^{PRO}$ do SARS-CoV-2 (código PDB: 7BQY). ....	71
<b>Figura 11.</b> Reancoragem molecular do ligante N3 no sítio de ligação $M^{PRO}$ para os dados cristalográficos encontrados no código	

PDB: 7BQY. A pose cristalográfica de referência é mostrada em roxo e a pose do encaixe é mostrada em laranja.....	72
<b>Figura 12.</b> Raiz quadrada do desvio quadrático médio (RMSD) da posição atômica da cadeia principal da proteína complexada com ligante, cadeia principal da proteína sem ligante e átomos pesados de ATN-161 (azul). (a) conformação fechada e (b) conformação aberta da integrina $\alpha_5\beta_1$ , no qual a cadeia A da proteína complexada está em marrom e cadeia B em roxo, a cadeia A da proteína sem ligante está em vermelho e a cadeia B em verde. (c) proteína S da variante ômicron/hACE2 e (d) complexo da M <sup>PRO</sup> do SARS-CoV-2, proteína complexada em laranja e proteína sem ligante em cinza. ....	74
<b>Figura 13.</b> Raiz quadrada da flutuação quadrática média (RMSF) da posição atômica por resíduo de aminoácido. As regiões que mostraram variação de RMSF maior que 0,3 nm foram destacadas. (a) RMSF da integrina de conformação fechada, (b) representação das regiões de maior variação na integrina conformação fechada. (c) RMSF da integrina de conformação aberta, (b) representação das regiões de maior variação na integrina conformação aberta. Em (a) e (c) as linhas sólidas referem-se à cadeia A e as linhas pontilhadas referem-se à cadeia B, em (b) e (d) as marcações de cor verde representam as maiores variações na cadeia A e em magenta na cadeia B. ....	75
<b>Figura 14.</b> Distribuição de ligações de hidrogênio em termos de ocupância relativa entre o ATN-161 e os alvos. (a) integrina $\alpha_5\beta_1$ de conformação fechada e (b) aberta, (c) proteína S da variante ômicron/hACE2. (d) M <sup>PRO</sup> do SARS-CoV-2. “Princ” refere-se a cadeia principal e “Lat” a cadeia lateral.....	76
<b>Figura 15.</b> Número de ligações de hidrogênio dos sistemas de ATN-161 e alvos (a) integrina $\alpha_5\beta_1$ com conformação fechada e (b) aberta, (c) proteína S da variante ômicron/hACE2. (d) M <sup>PRO</sup> do SARS-CoV-2. ....	77
<b>Figura 16.</b> Interações intermoleculares entre ATN-161 e alvos. (a) interações da integrina $\alpha_5\beta_1$ com conformação fechada, (b) ligações de hidrogênio da integrina $\alpha_5\beta_1$ com conformação fechada, (c) interações da integrina $\alpha_5\beta_1$ com conformação aberta, (d) interações da proteína S da variante ômicron/hACE2, (d) interações da M <sup>PRO</sup> do SARS-CoV-2 e (e) ligações de hidrogênio da M <sup>PRO</sup> do SARS-CoV-2. ....	79
<b>Figura 17.</b> Perfis de energia livre de Gibbs e histogramas obtidos pelo método de amostragem guarda-chuva e análise de histograma ponderado (WHAM) para ATN-161/Alvos. (a) potencial de força média e (b) histograma para integrina $\alpha_5\beta_1$ conformação fechada, (c) potencial de força média e (d) histograma para integrina $\alpha_5\beta_1$ conformação aberta, (e) Potencial de força média e (f) histograma para a proteína S da variante de ômicron/hACE2, (g) potencial de força média e (h) histograma para a M <sup>PRO</sup> do SARS-CoV-2. ....	82

**Figura 18.** Resultados estatísticos para o complexo (a) ATN-161 com M<sup>PRO</sup> e (b) GC376 com M<sup>PRO</sup>: Função de distribuição cumulativa (CDF), função de distribuição de tempo residência efetivo e teste de Kolmogorov-Smirnov (KS)..... 87

**Figura 19.** Variação do tempo de residência médio para cada replica. Complexo (a) ATN-161 – M<sup>PRO</sup> e (b) GC376 – M<sup>PRO</sup>. A caixa representa o intervalo dos valores médios entre as trajetórias e os valores discrepantes são representados por círculos. A barra mostra o desvio dos valores entre as trajetórias de cada replica. A média é mostrada pelas linhas vermelhas tracejadas e a mediana pela linha sólida laranja. .... 88

## Lista de tabelas

**Tabela 1.** Pontuação de energia de ligação ( $E_{\text{bind}}$ ) para a ancoragem do ATN-161 com as conformações fechada (código PDB: 4WK4) e aberta (código PDB: 7NWL) da integrina  $\alpha_5\beta_1$ , M<sup>PRO</sup> (código PDB: 7BQY) e proteína S da variante ômicron complexada ao hACE2 (código PDB: 7WPB) obtida por ancoragem molecular. .... 70

**Tabela 2.** Energia livre de ligação ( $\Delta G_{\text{bind}}$ ) do ATN-161 com as conformações fechada (código PDB: 4WK4) e aberta (código PDB: 7NWL) da integrina  $\alpha_5\beta_1$ , da proteína M<sup>PRO</sup> (código PDB: 7BQY) e da proteína S da variante do ômicron/hACE2 (código PDB: 7WPB), calculada por amostragem guarda-chuva. .... 83

## **Lista de Siglas e Abreviaturas**

SRAG - Síndrome respiratória aguda grave

SARS-CoV - Síndrome respiratória aguda grave de coronavírus

SARS-CoV-2 – Síndrome respiratória aguda grave de coronavírus 2

OMS - Organização Mundial de Saúde

DM – Dinâmica Molecular

OMS – Organização Mundial de Saúde

EUA– Estados Unidos da América

FDA– Food and Drugs Administration

AA – All atom

OPLS – Potenciais otimizados para simulações de líquidos

TIP4P – Potencial intermolecular transferível com 4 pontos

AMP – Peptídeo antimicrobiano

AVP – Peptídeo antiviral

RGD – Ácido arginilglicilaspártico

PRO – Prolina

HIS – Histidina

SER – Serina

CYS – Cisteína

ASN – Asparagina

2D – Bidimensional

3D – Tridimensional

ECM – Matriz extracelular

NSP – Proteína não estrutural

RBD – Domínio de ligação ao receptor

RMSD – Raiz Quadrada do Desvio Quadrático Médio

LGA – Algoritmo Genético Lamarckiano

LJ – Lennard-Jones

LINCS – Solucionador de Restrições Lineares

PMF– Potencial de força média

FF14SB – Force Field 14 Stony Brook

GAFF – Generalized Amber Force Field

vdW- van der Waals

*“A nossa maior glória não reside no fato de nunca cairmos, mas sim em  
levantarmo-nos sempre depois de cada queda.”*

Oliver Goldsmith



## Introdução

A Síndrome Respiratória Aguda Grave (SRAG) é uma doença que foi relatada pela primeira vez em humanos entre os anos de 2003 e 2004 na China (XU et al., 2004). Durante esse período esta infecção espalhou-se para mais de 28 regiões pelo mundo, acometendo um total de 8096 infectados e 774 vítimas fatais (SHI; HU, 2008). Esta doença epidêmica está intimamente associada a família de coronavírus, nos quais os morcegos são indiciados como hospedeiros, o que levou a defini-la como síndrome respiratória aguda grave de coronavírus, mais conhecida pelo termo em inglês (Severe Acute Respiratory Syndrome-Coronavirus, SARS-CoV) (MENACHERY et al., 2016).

No ano de 2019, um surto de infecções respiratórias na cidade de Wuhan na China chamou atenção das autoridades públicas de saúde. Estudos preliminares identificaram o surgimento de um novo vírus que foi nomeado provisoriamente como vírus 2019nCoV (ZHOU et al., 2020; ZHU et al., 2020). No ano seguinte, em 30 de janeiro de 2020, esta epidemia foi declarada como uma emergência de saúde pública de importância internacional pela Organização Mundial de Saúde (OMS) (WHO, 2022). Este problema de saúde, que se espalhou rapidamente por todo planeta, parou o mundo entre 2020 e 2021, tornando-se uma pandemia e um grande problema para economia mundial.

O fato de ser uma doença até então desconhecida e altamente transmissível, sem métodos eficientes de prevenção e combate, fez com que as pessoas se mantivessem em isolamento, tendo apenas trabalhos essenciais em funcionamento. Esta situação motivou vários pesquisadores e empresas farmacêuticas pelo mundo a voltarem suas atenções para a descoberta de novos fármacos e vacinas para o combate ao 2019nCoV, que posteriormente seria nomeado oficialmente como SARS-CoV-2 (EL-DIN ABUO-RAHMA et al., 2020; XIU et al., 2020; YANG; YANG, 2021).

Em 2020, a agência *Food and Drugs Administration* (FDA) dos Estados Unidos da América autorizou o uso emergencial do fármaco Remdesivir como primeiro tratamento contra o SARS-CoV-2 no mundo (RUBIN et al., 2020). As primeiras vacinas foram aprovadas entre o fim de 2020 e começo de 2021,

quando as campanhas de vacinação pelo mundo se iniciaram (STASI et al., 2022).

De acordo com a OMS, existem atualmente 11 vacinas aprovadas para uso emergencial e 145 em desenvolvimento clínico (STASI et al., 2022). Além disso, somente corticosteroides sistêmicos, bloqueadores de receptores IL-6 e Baricitinib são fortemente recomendados para casos críticos e graves, e apenas Nimatrelvir e Ritonavir para casos leves (WHO, 2022). Dados da OMS de maio de 2022 indicam que apenas 57 países vacinaram 70% de sua população, tendo quase 1 bilhão de pessoas que permanecem sem a vacina (WHO, 2022). De acordo com informativos da OMS de novembro de 2022, apenas aproximadamente 5 bilhões de pessoas concluíram a vacinação com todas as doses necessárias.

Apesar do desenvolvimento de medicamentos e vacinas, existe uma preocupação na comunidade científica com novas ondas de infecção, devido ao surgimento de novas variantes com potencial de alta transmissão (SINGANAYAGAM et al., 2022). Portanto, o cenário atual ainda remete a procura por terapias que apresentam eficiência no combate deste vírus. A busca por novos agentes terapêuticos contra o SARS-CoV-2 está intimamente associada com a inibição de uma ou mais etapas da infecção e replicação viral nas células (BENVENUTO et al., 2020; CHOWDHURY et al., 2020; MOTHAY; RAMESH, 2020; XIU et al., 2020).

O mecanismo de entrada do vírus do SARS-CoV-2 ocorre pela invasão da célula hospedeira através do reconhecimento da proteína S (do termo em inglês S de spike) do vírus por receptores de membrana celular, Enzima Conversora de Angiotensina (ACE2). Mais recentemente foi reportado que as integrinas possuem importância no mecanismo de entrada do vírus do SARS-CoV-2 (MAIER, H. J.; BICKERTON; BRITTON, 2015; MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021; OTHMAN et al., 2022). Após a ligação do vírus pela proteína S e a sua mudança conformacional, o envelope viral se funde com a membrana celular do hospedeiro, liberando o RNA viral na célula hospedeira (SHEREEN et al., 2020).

O conhecimento do mecanismo de entrada do vírus e a descoberta da estrutura tridimensional (3D) destas proteínas, possibilitou o planejamento racional de novos fármacos que podem atuar como inibidores do SARS-CoV-

2. Após a identificação do sítio ativo e as etapas fundamentais da infecção viral, pôde-se estudar as interações e as afinidades de ligação entre ligantes e receptores, e conseqüentemente, iniciar a descoberta de novas moléculas com atividade contra o SARS-CoV-2 através da química computacional (AROOJ et al., 2022).

Os peptídeos antimicrobianos (AMP) e os antivirais (AVP) tem se apresentado como uma boa alternativa, sendo uma das categorias mais promissoras para candidatos a novos fármacos para diversas doenças, por terem como características a especificidade mesmo em pequenas concentrações (TONK; RŮŽEK; VILCINSKAS, 2021), fácil síntese, e possibilidade de modificação para melhorar suas propriedades (AGARWAL; GABRANI, 2021). Estudos recentes apontaram o peptídeo ATN-161 como potencial inibidor do SARS-CoV-2 (AMRUTA et al., 2021; BEDDINGFIELD et al., 2021). Como será mostrado na Figura 1, esta molécula trata-se de um pentapeptídeo derivado da região de sinergismo da proteína extracelular fibronectina.

Como os estudos *in vitro* e *in vivo* são caros, e muitas vezes, não fornecem informações mecanísticas e de interação destes compostos ativos em alvos proteicos do SARS-CoV-2, os estudos computacionais destes peptídeos ativos através de simulações de ancoramento molecular (GARRETT M et al., 2012; MORRIS et al., 2014) e dinâmica molecular (DM), (ABRAHAM, Mark James et al., 2015; SPOEL, VAN DER et al., 2005) (LEMKUL; BEVAN, 2010) são de extrema importância e relevância para fornecer as informações necessárias para se entender os ensaios *in vitro* e *in vivo*. Estas ferramentas possibilitam o cálculo da afinidade de ligação e da cinética de ligação dos peptídeos ativos nos alvos proteicos do SARS-CoV-2. Sendo assim, esta dissertação teve como objetivo estudar as interações do peptídeo ATN-161 com sítios de ligação de potenciais alvos contra o SARS-CoV-2, obtendo as energias livres de ligação do ATN-161 com alguns alvos proteicos importantes e a cinética de ligação do ATN-161 no sítio de ligação de um dos alvos.

## 2

### Justificativa e motivação

Estudos para o uso de AMP's e AVP's tem-se tornado cada vez mais comuns (GAN et al., 2021). A necessidade da descoberta de medicamentos que não desenvolvam resistência frente a doenças infecciosas tem contribuído para o aumento destes estudos. Os AMPs são uma das principais categorias que não demonstram o desenvolvimento de resistência. Em 2019, a OMS publicou a lista de medicamentos essenciais, que inclui mais de 40 antimicrobianos (GAN et al., 2021). Esse tipo de agente terapêutico pode ser obtido tanto de forma direta da natureza como através de sínteses simples. Além disso, existe a possibilidade da otimização utilizando um peptídeo proveniente da natureza e modificando-o sinteticamente (GAN et al., 2021).

Estudos recentes relataram que o pentapeptídeo ATN-161 apresenta ação contra o SARS-CoV-2. Os estudos *in vitro* e *in vivo* tiveram como resultado uma diminuição da carga viral quando colocado em cultura de células e quando administrado a camundongos (AMRUTA et al., 2021; BEDDINGFIELD et al., 2021). Os resultados para estes estudos são bastante promissores, tendo em vista que o ATN-161 já possui diversos estudos para tratamento de câncer e outras patogenicidades (CIANFROCCA et al., 2006; EDWARDS et al., 2020; KHALILI et al., 2006; STOELTZING et al., 2003).

O ATN-161 trata-se principalmente de um inibidor da integrina  $\alpha_5\beta_1$ . Esta proteína por sua vez, demonstrou associação com a angiogênese tumoral, migração e adesão de células tumorais a nicho de metástase, além de ter efeitos na resistência terapêutica (SCHAFFNER; RAY; DONTENWILL, 2013). Estudos recentes indicam que as integrinas atuam como via alternativa para entrada do vírus do SARS-CoV-2. A alta expressão destas proteínas de adesão em todo o corpo, e especificamente em tecidos pulmonares, corroboram para a integrina  $\alpha_5\beta_1$  se tornar um bom alvo deste vírus (LIU, J. et al., 2022; MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021; OTHMAN et al., 2022; PARK, E. J. et al., 2021).

Apesar dos estudos *in vitro* e *in vivo* do ATN-161 indicarem atividade inibitória contra o SARS-CoV-2, algumas questões continuam em aberto, tais como: a atividade contra SARS-CoV-2 é devido apenas a interação com a

integrina? Existe a possibilidade deste peptídeo ligar-se a mais de um alvo? Se sim, qual a afinidade de ligação? Quais amino ácidos contribuem para a interação proteína-ligante? A utilização de ferramentas de Química Computacional torna possível a obtenção de respostas para estas perguntas, e podem contribuir significativamente para o desenvolvimento deste e outros fármacos.

Ainda que existam muitos estudos sobre a atuação inibitória do ATN-161 com integrinas, não há relatos na literatura que indicam a forma de interação deste peptídeo com um sítio de ligação específico. No entanto, existem sugestões de que este peptídeo liga-se fora do bolso de ligação RGD, tendo como principal sítio a região de sinergismo já que esse peptídeo tem origem nesta região da fibronectina (KHALILI et al., 2006; SCHAFFNER; RAY; DONTENWILL, 2013). Porém, as modificações estruturais sofridas para otimização do seu potencial terapêutico podem influenciar na região de interação deste peptídeo (SCHAFFNER; RAY; DONTENWILL, 2013).

As simulações computacionais possuem um papel fundamental no planejamento de um fármaco e estão cada vez mais sendo empregadas pela comunidade científica (KĘSKA; GUSTAW; STADNIK, 2021; MORADI et al., 2022). Este tipo de ferramenta possibilita realizar previsões farmacológicas e farmacocinéticas entre outros parâmetros. Além disso, pode permitir a obtenção de informações a nível molecular que geralmente não são possíveis por meio de métodos experimentais (KHAN et al., 2019).

A partir da obtenção das estruturas 3D das proteínas e do peptídeo, é possível realizar estudos que podem obter respostas para questões que continuam em aberto para esse tema. Sendo assim, este trabalho tem a finalidade de estudar a nível molecular a interação e afinidade do peptídeo ATN-161 com alvos de integrina e de SARS-CoV-2.

O estudo é realizado com integrina  $\alpha_5\beta_1$  em conformação aberta e fechada, a fim de entender se este peptídeo interage nas duas conformações e determinar o sítio de ligação para esta interação. A interação é também estudada com a proteína S da variante ômicron complexada a hACE2 e da enzima 3CL ( $M^{PRO}$ ), tendo como propósito verificar se ocorre a interação favorável para inibição da carga viral. Estes estudos contemplam a afinidade

de ligação do ATN-161 com cada alvo, bem como a região em que cada interação ocorre e os amino ácidos que contribuem para esta interação.

## 3

### Objetivos

#### 3.1

##### Objetivo geral

Este trabalho tem como objetivo investigar o comportamento do pentapeptídeo ATN-161 ao interagir com a integrina  $\alpha_5\beta_1$  e alvos do SARS-CoV-2 utilizando métodos de ancoramento molecular e dinâmica molecular. Além disso, calcular a energia livre de ligação do ATN-161 em cada complexo formado e os aminoácidos que participam da estabilização do ligante no sítio de interação utilizando o método de amostragem guarda-chuva.

#### 3.2

##### Objetivos específicos

- ❖ Formar complexos entre ATN-161 e integrina  $\alpha_5\beta_1$  conformação fechada e aberta, proteína S da ômicron complexada a hACE2 do SARS-CoV-2 e protease principal do SARS-CoV-2.
- ❖ Obter a configuração de menor energia pelo método de ancoramento molecular.
- ❖ Analisar a estabilidade de cada complexo formado utilizando simulações de dinâmica molecular.
- ❖ Analisar o comportamento do ATN-161 no sítio de interação de cada alvo durante o tempo de simulação.
- ❖ Identificar os aminoácidos que interagem com o ATN-161 no sítio de interação de cada alvo estudado.
- ❖ Calcular a energia livre de ligação do ATN-161 com cada alvo executando simulações de amostragem guarda-chuva.
- ❖ Calcular o tempo de residência relativo do ATN-161 do sistema mais estável estudado.

## 4

### Referencial Teórico

#### 4.1

##### Peptídeo antimicrobiano

Peptídeos antimicrobianos são uma classe de peptídeos que existem amplamente na natureza e são uma parte importante do sistema imune inato de diferentes organismos (HUAN et al., 2020). Os AMP's atualmente são os candidatos mais promissores a novos fármacos por serem altamente específicos mesmo em pequenas concentrações, por apresentarem baixa toxicidade para células eucarióticas, por terem estabilidade térmica, alta solubilidade, baixo peso molecular e não desenvolverem resistência (LI et al., 2019; TONK; RŮŽEK; VILCINSKAS, 2021).

Os AMP's são amplamente encontrados na natureza com origem em bactérias, plantas, insetos, peixes, pássaros e outros animais (LUO, Y.; SONG, 2021). Essa classe de moléculas é responsável por atuar no sistema imunológico inato e na primeira linha de defesa contra infecções por patógenos (PARK, S. et al., 2001; ZASLOFF, 2002). Uma das vantagens da utilização dos AMP's é sua fácil modificação, pois são constituídos por aminoácidos que são relativamente fáceis de modificar suas estruturas para melhorar suas propriedades (AGARWAL; GABRANI, 2021; COSTA, F. et al., 2011). Além disso, é também possível obter peptídeos totalmente sintéticos por síntese química ou usando expressão recombinante (RAMOS et al., 2013).

Quando peptídeos apresentam atividade biológica contra vírus são chamados de peptídeos antivirais (AVP's) (BAHAR; REN, 2013). Além da atividade contra doenças microbianas, alguns peptídeos também apresentam atividade anticancerígena (RAMOS et al., 2013). Nessa categoria encaixa-se o ATN-161 que possui atividade para diversas patogenicidades, entre elas o câncer (CIANFROCCA et al., 2006; RIDER et al., 2011).

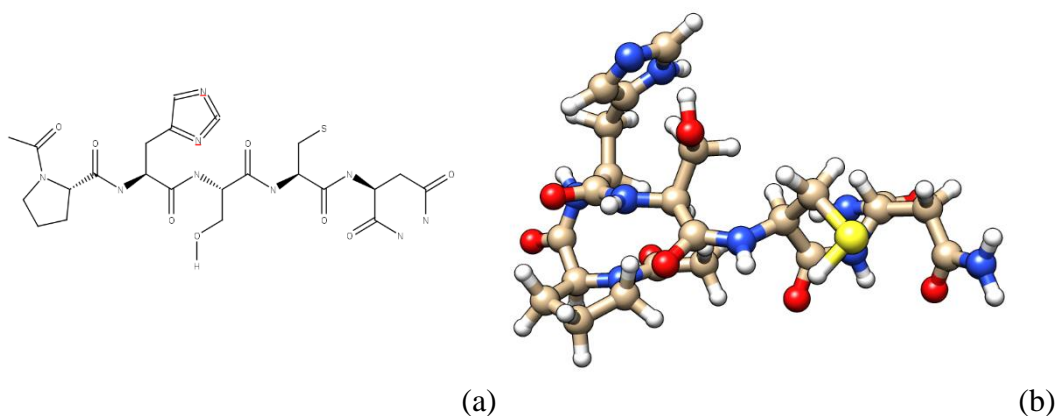
##### 4.1.1

##### ATN-161

O ATN-161 é um pentapeptídeo (Ac-PHSCN-NH<sub>3</sub>) derivado da região de sinergismo da fibronectina (PHSRN), que é responsável por potencializar a



interação com a integrina  $\alpha_5\beta_1$  (KHALILI et al., 2006). A princípio este AMP tinha PHSCN como sequência de aminoácidos, no qual alguns trabalhos o apontavam como um potencial agente tumoral (KHALILI et al., 2006;STOELTZING et al., 2003). Estudos posteriores mostraram que ele torna-se bem mais potente quando acetilado e amidado em seus terminais (Ac-PHSCN-NH<sub>3</sub>) (Figura 1) (SCHAFFNER; RAY; DONTENWILL, 2013).



**Figura 1.** Formula Estrutural do ATN-161 (a) 2D e (b) 3D. Figura extraída de (SILVA et al., 2022).

Este peptídeo possui diversos relatos na literatura para ação terapêutica: 1) no bloqueio da recorrência do tumor da próstata, metástases e micrometástases; 2) redução da metástase hepática colorretal e melhoria na sobrevida quando administrada em adição à quimioterapia; 3) bloqueio ao crescimento e metástase do câncer de mama em modelos pré-clínicos de camundongos; e 4) bloqueio eficiente da neovascularização da coroide (DHALIWAL; SHEPHERD, 2022;KHALILI et al., 2006;SCHAFFNER; RAY; DONTENWILL, 2013).

As diversas atividades farmacológicas relatadas na literatura estão particularmente associadas a interação não competitiva com as integrinas, que são moléculas relacionadas a adesão celular (STOELTZING et al., 2003). Um fator determinante para ser um agente terapêutico singular é não ser derivado do domínio de ligação RGD para não afetar a adesão celular (STOELTZING et al., 2003). Recentemente, o ATN-161 foi apontado como candidato a agente terapêutico contra o SARS-CoV-2, motivando a realização de estudos mais

aprofundados sobre o tema (AMRUTA et al., 2021;BEDDINGFIELD et al., 2021).

## **4.2**

### **Moléculas de adesão celular**

A adesão celular é uma etapa fundamental no desenvolvimento e geração de órgãos, no movimento celular e distribuição celular no corpo (GAHMBERG; GRÖNHOLM; MADHAVAN, 2022). Este grupo de moléculas está envolvido em uma gama de processos entre eles as interações células-células, células-matriz, migração celular, ciclo celular, sinalização, morfogênese e regeneração tecidual (TRZPIS et al., 2007). As moléculas de adesão celular tem ganhado um papel de destaque nas últimas décadas devido a sua associação a diversas patogenicidades, desde câncer, inflamações, infecções patogênicas até doenças autoimunes (OKEGAWA et al., 2004).

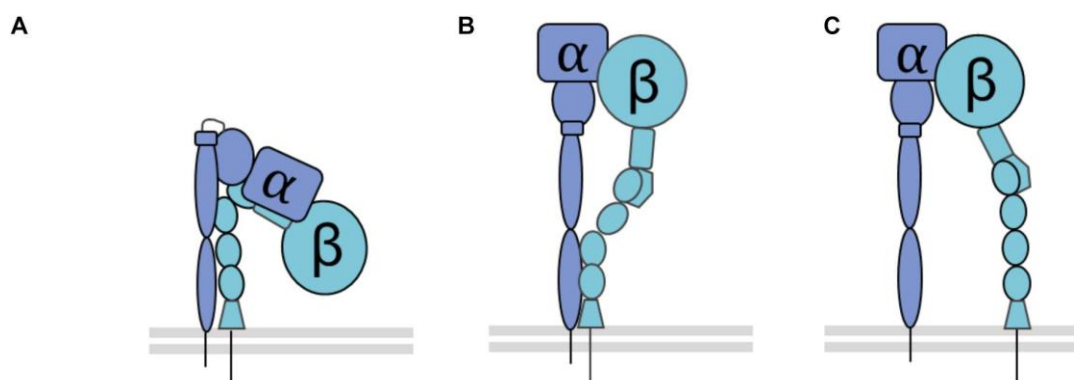
As entidades responsáveis pela adesão celular são denominadas como moléculas de adesão celular, neste grande grupo inclui-se as famílias de proteínas Selectinas, Imunoglobulinas, Caderinas e Integrinas (ELANGBAM; QUALLS; DAHLGREN, 1997;TRZPIS et al., 2007). As Selectinas são uma família de proteínas que atuam em conjunto com as Integrinas, participando de forma fundamental na migração dos leucócitos através dos vasos sanguíneos, tendo como função ajudar na fixação destes no endotélio do vaso (MCEVER, 1994). As Imunoglobulinas e caderinas são moléculas de adesão celular que ligam as células vizinhas através de interações homofílicas e heterofílicas para a primeira e apenas homofílica para segunda (ROY, VAN; BEX, 2008;WILLIAMS, 1987). Já as Integrinas são moléculas de adesão celular dependentes de cálcio ou magnésio que ligam as células à matriz extracelular (SCHAFFNER; RAY; DONTENWILL, 2013).

#### **4.2.1**

##### **Integrinas**

As Integrinas são proteínas transmembranas heterodiméricas, que possuem como função intermediar a comunicação entre o citoesqueleto celular e proteínas plasmáticas ou da matriz extracelular, por meio da adesão célula-célula (CLEGG, 2003;DARRIBÈRE et al., 2000). As Integrinas são

glicoproteínas heterodiméricas formadas por duas subunidades diferentes ligadas de forma não covalente  $\alpha$  e  $\beta$  (GAHMBERG; GRÖNHOLM; MADHAVAN, 2022). Geralmente este grupo de proteínas apresentam até 3 estados (Figura 2), sendo eles inativo (conformação fechada), intermediário e ativo (conformação aberta) (LUO, B.-H.; SPRINGER, 2006).



**Figura 2.** Estados conformacionais das integrinas. A. Conformação inativa, B. Conformação intermediária, C. conformação Ativa. Figura adaptada de (PETERSON; KOVAL, 2021)

A conformação fechada é caracterizada pela posição das cabeças da integrina voltadas para membrana lipídica (GAHMBERG; GRÖNHOLM; MADHAVAN, 2022). As integrinas podem ser ativadas 1) de dentro pra fora, quando se origina um sinal de um receptor diferente da integrina e é transmitido a elas; ou 2) de fora pra dentro quando um ligante extracelular se liga a integrina (GAHMBERG; GRÖNHOLM; MADHAVAN, 2022). Quando as integrinas são ativada ela se estendem e abrem os locais de ligação, aumentando a capacidade de ligação (LUO, B.-H.; SPRINGER, 2006).

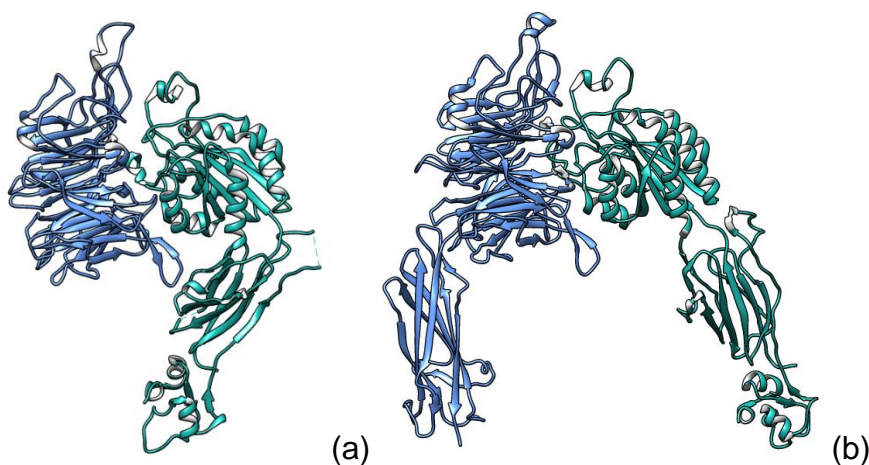
Os ligantes para integrinas podem ser divididos em quatro classes que se ligam: ao colágeno, aos leucócitos, a laminina e aos receptores RGD (TRABOCCHI; GUARNA, 2014). Os receptores RGD são a fibronectina, fibrinogênio e vitronectina. Estes são encarregados de mediar a adesão celular à matriz extracelular (ECM) (MEYER et al., 2006). A sequência Arg-Gly-Asp (RGD) é bastante utilizada na projeção de biomoléculas, tendo como intuito mimetizar a interação destes receptores com as integrinas (ASSA-MUNT et al., 2001; DANHIER; BRETON, LE; PRÉAT, 2012; HARBOTTLE et al., 1998; RUOSLAHTI, 1996; SANI et al., 2021). As integrinas tornam-se ótimos alvos por participarem de vários processos fisiológicos relevantes, como

agregação plaquetária e angiogênese, e processos patológicos como doenças cardiovasculares e formação de tumores (DANHIER; BRETON, LE; PRÉAT, 2012; DESGROSELLIER; CHERESH, 2010).

A integrina  $\alpha_5\beta_1$ , que também é conhecida como receptor de fibronectina, tem se destacado como um alvo importante em estudos para câncer. A alteração desta integrina está intimamente relacionada a progressão de vários tipos de câncer humano, incluindo proliferação celular, angiogênese, metástase tumoral e carcinogênese (HOU et al., 2020). Para que sua ação biológica seja eficiente, é necessário a funcionalização de ambas as subunidades ( $\alpha$  e  $\beta$ ). A alta afinidade e especificidade de ligação da fibronectina é proporcionada por uma interação primária através do sítio RGD e uma interação secundária no sítio de sinergismo da fibronectina (AOTA; NOMIZU; YAMADA, 1994; OBARA; KANG; YAMADA, 1988).

A integrina  $\alpha_5\beta_1$  é pró-angiogênica, e isso tem motivado a busca por novos antagonistas pretendendo entender o comportamento fisiopatológico em estudos pré-clínicos em células endoteliais e tumores—(AOUDJIT; VUORI, 2012). Dentre os antagonistas propostos é possível dividi-lo em pelo menos três principais categorias: moléculas semelhantes a RGD, moléculas não semelhantes a RGD e anticorpos (SCHAFFNER; RAY; DONTENWILL, 2013).

Estruturas tridimensionais tanto da conformação fechada como aberta da integrina  $\alpha_5\beta_1$  (Figura 3) estão disponíveis no Protein Data Bank (PDB). Isto facilita a triagem de ligantes por meio de estudos de modelagem molecular.



**Figura 3.** Representação da estrutura tridimensional da integrina  $\alpha_5\beta_1$  em modelo Cartoon. (a) conformação fechada (código PDB: 4WK4) e (b)

conformação aberta (código PDB: 7NWL). As cadeias da subunidade  $\alpha_5$  estão representadas na cor azul e  $\beta_1$  na cor verde.

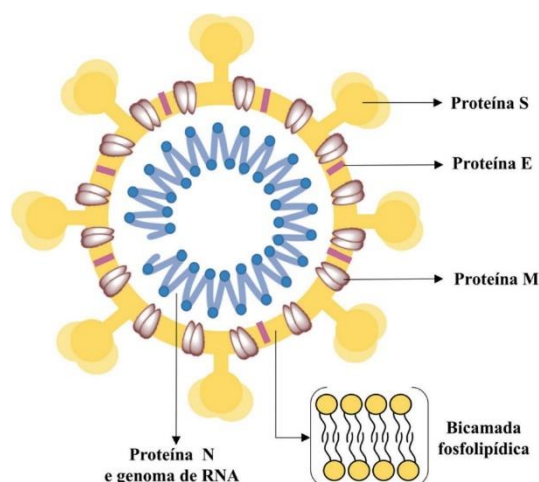
A elucidação teórica da estrutura cristalina do ectodomínio da integrina  $\alpha_5\beta_1$  tende a contribuir bastante para o projeção de antagonistas mais potentes e específicos (SCHUMACHER et al., 2021; XIA; SPRINGER, 2014).

### 4.3

#### SARS-CoV-2

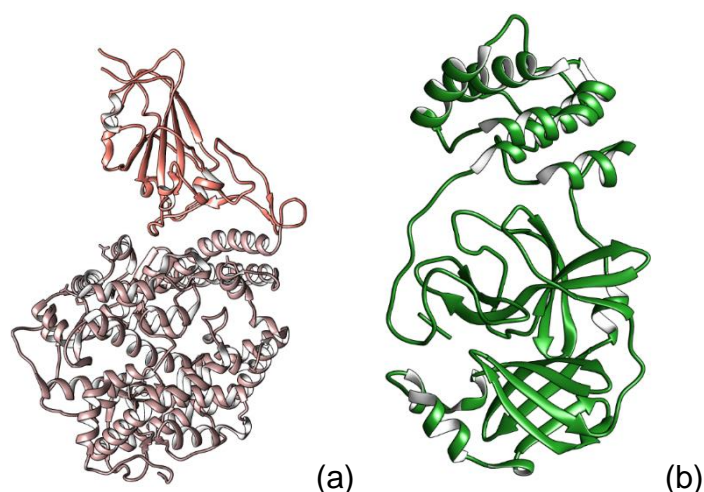
O SARS-CoV-2 é um vírus da família coronavírus, que pertence a mesma família do SARS-CoV que possui estrutura semelhante, e ao vírus da síndrome respiratória aguda do oriente médio (MERS-CoV) (WU et al., 2020). Esta família de coronavírus possui até então 7 vírus que podem ser transmitidos para humanos. O SARS-CoV-2 é aquele que apresenta maior eficiência de transmissão (MADEWELL et al., 2020). No entanto, a maior eficiência de transmissão não se traduz em maior mortalidade, sendo assim sua taxa de mortalidade é relativamente menor se comparada ao SARS-CoV e ao MERS-CoV (LIU, Z. et al., 2020; ZUMLA; HUI; PERLMAN, 2015).

O SARS-CoV-2 é um vírus de RNA de fita simples com sentido positivo (ssRNA+). Esta característica genômica o torna muito semelhante ao RNA mensageiro humano (PENG et al., 2021). Este genoma é responsável por codificar três tipos de classes de proteínas, sendo elas, duas grandes poliproteínas, pp1a e pp1ab, que são clivadas em 16 proteínas não estruturais (NSP) necessárias para síntese do RNA viral. Além disso, é responsável por codificar quatro proteínas estruturais: proteína S, proteína envelope (E), proteína de membrana (M) e proteína de nucleocapsídeo (N) (Figura 4), que são importantes para entrada e montagem viral, assim como nove proteínas acessórias que possivelmente neutralizam a imunidade do hospedeiro durante a infecção (V'KOVSKI et al., 2021).



**Figura 4.** Representação da estrutura viral do SARS-CoV-2. Onde as proteínas estruturais S (spike), E (envelope), M (membrana) e N (nucleocapsídeo) são mostradas. Figura extraída de (LIMA et al., 2020)

Entre as estruturas citadas, a proteína não estrutural nsp5 ( $M^{\text{PRO}}$ ) e a proteína estrutural S são as mais estudadas como alvo no planejamento de fármacos inibidores de SARS-CoV-2. A triagem de moléculas ativas contra o SARS-CoV-2 é possibilitada a partir do depósito de dados cristalográficos que permitem o estudo por modelagem molecular. A estrutura tridimensional da proteína S e da protease principal  $M^{\text{PRO}}$  estão disponíveis no banco de dados de proteínas (PDB), e são mostradas na Figura 5.



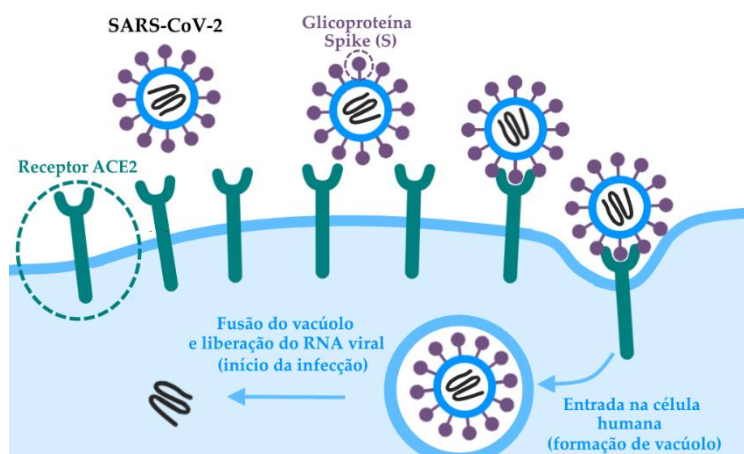
**Figura 5.** Representação tridimensional em modelo Cartoon da Estrutura Cristalográfica da (a) proteína S, na cor bege complexada a Enzima Conversora de Angiotensina Humana (hACE2) na cor marrom e (b) protease principal ( $M^{\text{PRO}}$ ) na cor verde.

As estruturas mostradas na Figura 5 podem ser obtidas pelos códigos 7WPB (YIN et al., 2022) e 7BQY (JIN et al., 2020) no banco de dados de proteínas respectivamente.

#### 4.3.1

##### Mecanismo de entrada

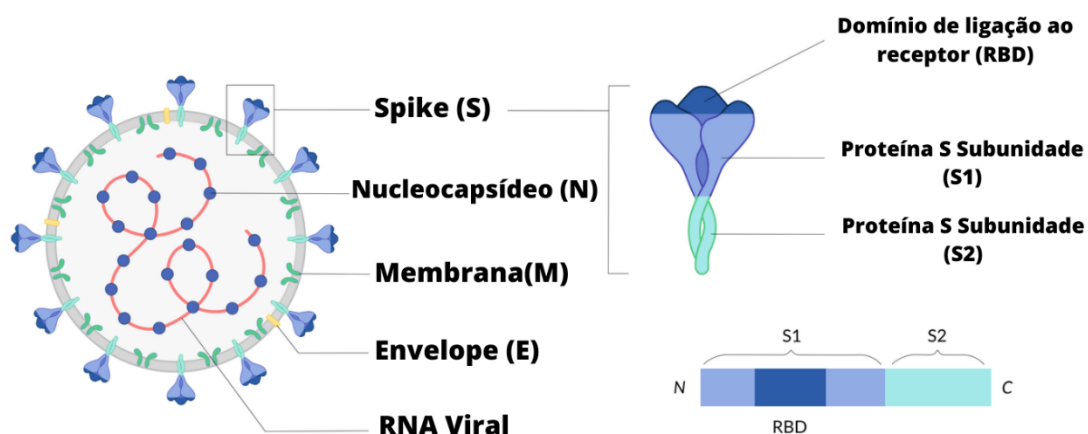
O processo de entrada é uma das etapas mais importantes para infecção viral e ciclo de vida do vírus, que é realizado pela proteína S no envelope do SARS-CoV-2, que reconhece a célula hospedeira e media a fusão da membrana para permitir que o conteúdo genômico viral seja liberado no citoplasma (BELOUZARD et al., 2012). O mecanismo de entrada está representado na Figura 3.



**Figura 6.** Representação do Mecanismo de Entrada do Virus do SARS-CoV-2 na Célula hospedeira. Figura extraída do Glossário Fiocruz.

A proteína S do SARS-CoV-2 é uma glicoproteína de fusão homotrimérica de classe I. Esta proteína pode ser dividida em duas regiões funcionais distintas, sendo elas S1 e S2, conforme mostra a Figura 6 (V'KOVSKI et al., 2021).





**Figura 7.** Representação da estrutura da Glicoproteína S do SARS-CoV-2, domínio de ligação ao receptor (RBD), e regiões das subunidades S1 e S2. Figura adaptada de Lubio Science.

A região S1 é a parte que possui o domínio de ligação ao receptor (RBD). Esta parte do trímero é responsável pela determinação do tropismo e patogenicidade da célula viral (LAN et al., 2020). A região denominada de S2 possui o domínio transmembrana e de regiões de um número de sete repetições do peptídeo de fusão (BELOUZARD et al., 2012). Estes componentes são responsáveis por mediar a fusão das membranas virais e celulares após longos rearranjos conformacionais (LETKO; MARZI; MUNSTER, 2020).

#### 4.3.2

##### Principais receptores

Após o surto do SARS-CoV nos anos de 2001 a 2003, a enzima conversora de angiotensina 2 (ACE2) foi apontada como principal receptor para entrada deste vírus no hospedeiro (LETKO; MARZI; MUNSTER, 2020). Com o surgimento do SARS-CoV-2, estudos mostraram que há uma alta conservação genômica e estrutural da proteína S do SARS-CoV e do SARS-CoV-2 (PENG et al., 2021). Os aminoácidos do RBD destas proteínas possuem identidade de 74%. Logo, assim como o SARS-CoV, o SARS-CoV-2 apresenta a ACE2 como receptor chave para entrada nas células (PENG et al., 2021). Além disso, a diferença em alguns aminoácidos chaves contribuem para um maior afinidade de ligação e transmissão mais eficiente em humanos (WANG, Q. et al., 2020).



A ACE2 é uma carboxipeptidase que possui 805 aminoácidos (JACKSON et al., 2022). A sua principal função no metabolismo é converter angiotensina I e II em angiotensina 1-9 e angiotensina 1-7 respectivamente (RIVELLESE; PREDILETTO, 2020). Uma desregulação deste processo de conversão de angiotensina II pode desencadear a liberação de uma série de citocinas pró-inflamatório, como IL-6, TNF $\alpha$  entre outras. Logo, a angiotensina II está intimamente associada a lesão pulmonar inflamatória (RIVELLESE; PREDILETTO, 2020). Portanto a regulação negativa desta enzima, com o intuito de evitar a entrada do vírus podem aumentar a gravidade da doença ao perturbar o sistema renina-angiotensina-aldosterona.

Apesar da ACE2 ter sido identificada rapidamente como receptor primário para SARS-CoV-2, e comprovadamente ter apresentado maior afinidade de ligação em comparação com SARS-CoV, estudos recentes indicam a existência de interações adicionais com outros receptores de superfície celular (MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021). Logo, sugere-se que estas interações ocorrem devido a presença do motivo RGD na proteína S do SARS-CoV-2 que não existe na SARS-CoV, o que facilitaria a entrada do vírus por vias alternativas (MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021). As integrinas são uma das vias alternativas relatadas na literatura, no qual 8 de 24 proteínas desta família reconhecem o motivo de RGD no ligantes nativos (CARVACHO; PIESCHE, 2021).

O motivo RGD já foi relatado como entrada para infecções por outros vírus, tais quais vírus de Epstein-Barr, rotavírus e Ebola (HUSSEIN et al., 2015). Portanto, a interação com integrinas pode ser uma das explicações para o SARS-CoV-2 ter uma maior eficiência na transmissão, e consequentemente se espalha de forma mais rápida e agressiva que o SARS-CoV (CARVACHO; PIESCHE, 2021). Neste sentido, as integrinas podem atuar como vias alternativas ou co-receptores para entrada na célula (MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021).

Estudos indicam que no início da infecção, quando apenas sintomas respiratórios são observados, a partícula viral liga-se as integrinas gerando uma desregulação no processo de coagulação e angiogênese (MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021). Em etapas seguintes, a replicação viral acelera e a barreira epitelial-endotelial é comprometida, levando a presença de

partículas virais no sangue em 15% dos pacientes (CHEN et al., 2020; WIERSINGA et al., 2020; ZHOU et al., 2020).

### **4.3.3**

#### **Tratamentos disponíveis**

Após a descoberta e aprovação de algumas vacinas no fim de 2020, o seu uso através da vacinação tornou-se a maneira mais eficaz de prevenir-se da infecção pelo SARS-CoV-2 (STASI et al., 2022). Atualmente o tratamento para caso graves ou críticos é baseado no uso de corticosteroides sistêmicos, tocilizumabe ou sarilumabe que são bloqueadores de receptores Interleucina-6 (IL-6) e Baricitinib como alternativa a bloqueadores de IL-6 (WHO, 2022). Para casos não graves, mas com risco de hospitalização apenas o Nirmatrelvir-Ritonavir é fortemente recomendado (WHO, 2022). Demais tratamentos aprovados pela OMS são colocados para uso condicional, considerando respostas clínicas ou casos específicos (WHO, 2022).

## **4.4**

### **Simulação molecular**

O surgimento de epidemias e outras doenças, tais como COVID-19, câncer e diabetes, ameaçam a vida e a saúde humana. Portanto, é frequentemente necessário o planejamento e desenvolvimento de novos fármacos (ADELUSI et al., 2022), que é um processo caro e longo, iniciando na identificação, passando pelo desenvolvimento, ensaios in vitro, in vivo, clínicos e finalmente aprovação. Este processo pode custar até 2,6 bilhões de dólares e leva em média 14 anos (MYERS; BAKER, 2001). Os avanços computacionais das últimas décadas possibilitam a realização de algumas etapas do desenvolvimento e planejamento de fármacos através de simulações computacionais (MIRZAEI, 2020), reduzindo o custo e o tempo.

Atualmente a técnica de modelagem molecular é uma ferramenta bem integrada ao processo de planejamento de fármacos na indústria farmacêutica, principalmente auxiliando na rápida triagem e descoberta de candidatos terapêuticos com menor custo possível (ABRAMOV; SUN; ZENG, 2022; ADELUSI et al., 2022). A técnica de modelagem molecular abrange uma grande quantidade de técnicas como ancoramento molecular, dinâmica

molecular e predições de adsorção, distribuição, metabolismo, excreção e toxicidade (ADMET) no desenvolvimento e planejamento de fármacos (ADELUSI et al., 2022).

As simulações moleculares são amplamente utilizadas em sistemas químicos e biológicos com o intuito de compreender fenômenos que geralmente não são visualizados em estudos experimentais (HERWIG, 2014; SILVA et al., 2022). Elas podem fornecer informações a níveis moleculares e atômicos das interações entre uma alvos e um ligante, entre moléculas e modelos membranas, das rotas de síntese de um certo fármaco, da transição de fase e agregação de moléculas, e de vários outros processos importantes (LYUBARTSEV; RABINOVICH, 2016; MASERAS; MOROKUMA, 1995; MOROKUMA et al., 2001; SENN; THIEL, 2009; SILVA et al., 2022).

#### **4.4.1**

##### **Ancoramento molecular**

O ancoramento molecular é uma técnica que visa obter informações sobre a interação entre componentes químicos e biológicos ou entre dois componentes biológicos, tais como interações entre uma proteína e outra proteína, entre um ácido nucleico e uma proteína e entre uma molécula pequena (ligante) e uma macromolécula (PAGADALA; SYED; TUSZYNSKI, 2017). Esta ferramenta prediz o comportamento destes componentes, contabilizando as interações entre as duas entidades estudada, gerando conformações e poses que podem ser avaliadas através de ranqueamento por funções de pontuação (GUEDES; MAGALHÃES, DE; DARDENNE, 2014).

Em sistemas alvo-ligante, esta técnica tem três objetivos principais: a predição do modo de ligação (pose), triagem virtual e estimativa de afinidade de ligação (JAIN; NICHOLLS, 2008). Estes objetivos são alcançados a partir da contabilização das interações de van der Waals, interações eletrostáticas e ligações de hidrogênio (PAGADALA; SYED; TUSZYNSKI, 2017).

As metodologias de ancoramento molecular são constituídas principalmente por duas ferramentas básicas, um algoritmo de busca e uma função de pontuação de energia (GUEDES; MAGALHÃES, DE; DARDENNE, 2014). No caso de uma molécula pequena (ligante) e uma proteína, o algoritmo de busca é responsável por explorar as orientações e conformações de um

ligante, enquanto que a função de pontuação é responsável por avaliar a qualidade destas conformações e ordená-las de acordo com a afinidade do ligante em um alvo (BROOIJMANS; KUNTZ, 2003; SPERANDIO et al., 2006).

Para ser considerada como adequada para dado sistema, a metodologia de ancoramento tem que ser capaz de prever a posição, as interações moleculares de um ligante ativo no sítio de ligação de um alvo, e reproduzir dados observados experimentalmente dentro de uma margem de tolerância (GUEDES; MAGALHÃES, DE; DARDENNE, 2014). Este procedimento de validação é comumente denominado de reancoragem, e sua margem de tolerância geralmente é medida pela raiz quadrada do desvio quadrático médio (RMSD) (MORRIS et al., 2009; 2014).

#### **4.4.1.1**

##### ***Autodock***

O software AutoDock é dentre muitas ferramentas, uma ferramenta amplamente de ancoragem molecular utilizada na previsão de complexos proteína-ligante (GARRETT M et al., 2012). O Autodock mostrou-se capaz de prever com rapidez e precisão modos de ligação e afinidade de ligação para sistemas de proteínas com ligantes (MORRIS et al., 2009). Esta técnica combina um campo de força de energia livre semiempírica e um algoritmo de busca lamarckiano para gerar resultados com rapidez para cada modo de ligação e afinidade associada (MORRIS et al., 1998a).

Para prever os modos de ligação, o Autodock usa um método baseado em grades para compreender todo espaço conformacional disponível para o ligante na proteína (MORRIS et al., 2009). Neste método, a proteína é incorporada a uma grade, em seguida cada ponto da grade é sequenciado por um átomo de sonda (MORRIS et al., 2014). Então, a energia de interação entre a sonda e o alvo é calculada. Este valor é armazenado na grade e por fim esta grade de energias calculadas é usada durante a simulação (MORRIS et al., 2014).

O algoritmo de busca conformacional utilizado neste método é baseado de maneira análoga a evolução biológica (MORRIS et al., 1998a). O Algoritmo Genético Lamarckiano (LGA) considera uma população inicial, onde as conformações experimentais são criadas e, em gerações sucessivas esses

indivíduos sofrem mutações e trocam parâmetros conformacionais. Por fim, o algoritmo seleciona os indivíduos de menor energia de ligação (MORRIS et al., 2009). O uso do LGA permite uma busca local de conformações, encontrando mínimos locais que são passados para gerações posteriores. Logo, o algoritmo de busca LGA é um método híbrido em que realiza uma busca global e local (MORRIS et al., 1998a).

Para prever a energia de ligação, o módulo Autodock4 usa um campo de força de energia livre semiempírico (MORRIS et al., 2014). O campo de força utilizado é baseado em um modelo termodinâmico amplo que permite a incorporação de energias intramoleculares na energia de ligação prevista (HUEY et al., 2007). Nesta abordagem, a energia de ligação é estimada pela diferença entre a energia do ligante e da proteína no estado separado, e a energia do ligante no estado complexado (HUEY et al., 2007; MORRIS et al., 2009).

As funções de pontuação atuais são otimizadas pela adição de funções de energia empírica, na qual são adicionados termos entrópicos as equações de mecânicas moleculares. A equação 1 principal está representada abaixo:

$$\Delta E_{bind} = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L}) + \Delta S_{conf} \quad \text{Eq. 1}$$

onde L e P referem-se ao ligante e à proteína. V são as energias associadas aos pares estabelecidos pelo método.  $\Delta S_{conf}$  é a perda de entropia conformacional após a ligação do ligante à proteína, que é calculada a partir da soma dos graus de liberdade de torção, da contribuição da água (dispersão/repulsão, ligação de hidrogênio, eletrostática e dessolvatação) e de aproximações de dados empíricos em termos de pares atômicos conforme apresentado na equação 2 (HUEY et al., 2007; Morris et al., 2009):

$$V = W_{vdw} \sum_{ij} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hbound} \sum_{ij} E(t) \left( \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{ij} \frac{q_i q_j}{\epsilon (r_{ij}) r_{ij}} + W_{sol} \sum_{ij} (S_i V_j + S_j V_i) e^{(-r_{ij}^2 / 2\sigma^2)}$$

**Eq. 2**

onde  $W_{vdw}$ ,  $W_{bond}$ ,  $W_{elec}$  e  $W_{sol}$  são os termos de ponderação,  $i$  e  $j$  são os pares atômicos,  $A_{ij}$ ,  $B_{ij}$ ,  $C_{ij}$  e  $D_{ij}$  são os parâmetros de van der Waals,  $r_{ij}$  é a distância entre os pares atômicos  $i$  e  $j$ ,  $\epsilon_{ij}$  é a permissividade elétrica do meio onde o campo existe,  $q_i$  e  $q_j$  são as cargas dos pares atômicos  $i$  e  $j$ ,  $S$  é o parâmetro de solvatação e  $\sigma$  é uma constante. As constantes de ponderação  $W$  são usadas para calibrar a energia livre empírica com base em um conjunto de complexos caracterizados experimentalmente (HUEY et al., 2007; Morris et al., 2009).

#### 4.4.2

##### Dinâmica molecular

A dinâmica molecular (DM) é um ramo da ciência baseada na mecânica clássica que estuda o comportamento de um sistema de muitas partículas com o decorrer do tempo (KARPLUS; MCCAMMON, 2002). O estudo da DM possibilita obtenção de informações sobre o comportamento dinâmico microscópico do sistema. Além disso, com uso de mecânica estatística é possível obter propriedades macroscópicas como energia interna, pressão, volume, temperatura, entropia, energia livre e etc. (NAMBA; SILVA; SILVA, 2008).

As simulações de DM aplicadas a sistemas biomoleculares são capazes de descrever uma grande variedade de processos importantes, entre eles: a mudança conformacional, a ligação de moléculas com alvos macromoleculares, o dobramento e a desnaturação de proteínas, bem como prever perturbações de nível atômico tais como mutação e fosforilação (HOLLINGSWORTH; DROR, 2018). Além disso, pode-se evidenciar a posição de cada átomo em uma resolução de femtosegundos.

A DM é baseada em equações de mecânica molecular (MM), onde as moléculas são consideradas como um conjunto de átomos que podem ser descritos por forças newtonianas, ou seja, os átomos são tratados como partículas que são conectadas por forças elásticas ou harmônicas (SANT'ANNA, 2009). A simulação de DM consiste na resolução numérica das equações clássicas do movimento. Em um sistema simples, a equação de movimento pode ser escrita para cada átomo  $i$  de um sistema molecular da seguinte forma (equação 3):

$$F_i = m_i a_i = \frac{d^2 \vec{r}_i}{dt^2} m_i = \frac{d^2 \vec{v}_i}{dt} m_i \quad \text{Eq. 3}$$

onde  $F_i$  é a força sobre o átomo  $i$  na posição  $r_i$ ,  $t$  é o tempo,  $v_i$  e  $a_i$  são a velocidade e a aceleração do átomo  $i$  de massa  $m_i$ . A força também pode ser obtida diretamente da derivada da energia potencial  $V$  em relação as coordenadas  $r_i$  dos átomos conforme a equação 4:

$$F_i = \frac{\partial V}{\partial \vec{r}_i} \quad \text{Eq. 4}$$

A resolução da equação 4 fornece diretamente a aceleração de cada átomo. Associando as equações de movimento é possível determinar as velocidades, a partir das quais podem ser integradas para obter a mudança de posição de cada átomo. Com as posições e velocidades de cada átomo, pode-se determinar as energias potencial e cinética do sistema.

Desta forma, a resolução destas equações de movimento representa a evolução no tempo dos movimentos moleculares, denominada de trajetória, que é o conjunto de posições e velocidades de cada átomo com o decorrer do tempo (NAMBA; SILVA; SILVA, 2008; SANT'ANNA, 2009).

#### 4.4.2.1

##### Campo de Força

Nas simulações de DM, os átomos são considerados como centros pontuais que interagem através de potenciais de interação de pares ou multicorpos. Um conjunto completo de potenciais de interação entre as partículas é denominado de campos de força (NAMBA; SILVA; SILVA, 2008). Nos dias de hoje, a aplicação do campo de força torna-se necessária quando pretende-se executar simulações de sistemas grandes (com mais de 100.000 átomos) com escala espacial de dezenas nanômetros e resolução temporal de até microssegundos. (BAI et al., 2022; FU, H. et al., 2022).

Os primeiros campos de força surgiram por volta de 1960. A principal finalidade era o estudo da estrutura, espectros vibracionais e entalpia de pequenas moléculas orgânicas (hidrocarbonetos, álcoois, éter e etc.) (GONZÁLEZ, M.A., 2011). Com o avanço da computação, tanto em questão de software como de hardware, diversos campos de força têm surgido atendendo sistema mais complexos e específicos. Entre os campos de força mais

populares estão o AMBER (CORNELL et al., 1995), CHARMM (HUANG; MACKERELL, 2013; MACKERELL et al., 1998), OPLS (JORGENSEN; MAXWELL; TIRADO-RIVES, 1996), GROMOS (OOSTENBRINK et al., 2004) e COMPASS (SUN, 1998). Um fato que pode potencializar a popularidade destes campos de força é a aplicabilidade geral. Cabe ressaltar que a maioria destes campos de força são muitas vezes específicos para um tipo de sistema e são frequentemente atualizados constantemente.

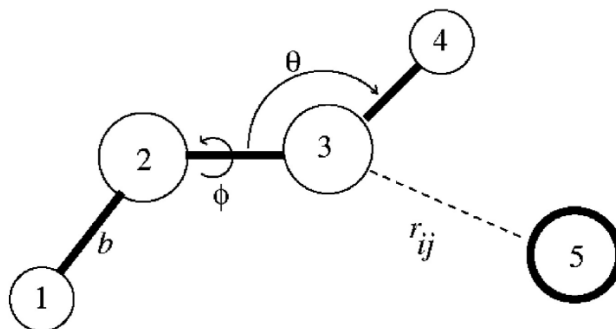
Apesar dos campos de força gerais atender uma variedade de sistemas, algumas classes de compostos e sistemas específicos requerem potenciais específicos tais como carboidratos e ácido ribonucleicos (FOLEY; TESSIER; WOODS, 2012; SALSBURY; LEMKUL, 2021). Dentre essas classes, deve se considerar a água como de grande importância, tal que vários modelos de água foram propostos para otimizar as simulações de DM (BARKER; WATTS, 1969). Os modelos de água mais comuns são: TIP3P, TIP4P, TIP5P, SPC e SPC/E (VEGA et al., 2009). Alguns estudos que realizam comparações entre modelos indicam que o modelo de água TIP4P, fornece a melhor descrição de 9 de 10 propriedades experimentais investigadas (VEGA et al., 2009). Uma das propriedades mais desafiadoras para reconciliar os dados experimentais e teóricos com diversos campos de força e modelos de água diferentes é o coeficiente de partição água-octanol (COSTA, R. K. M. et al., 2023).

Para que um campo de força seja considerado ideal, este deve ter parâmetros que consigam reproduzir com eficiência as propriedades de interesse do sistema estudado. Portanto, existem vários campos de forças com complexidades diversificadas, buscando atender a diferentes sistemas com características específicas (GONZÁLEZ, M.A., 2011). Os parâmetros geralmente são obtidos através de cálculos de mecânica quântica *ab initio* ou semiempírico, bem como utilizando informações experimentais, tais como raios-X e difração de elétrons, ressonância magnética nuclear, infravermelho, Raman e espectroscopia de nêutrons, entre outros.

Um campo de força é uma expressão matemática, que refere-se à forma funcional e conjunto de parâmetros que associados podem calcular a energia potencial de um sistema às coordenadas de suas partículas (CORNELL et al., 1995; GONZÁLEZ, M.A., 2011). Os campos de força clássicos são baseados na interpretação física de 3 termos de energia potencial para átomos ligados:



estiramento de ligação, deformação angular da ligação e de torção em torno de ângulos diedros. Em termos de átomos não ligados, existem 2 termos de energia potencial em relação as: interações de van der Waals de repulsão e dispersão; e as interações eletrostáticas ou de Coulomb (MONTICELLI; TIELEMAN, 2013).



**Figura 8.** Esquema das interações em um campo de força. As ligações covalentes são representadas por linhas sólidas e as interações intermoleculares são indicadas como linhas tracejadas.  $b$  representa o estiramento da ligação entre as partículas 1 e 2,  $\phi$  são as torções das ligações que envolve as partículas 1, 2, 3 e 4,  $\theta$  representa as deformações angulares entre as partículas 2, 3 e 4, e  $r_{ij}$  representa as interações intermoleculares de van der Waals e Coulomb entre a partícula externa 5 e partícula 3. Adaptada de (PONDER; CASE, 2003).

Uma expressão típica para energia potencial de um sistema molecular simples é apresentada na equação 5:

$$V(r) = \sum V_b + \sum V_\theta + \sum V_\phi + \sum V_{vdW} + \sum V_{elet}$$

**Eq. 5**

onde  $V_b$  é o potencial do estiramento de ligações,  $V_\theta$  é o potencial da deformação angular da ligação,  $V_\phi$  é o potencial de torção de ligação,  $V_{vdW}$  é o potencial de interações de van der Waals e  $V_{elet}$  é o potencial de interações eletrostáticas ou Coulombianas. Cada termo da equação 5 será detalhado na próxima seção.

#### 4.4.2.1.1

## Estiramento de ligação

A energia potencial para o alongamento da ligação geralmente é representada por um potencial harmônico, conforme mostra a equação 6:

$$V(b) = \sum \frac{k_b}{2} (b - b_0)^2 \quad \text{Eq. 6}$$

onde  $k_b$  é constante de força da ligação,  $b$  e  $b_0$  são comprimento de ligação e comprimento de ligação de referência, respectivamente. Cabe dizer que geralmente  $k_b$  tem um valor grande, o que se traduz em uma grande quantidade de energia para esticar ou comprimir significativamente uma ligação química (MONTICELLI; TIELEMAN, 2013).

O potencial harmônico é uma aproximação ruim quando há grandes desvios no comprimento de ligação de referência (GONZÁLEZ, M.A., 2011). Ocasionalmente, outras formas funcionais são empregadas para melhorar a precisão nestes casos, em particular o potencial de Morse apresenta boa precisão (MONTICELLI; TIELEMAN, 2013). Como grandes deslocamentos de ligação de referência são incomuns em simulações de macromoléculas biológicas, funções de energia potencial desse tipo raramente são usadas.

### 4.4.2.1.2

#### Deformação angular de ligação

A energia potencial para a deformação angular da ligação também é descrita por um potencial harmônica (equação 7):

$$V(\theta) = \sum \frac{k_\theta}{2} (\theta - \theta_0)^2 \quad \text{Eq. 7}$$

onde  $k_\theta$  é a constante de força da deformação angular,  $\theta$  e  $\theta_0$  são a variação angular e o ângulo no equilíbrio, respectivamente.

Diferente da constante de força para comprimento de ligação,  $k_\theta$  geralmente possui valores menores em comparação com  $k_d$ . Logo, é necessário menos energia para que ocorra uma variação no ângulo de ligação em relação ao valor de referência (MONTICELLI; TIELEMAN, 2013). Outros termos podem ser adicionados para melhorar a precisão em casos particulares. Entretanto, esta é uma boa aproximação no geral.

#### 4.4.2.1.3

##### Torsões de ligação

Sempre que uma molécula contém mais de 4 átomos consecutivos é necessário adicionar o termo de torção de ligação. Este termo descreve a energia potencial do sistema molecular em relação a rotação em torno de cada ângulo diedro (MONTICELLI; TIELEMAN, 2013). Na função da energia potencial de deformação angular da ligação apresentada na equação 8, o primeiro somatório refere-se a energia potencial de torção e o segundo somatório refere-se a energia potencial de torções impróprias:

$$V(\phi) = \sum \frac{k_{\phi}}{2} (1 + \cos(n\phi - \phi_0))^2 + \sum \frac{k_{\psi}}{2} (\psi - \psi_0)^2 \quad \text{Eq. 8}$$

onde  $k_{\phi}$  representa a constante do diedro que está relacionada com a altura da barreira,  $n$  é a multiplicidade ou número de mínimos da função potencial,  $\phi_0$  é o ângulo diedro de referência. No segundo termo temos a  $k_{\psi}$  que é a constante do diedro improprio e  $\psi_0$  é o ângulo diedro improprio de referência.

As energias envolvidas nas torções são significativamente menores em relação ao comprimento de ligação e deformação angular, o que lhe confere centenas de vezes menos rigidez (GONZÁLEZ, M.A., 2011). Logo, os termos torcionais ou diedros são muito importantes para assegurar a conformação correta de moléculas e para reproduzir as principais mudanças conformacionais ocasionadas pelas rotações sobre as ligações.

#### 4.4.2.1.4

##### Potencial de van der Waals

O potencial de van der Waals é um dos termos de átomos não ligados. Ele deve ser calculado para pares de átomos separados por no mínimo 3 ligações e entre átomos de diferentes moléculas (MACKERELL, 2004). Geralmente, esse termo é representado pelo potencial de Lennard-Jones (LJ), como mostra a equação 9 para os átomos  $i$  e  $j$ :

$$V_{vdW} = \sum 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad \text{Eq. 9}$$

onde  $\varepsilon_{ij}$  representa a profundidade do poço, ou seja, o potencial entre barreira atrativa e repulsiva, o termo  $\sigma_{ij}$  é o diâmetro ou distância na qual o potencial tende a zero.

O  $V_{vdW}$  possui um termo de atração e um termo repulsivo. O primeiro tem como origem as forças de dispersão, que são geradas entre dipolos induzidos que nascem de variações na distribuição de cargas eletrônicas em todas as moléculas (GONZÁLEZ, M.A., 2011). Já o termo repulsivo surge da observação de que abaixo de uma certa distância, geralmente em torno de 0,3 nm, os átomos tendem a se repelir, devido a sobreposição das nuvens de elétrons de cada átomo (GONZÁLEZ, M.A., 2011; MONTICELLI; TIELEMAN, 2013).

#### 4.4.2.1.5

##### Potencial eletrostático

O último termo da equação 5 é do potencial eletrostático, onde comumente é atribuído uma carga atômica parcial ao núcleo do átomo. Geralmente, a lei de Coulomb é utilizada para calcular esta contribuição conforme mostra equação 10 (MONTICELLI; TIELEMAN, 2013):

$$V_{elet} = \sum \frac{q_i \times q_j}{4\varepsilon_0 \varepsilon_r r_{ij}} \quad \text{Eq. 10}$$

onde  $q_i$  e  $q_j$  representam as cargas parciais para os átomos  $i$  e  $j$ ,  $r_{ij}$  é a distância entre os núcleos dos átomos,  $\varepsilon_0$  é a permissividade do vácuo e  $\varepsilon_r$  é a constante dielétrica relativa do meio.

Uma alternativa mais precisa para calcular a contribuição eletrostática é o uso de métodos quânticos, pois pode-se calcular este potencial diretamente da densidade eletrônica. Porém, o alto custo computacional no uso destes métodos para simulações de dinâmica molecular torna a aplicação inviável (GONZÁLEZ, M.A., 2011).

#### 4.4.2.2

##### Condições iniciais do sistema em dinâmica molecular

Para executar uma simulação de DM é necessário especificar as posições iniciais das partículas que fazem parte do sistema. Portanto, é conveniente que as coordenadas iniciais estejam situadas dentro de uma rede

cristalina, evitando assim que haja sobreposições atômicas indesejadas (NAMBA; SILVA; SILVA, 2008). Cabe ressaltar que podem ser usadas diferentes geometrias para a caixa de simulação, e que estas não influenciam ou interferem significativamente no resultado da simulação (NAMBA; SILVA; SILVA, 2008). As geometrias mais utilizadas são cúbicas, octaédricas truncada e paralelepipedica.

Após a solvatação do sistema, sempre é necessário minimizar a energia do sistema para eliminar maus contatos entre os átomos. A minimização de energia visa encontrar um conjunto de posições atômicas que minimizam a energia potencial do sistema. Os métodos mais utilizados são o algoritmo de minimização de descida mais íngreme (do termo em inglês *steepest descent*) e o método dos gradientes conjugados (LINDAHL, 2008). Após a minimização de energia, há um relaxamento do sistema, no qual a força aplicada sobre cada átomo é pequena. Desta forma, obtém-se a estrutura de partida para a simulação (NAMBA; SILVA; SILVA, 2008).

#### 4.4.2.2.1

##### Algoritmo de descida mais íngreme

O método de descida mais íngreme é um dos mais usados por ser considerado um método robusto. Este método utiliza a primeira derivada da função de energia para determinar a direção para o mínimo. Apesar de não ser particularmente eficiente, esta técnica tem um bom desempenho na minimização de estruturas que estão longe do mínimo de energia (TONNER, 2011). Este algoritmo utiliza a equação 11 para encontrar o mínimo local de um sistema:

$$x_{k+1} = x_k - \alpha_k g_k \quad \text{Eq. 11}$$

onde  $x_k$  é o ponto atual,  $g_k$  é a direção do gradiente e  $\alpha_k$  é o tamanho do passo. O cálculo do próximo ponto  $x_{k+1}$  é feito a cada iteração pela equação 11 até que seja encontrado um mínimo local da função (MEZA, 2010). No software Gromacs, o método de descida mais íngreme é implementado pela equação 12:

$$r_{n+1} = r_n + \frac{F_n}{\max(|F_n|)} h_n \quad \text{Eq. 12}$$

onde  $r$  é um vetor definido por  $3N$  coordenadas,  $F_n$  é a força e  $h_n$  é o deslocamento máximo. A notação de  $\max(|F_n|)$  significa a maior força que pode ser aplicada a um átomo. Com isso as energias e forças são calculadas até que exista uma convergência com a força máxima definida.

#### 4.4.2.2.2

#### Método de gradientes conjugados

Em comparação com o método de descida mais íngreme, a minimização por gradientes conjugados é mais robusta para encontrar o mínimo. Porém, é um cálculo mais demorado nos estágios iniciais da minimização. Este método também usa a primeira derivada para determinar o mínimo de energia. A equação principal para o método de gradientes (equação 13) é semelhante a equação 10, mas são consideradas as seguintes condições (GILBERT; NOCEDAL, 1992):

$$x_{k+1} = x_k - \alpha_k d_k \quad \text{Eq 13.}$$

$$d_k = \begin{cases} -g_k & \text{para } k = 1 \\ -g_k + \beta_k d_{k-1} & \text{para } k \geq 2 \end{cases}$$

onde  $\beta_k$  é um escalar e  $\alpha_k$  é um passo obtido por uma busca unidimensional. Como dito anteriormente, este método apresenta algumas limitações na busca por mínimos globais, sendo necessárias minimizações anteriores.

#### 4.4.2.3

#### Ensembles

Um estado macroscópico é consequência da associação de vários microestados, que interagem para formarem um equilíbrio dinâmico estatístico (LUSTIG, 1994). Logo, um ensemble é um conjunto de microestados que representa o estado do sistema. Ele permite que um conjunto de propriedades termodinâmicas e configurações seja mantido constante durante a integração das equações de Newton (TOBIAS; MARTYNA; KLEIN, 1993).

O ensemble microcanônico (NVE) mantém constante o número de partículas ( $N$ ), o volume ( $V$ ) e a energia total do sistema ( $E$ ) durante a simulação (NAMBA; SILVA; SILVA, 2008). Porém, este ensemble possui a desvantagem de que algumas propriedades experimentais não são mensuráveis através de

estudos por simulação (TOBIAS; MARTYNA; KLEIN, 1993). Geralmente, estudos experimentais são feitos sob a condição de número de partículas, temperatura, volume ou pressão constantes, o que levou ao desenvolvimento de novos ensembles.

Para controlar de forma separada a temperatura (T), o volume (V) e a pressão (P), foram criados o ensemble canônico e o isotérmico-isobárico. O ensemble canônico (NVT) controla o número de partículas (N), Volume (V) e Temperatura (T) constantes durante a simulação. Já o ensemble isotérmico-isobárico (NPT) mantém constantes o número de partículas (N), a pressão (P) e a temperatura (T). Os ensembles NVT e NPT tem como potencial termodinâmico característico a energia livre de Helmholtz e energia livre de Gibbs, respectivamente, que são bem mais fáceis de serem calculados por simulação do que a entropia, que é a propriedade termodinâmica característica do ensemble NVE (TOBIAS; MARTYNA; KLEIN, 1993).

#### **4.4.2.4**

##### **Termostatos**

As simulações de DM necessitam do controle de algumas variáveis no ensemble, entre elas a temperatura no NVT e a temperatura e a pressão no NPT. Para esta finalidade, usa-se uma ferramenta chamada de termostato que tem como função realizar o controle da temperatura do sistema. O termostato impede que a temperatura do sistema varie acima dos limites pré-definidos. Os termostatos mais comuns em simulações de DM são o de Berendsen (BERENDSEN et al., 1984), Nosé-Hoover (EVANS; HOLIAN, 1985) e o de reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007).

##### **4.4.2.4.1**

##### **Termostato de Berendsen**

O termostato de Berendsen controla a temperatura do sistema através de um banho térmico que é acoplado ao sistema, onde a temperatura  $T_0$  é mantida fixa (BERENDSEN et al., 1984). A cinética deste termostato é de primeira ordem conforme mostra a equação 14:

$$\frac{dT}{dt} = \frac{1}{\tau} (T_0 - T) \quad \text{Eq. 14}$$

onde  $\tau$  é o tempo de cada passo da simulação ou constante de acoplamento,  $T$  é a temperatura do sistema no tempo  $t$ . Este termostato aplica uma renormalização das velocidades em cada passo da simulação para controlar a temperatura até que a temperatura desejada seja alcançada (BERENDSEN et al., 1984). O valor de  $\tau$  deve ser testado para averiguar a rapidez com que o sistema chega à temperatura desejada. Cabe ressaltar que quanto maior o valor de  $\tau$  mais tempo levará para a temperatura ser alcançada. Embora tenha uma forma simples, geralmente este não é o melhor termostato, pois por vezes não simula a temperatura corretamente.

#### 4.4.2.4.2

##### Termostato de Reescalonamento de Velocidade

O termostato de reescalonamento de velocidade é uma otimização do termostato de Berendsen. Uma das desvantagens ou falha no termostato de Berendsen é que as flutuações presentes no sistema não satisfazem o ensemble canônico (BERENDSEN et al., 1984; BUSSI; DONADIO; PARRINELLO, 2007). Logo, este termostato tem como proposta corrigir as flutuações do NVT através de flutuações estocásticas, conforme mostra a equação 15:

$$dK = (K_0 - K) \frac{dt}{\tau} + \left( \frac{KK_0^{\frac{1}{2}}}{N_l} \right) \frac{dW}{(\tau)^{\frac{1}{2}}} \quad \text{Eq. 15}$$

onde  $K$  representa a energia cinética,  $N_l$  corresponde aos graus de liberdade,  $dW$  representa um processo de Wiener, ou seja, uma distribuição de movimento browniano, e  $\tau$  é a constante de acoplamento (BUSSI; DONADIO; PARRINELLO, 2007).

#### 4.4.2.4.3

##### Termostato de Nosé-Hoover

O termostato de Nosé-Hoover é um algoritmo que utiliza a adição de uma variável dinâmica imaginária. Ele atua modificando a cinética das partículas do sistema, aumentando ou diminuindo a aceleração até que estas atinjam a temperatura definida (EVANS; HOLIAN, 1985). Portanto, uma variável  $s$  e o seu momento conjugado são adicionados as equações de movimento. Estas



variáveis representam um reservatório térmico através da energia potencial  $U_s$ , e cinética  $K_s$ .

$$U_s = (3N - 2)k_b t \ln(s) \quad \text{Eq. 16}$$

$$K_s = \frac{Q_s^2}{2} \quad \text{Eq. 17}$$

onde o grau de liberdade é dado por  $(3N-2)$ ,  $k$  representa a constante de Boltzmann e  $Q_s$  é a inércia térmica. A equação clássica (equação 18) do termostato de Nosé-Hoover é uma função de Lagrange estendida para incluir as coordenadas de  $s$  e sua derivada do tempo.

$$\mathcal{L} = s^2 \sum_{i=1}^N \frac{p_i^2}{2m_i} - U + \frac{Q^2}{2} - k_b T L \ln s \quad \text{Eq. 18}$$

onde  $Q$  é a massa efetiva associada a  $s$ ,  $L$  é definido pelo número de graus de liberdade. Este termostato não é ergódico para um oscilador harmônico. Logo, um termostato ou cadeias de termostatos podem ser implementados para melhorar a ergodicidade e levar em conta leis de conservação adicionais (MARTYNA; KLEIN; TUCKERMAN, 1992).

#### 4.4.2.5

##### Barostatos

Assim como a temperatura, a pressão também pode ser controlada em um ensemble. Neste caso, barostatos são usados para manter a pressão constante e próxima da pressão desejada. O barostato usa a variável  $P$ . Os barostatos mais comuns são o de Berendsen (BERENDSEN et al., 1984) e o de Parrinello-Rahman (PARRINELLO; RAHMAN, 1981).

#### 4.4.2.5.1

##### Barostato de Berendsen

A funcionalização deste barostato ocorre pelo fraco acoplamento de um banho térmico externo, no qual é fundamentado no princípio da menor perturbação local. Semelhante ao termostato de Nosé-Hoover, um termo extra é adicionado às equações de movimento que ocasionam uma mudança de pressão no sistema. A função que representa este barostato é apresentada na equação 19:

$$\frac{dP}{dt} = \frac{p_0 - p}{\tau_p} \quad \text{Eq. 19}$$

onde a pressão de referência é  $p_0$ ,  $p$  é a pressão no tempo  $t$  e  $\tau_p$  é a constante de tempo para o acoplamento. Cabe ressaltar que o barostato de Berendsen trabalha com o escalonamento de coordenadas com o escalonamento do volume para minimizar os distúrbios locais (BERENDSEN et al., 1984). Logo, é adicionando um termo extra ( $x'$ ) a equação de movimento  $x=v$  conforme mostrado na equação 20:

$$x' = v + \alpha x \quad \text{Eq. 20}$$

O ajuste do volume ocorre pela seguinte equação 21:

$$V' = 3\alpha V \quad \text{Eq. 21}$$

Já a mudança de pressão está relacionada com a compressibilidade isotérmica  $\beta$  conforme a equação 22:

$$\frac{dp}{dt} = \frac{1}{\beta V} - \frac{dV}{dt} = -\frac{3\alpha}{\beta} \quad \text{Eq. 22}$$

O fator  $\alpha$  é determinado pela equação 23:

$$\alpha = \frac{-\beta(p_0 - p)}{3\tau_p} \quad \text{Eq. 23}$$

Logo, a equação de movimento (equação 20) modificada fica da seguinte forma combinando as equações anteriores.

$$x' = v - \frac{\beta(p_0 - p)}{3\tau_p} x \quad \text{Eq. 24}$$

Apesar da compressibilidade é muito imprecisa, ela não apresenta consequência para dinâmica (BERENDSEN et al., 1984) uma vez que  $\beta$  influencia apenas a precisão de  $\tau_p$ .

#### 4.4.2.5.1

##### Barostato de Parrinello-Rahman

O barostato de Parrinello-Rahman foi proposto pois não é suficiente dimensionar apenas o volume quando estruturas cristalinas são simuladas.

Este barostato foi projetado como uma extensão do método de Andersen, de modo que tanto o volume quanto a forma da caixa de simulação mudam com o tempo (KE et al., 2022;PARRINELLO; RAHMAN, 1981). A expressão para o barostato de Parrinello-Rahman é demonstrada na equação 25:

$$\frac{d^2 r_i}{dt^2} = V W^{-1} b^{-1} (P - P_{ref}) \quad \text{Eq. 25}$$

onde  $b$  é a matriz que representa os vetores da caixa,  $V$  é o volume da Caixa, as pressões de referência e no tempo  $t$  são representadas por  $P$  e  $P_{ref}$ , respectivamente.  $W$  é uma matriz de parâmetros de massa que estabelece a força do acoplamento, que é definida conforme a equação 26:

$$W = \frac{3\tau_p^2 L}{4\pi^2 \beta} \quad \text{Eq. 26}$$

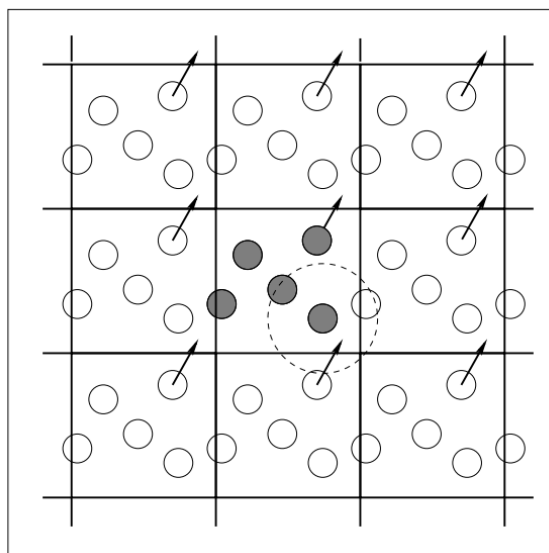
onde  $\tau_p$  é a constante de tempo do barostato,  $L$  é o maior comprimento da caixa, e  $\beta$  é a compressibilidade isotérmica. Cabe dizer que este barostato representa um ensemble NPT corretamente. Apesar disso, quando o sistema se desvia muito do equilíbrio, pode haver oscilações que são capazes de levar o sistema ao colapso (PARRINELLO; RAHMAN, 1981). Uma das saídas para este problema é aumentar o valor da  $\tau_p$  e ao simular o ensemble canônico usar o termostato de reescalonamento de velocidade, atingindo o equilíbrio com maior eficiência.

#### 4.4.2.6

##### Condições periódicas de contorno

As condições periódicas de contornos são uma ferramenta utilizada em sistemas simulados por DM quando não há interesse em algum efeito particular da superfície. Geralmente, simulações de DM possuem entre  $10^2$  a  $10^6$  partículas devido aos custos computacionais. Em comparação com sistemas reais (macroscópicos), essa quantidade de partículas é uma quantidade insignificante (NAMBA; SILVA; SILVA, 2008). Considerado um sistema com 1000 partículas em uma caixa cubica de aresta 10, quase metade das partículas estão nas faces externas, o que ocasiona um grande efeito nas propriedades mensuráveis neste sistema (ALLEN.P, 2004).

Existem várias maneiras para solucionar os problemas gerados pelas condições de contorno em uma simulação de DM. O uso das condições periódicas de contorno é um dos métodos utilizados para solucionar os problemas (KUMAR, H.; MAITI, 2011). Neste caso, a célula é replicada em todas as direções conforme mostra a Figura 9.



**Figura 9.** Representação de um sistema com condições periódicas de contorno. A figura central é o sistema estudado e nas adjacências os sistemas espaciais, a seta indica que o movimento executado na figura central é repetida nas imagens espaciais. Figura extraída da literatura (ALLEN.P, 2004)

A Figura 9 representa o esquema utilizado no método de condições periódicas de contorno. Esta representação mostra que à medida que uma partícula sai da caixa de simulação, sua imagem periódica se move da mesma maneira para substituí-la, mantendo o número de átomos na caixa central conservado (KUMAR, H.; MAITI, 2011). Quando as partículas estiverem dentro de um intervalo, é possível computar as interações entre os átomos reais e de imagem periódica (ALLEN.P, 2004).

#### 4.4.2.7

##### Algoritmos de integração em dinâmica molecular

As simulações de DM têm como principal finalidade gerar várias trajetórias de um sistema com o decorrer do tempo, possibilitando calcular as

propriedades termodinâmicas médias ou determinar a evolução temporal do sistema (KUMAR, H.; MAITI, 2011). Logo, é necessário o uso de operadores numéricos que sejam capazes de integrar as equações de movimento. Geralmente, algoritmos são utilizados para realizar essa tarefa (KUMAR, H.; MAITI, 2011). Os algoritmos de integração mais utilizados em simulações de DM são os algoritmos de Verlet e de Leap-Frog que serão detalhados nas próximas seções.

#### 4.4.2.7.1

##### Algoritmo de Verlet

O algoritmo de integração de Verlet é um operador matemático baseado no método das diferenças finitas, onde a integração é dividida em intervalos curtos de tempos. Estes intervalos são denominados de passos de integração ( $\Delta t$ ), que permitem simular os movimentos de maior frequência do sistema (VERLET, 1967). A funcionalidade deste algoritmo é baseada na equação 27:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2 \quad \text{Eq. 27}$$

onde  $a$  é aceleração,  $r$  é posição da partícula e  $t$  é o tempo. Esta equação determina as novas posições dos átomos,  $r(t + \Delta t)$ , no tempo de simulações ( $t + \Delta t$ ), usando acelerações e posições das partículas no tempo  $t$  e posições no passo anterior  $r(t - \Delta t)$  (VERLET, 1967). Esta equação pode ser obtida pela expansão de Taylor, conforme mostra as equações 28 e 29. A expansão de Taylor para a posição em torno do tempo  $t$  é representada na equação 28:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} a(t)\Delta t^2 + b(t)\Delta t^3 + O\Delta t^4 \quad \text{Eq. 28}$$

Se invertemos o tempo teremos

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{1}{2} a(t)\Delta t^2 - b(t)\Delta t^3 + O\Delta t^4 \quad \text{Eq. 29}$$

Combinando as duas equações e mantendo os termos até  $\Delta t^2$  temos,

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2 + O\Delta t^4 \quad \text{Eq. 30}$$

É importante mencionar que o algoritmo de Verlet é capaz de determinar a posição com precisão até  $\Delta t^4$  e a velocidade até  $\Delta t^2$  (VERLET, 1967). Para

o cálculo da velocidade, utilizam-se as diferenças de posições nos tempos  $(t - \Delta t)$  e  $(t + \Delta t)$ , dividido pelo tempo percorrido (Equação 31):

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} \quad \text{Eq. 31}$$

Portanto, temos as equações de posição e velocidade do algoritmo de Verlet (equações 30 e 31, respectivamente).

#### 4.4.2.7.2

##### Algoritmo de Leap-Frog

O algoritmo de Leap-Frog pode ser considerado uma otimização do algoritmo de Verlet, pois este foi projetado para corrigir a precisão numérica que era afetada por envolver erros grandes e pequenos. Neste algoritmo tanto a velocidade quanto a posição alternam simultaneamente. A primeira velocidade é calculada em meio passo de tempo e esta é utilizada para calcular a posição em passo de tempo completo (GUNSTEREN, VAN; BERENDSEN, 1988; KUMAR, H.; MAITI, 2011). As expressões para o algoritmo de Leap-Frog são apresentadas nas equações 32 e 33:

$$r(t + \Delta t) = r(t) + v\left(t + \frac{1}{2}\Delta t\right)\Delta t \quad \text{Eq. 32}$$

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + \frac{1}{m}F(t)\Delta t \quad \text{Eq. 33}$$

Observa-se que ao combinar as equações 32 e 33, a equação obtida é equivalente a equação 27 do algoritmo de Verlet. O algoritmo de Leap-Frog necessita da determinação da velocidade do passo de tempo anterior que pode ser determinada pela equação 34:

$$v\left(t_0 - \Delta t\right) = v(t_0) - \frac{1}{m}F(t_0)\frac{1}{2}\Delta t \quad \text{Eq. 34}$$

Para calcular a velocidade de uma partícula no tempo atual pode ser usada uma interpolação simples. Como dito anteriormente uma das principais vantagens do esquema de integração Leap-Frog é que não precisamos somar erros de ordens diferentes em  $\Delta t$  (KUMAR, H.; MAITI, 2011).

#### 4.4.2.8

##### Soma de Ewald (PME)

A forma das interações que é computada em uma simulação de DM associa-se diretamente ao custo computacional e a precisão dos resultados (ABRAHAM, Mark J.; GREASY, 2011). Em sistemas grandes, como por exemplo um sistema biomolecular, somar todas as contribuições de interações Coulombianas de pares de átomo, a fim de representar a natureza eletrostática de longo alcance é impraticável, devido ao elevado custo computacional (KUMAR, H.; MAITI, 2011).

A soma de Ewald é amplamente utilizada para tratar das interações em sistemas periódicos. A sua implementação dá-se através da equação 35:

$$E = \frac{f}{2} \sum_{n_x} \sum_{n_y} \sum_{n_z} \sum_i \sum_j \frac{q_i q_j}{r_{ij,n}}, (n_x + n_y n_z) = n \neq 0$$

**Eq. 35**

A equação 34 determina a energia eletrostática para N partículas e suas imagens periódicas. Os vetores de índice da caixa são representados por  $n_x, n_y, n_z$ , e a distância real entre as cargas é  $r_{ij,n}$  (ESSMANN et al., 1995). Para aumentar a velocidade de convergência a equação 34, esta é decomposta como o somatório de três termos independentes:  $E_{dir}$  calcula a energia de forma direta,  $E_{rec}$  fornece uma distribuição gaussiana sobre a energia e a transformada de Fourier e  $E_{corr}$  garante a neutralidade de carga da célula unitária (ESSMANN et al., 1995). Esses termos são apresentados nas equações 36, 37 e 38:

$$E_{dir} = \frac{f}{2} \sum_{i,j} \sum_{n_x} \sum_{n_y} \sum_{n_z} q_i q_j \frac{erfc(\beta r_{ij,n})}{r_{ij,n}}$$

**Eq. 36**

$$E_{rec} = \frac{f}{2\pi V} \sum_{i,j} q_i q_j \sum_{m_x} \sum_{m_y} \sum_{m_z} \frac{\exp(-(\frac{\pi m}{\beta})^2 + 2\pi i m (r_i - r_j))}{m^2}$$

**Eq. 37**

$$E_{corr} = -\frac{f\beta}{\sqrt{\pi}} \sum_i q^2$$

**Eq. 38**

onde o termo  $E_{\text{dir}}$  é referido como energia direta ou de curto alcance,  $E_{\text{rec}}$  é referido como energia recíproca ou de longo alcance e  $E_{\text{corr}}$  é referido como energia de correção. Nas equações 36, 37 e 38, o termo  $\beta$  é o parâmetro que define o peso relativo das somas direta e recíprocas,  $\text{erfc}$  é uma função de erro,  $V$  é o volume da célula unitária,  $m$  é o número de células unitárias e  $q$  é a carga das partículas.

#### 4.4.2.9

##### Algoritmo de Restrição

Em simulações de DM, geralmente as oscilações nas ligações são substituídas por restrições holonômicas, pois estas acontecem em frequência alta e amplitude baixa, o que limita o intervalo de tempo em uma simulação (HESS et al., 1997). Nesta situação, o comprimento de ligação entre os átomos fica longe do ideal, o que resulta na geração de energia que aquece o sistema. Esta situação geralmente ocorre com partículas pequenas e de pouca massa, visto que o momento médio do sistema produz grande movimento das partículas. Para resolver esse problema geralmente são aplicados algoritmos de restrições, os mais comuns são o LINCS e SHAKE.

##### 4.4.2.9.1

##### Algoritmo de Restrição LINCS

O algoritmo de restrição LINCS (solucionador de restrições lineares) foi projetado para otimizar o cálculo da equação matricial, executar correção de velocidade, evitando o alongamento rotacional e corrigindo o comprimento, e melhorando a precisão e a estabilidade (HESS et al., 1997). O método de LINCS é implementado pela equação 39, onde é considerado um sistema com  $N$  partículas, e posições dadas por um vetor  $3N$   $r(t)$ .

$$\frac{d^2r}{dt^2} = M^{-1} f \quad \text{Eq. 39}$$

onde  $f$  é a vetor de força  $3N$  e  $M$  uma matriz diagonal  $3N \times 3N$  contendo as massas das partículas. Geralmente, um sistema é limitado por  $K$  equações independentes do tempo na condição  $g_i(r) = 0$ ,  $i = 1, \dots, K$ . Adicionando as restrições como um potencial  $V(r)$ , e multiplicando pelos multiplicadores de Lagrange ( $\lambda_i$ ), temos a equação 39:



$$-M \frac{d^2 r}{dt^2} = \frac{d}{dr} (V - \lambda g) \quad \text{Eq. 40}$$

onde  $g$  é uma função de restrição  $g_i = (r_a - r_b)^2 - d^2 = 0$ , e  $r$  são as posições das partículas relacionadas a restrição, e  $d$  é o comprimento ideal da ligação.

#### 4.4.2.9.2

##### Algoritmo de restrição SHAKE

O método SHAKE utiliza a aplicação de restrição de distância, empregando coordenadas cartesianas para determinar os multiplicadores de Lagrange (KRAUTLER; GUNSTEREN, VAN; HUNENBERGER, 2001). Este algoritmo baseia-se na linearização e desacoplamento das equações de movimento, ou seja, as equações que redefinem as coordenadas atômicas dos sistemas (ELBER; RUYMGAART; HESS, 2011). A principal expressão matemática deste algoritmo é mostrada na equação 41:

$$\sigma_{\alpha \equiv (jk)}(X) = r_{jk}^2 - r_{jk,0}^2 = 0 \quad \forall \alpha = 1, \dots, L \quad \text{Eq. 41}$$

onde  $r_{jk}^2$  é o quadrado das distâncias entre os átomos  $j$  e  $k$  que são mantidas restritas,  $r_{jk,0}^2$  é o valor de referência ao qual o quadrado da distância é comparado e  $L$  é o número de distancias restritas. Por fim,  $\alpha$  é um índice que denota as restrições.

#### 4.4.2.10

##### Amostragem guarda-chuva

O método de amostragem guarda-chuva (do inglês umbrella sampling) é uma técnica que calcula a energia livre de um sistema ao longo de uma coordenada de reação ( $\xi$ ) (MILLS; ANDRICIOAEI, 2008). Geralmente  $\xi$  é definido em termos de geométricos, como distâncias, torsões e RMSD de dois estados de referências. Após a definição da  $\xi$  é possível calcular a energia livre ( $G$ ) de um processo pela equação 42:

$$G(\xi) = -\frac{1}{\beta} \ln Q(\xi) \quad \text{Eq. 42}$$

onde  $G(\xi)$  é também conhecido como potencial de força média (PMF). O termo  $\beta$  é igual ao inverso da constante de Boltzmann ( $k_b$ ) multiplicada pela temperatura absoluta  $T$ , e  $Q$  é a função de partição canônica. Para determinar

$G(\xi)$  em simulações computacionais, é necessário ter uma amostragem boa dos estados intermediários (KÄSTNER, 2011). Com a suposição de ter um sistema ergódico, a  $Q(\xi)$  é igual a função de distribuição  $P(\xi)$ , que é calculada pela equação 43:

$$P(\xi) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^1 \rho[\xi(t')] dt' \quad \text{Eq. 43}$$

onde  $t$  denota o tempo, e  $\rho$  é um termo que conta a ocorrência dos estados em um intervalo de  $\xi$ .

O método de amostragem guarda-chuva utiliza a aplicação de potenciais harmônicos (equação 44) ao longo de uma coordenada para que o sistema seja conduzido de um estado termodinâmico para outro (TORRIE; VALLEAU, 1977). A amostragem deste processo é geralmente realizada através de uma série de janelas para mostrar o intervalo de cada  $\xi$ .

$$\omega_i(\xi) = \frac{K}{2} (\xi - \xi_{ref})^2 \quad \text{Eq. 44}$$

onde  $w_i$  é o potencial de polarização da janela  $i$ ,  $K$  é a constante de força harmônica e  $\xi_{ref}$  é a coordenada de reação de referência. A aplicação do potencial em cada janela serve para o manter o sistema próximo ao ponto de referência  $\xi_{ref}$  da respectiva janela  $i$ . No método de amostragem guarda-chuva, a energia livre ( $G$ ) é determinada indiretamente através da função de distribuição não viesada ( $P_i^u$ ), conforme mostra as equações 45 e 46:

$$P_i^u = P_i^b(\xi) \exp[\beta \omega_i(\xi)] \langle \exp[-\beta \omega_i(\xi)] \rangle \quad \text{Eq. 45}$$

$$G(\xi) = -\frac{1}{\beta} \ln P_i^b(\xi) - \omega_i(\xi) + F_i \quad \text{Eq. 46}$$

onde  $P_i^b$  é a função de distribuição enviesada e  $F_i$  é uma ponderação que é calculado por  $F_i = -\frac{1}{\beta} \ln \langle e^{-\beta \omega(\xi)} \rangle$ . Geralmente, esse termo é obtido através de técnicas de análises de amostragem (KÄSTNER, 2011).

#### 4.4.2.11

##### Método da análise de histograma ponderado (WHAM)

Diversos métodos foram desenvolvidos para calcular o  $F_i$ , entre eles o método da análise de histograma ponderado (WHAM). Este é um método promissor que visa minimizar o erro estatístico associado a uma função de

distribuição global  $P^u(\xi)$  (KUMAR, S. et al., 1992). A determinação da distribuição  $P^u(\xi)$  é calculada através de uma média ponderada das distribuições  $P_i^u(\xi)$  das janelas individuais  $i$ , conforme a equação 47:

$$P^u(\xi) = \sum_i^{Janelas} p_i(\xi) P_i^u(\xi) \quad \text{Eq. 47}$$

onde  $p_i(\xi)$  é um peso escolhido para minimizar o erro estatístico da distribuição global (equação 48 e 49):

$$p_i = \frac{a_i}{\sum_j a_j} \quad \text{Eq. 48}$$

$$a_i(\xi) = N_i \exp[-\beta \omega_i(\xi) + \beta F_i] \quad \text{Eq. 49}$$

no qual  $N_i$  é o número de total de etapas amostradas para cada janela  $i$ . Logo, é possível notar a associação da ponderação  $F_i$  com a distribuição global  $P^u(\xi)$  (equação 47) e o peso  $p_i(\xi)$  pelas equações 48 e 49. No WHAM, a ponderação  $F_i$  é calculada pela equação 50:

$$\exp(-\beta F_i) = \langle \exp(-\beta \omega(\xi)) \rangle = \int P^u(\xi) \exp[-\beta \omega_i(\xi)] d\xi \quad \text{Eq. 50}$$

Ao observar a equação 47 onde tem-se  $P^u(\xi)$  e as equações 48 e 49 que mostram a associação de  $F_i$  com  $P^u(\xi)$ , a presença dos dois termos na equação 50 sugere que estes devem ser iterados até a convergência para obtenção do PMF (KÄSTNER, 2011; KUMAR, S. et al., 1992).

#### 4.4.2.12

##### Dinâmica molecular de aceleração randômica ( $\tau$ -RAMD)

O tempo de residência ( $\tau$ ) é definido como o tempo em que um fármaco permanece ligado ao sítio de ligação de um alvo, podendo ser calculado conforme a equação 51:

$$\tau = \frac{1}{k_{off}} \quad \text{Eq. 51}$$

onde  $k_{off}$  é constante de taxa de dissociação. Este parâmetro trata-se de uma constante cinética de primeira ordem que geralmente é expressa em  $s^{-1}$ . Para um fármaco ter efeito sobre um alvo, este tem que ter um tempo de residência

suficientemente longo para poder interferir na função da mesma (SOHRABY; NUNES-ALVES, 2022).

Uma das abordagens para determinar o tempo de residência relativo ( $\tau$ ) é chamada dinâmica molecular de aceleração aleatória  $\tau$  ( $\tau$ -RAMD). Esta técnica de cálculo de  $\tau$  é baseada em um procedimento de amostragem melhorada, onde uma pequena força adicional orientada aleatoriamente sobre a molécula ligante é aplicada, facilitando a sua desvinculação do alvo (KOKH et al., 2018). A direção da força aplicada é atribuída de forma aleatória toda vez que o ligante apresenta um movimento em um intervalo de tempo em que uma distância mínima definida não é respeitada.

Uma das vantagens de utilizar o método  $\tau$ -RAMD é o fato de que esta técnica não requer conhecimento prévio da via de dissociação, bem como um ajuste extensivo de parâmetros. Portanto, o único parâmetro a ser definido pelo usuário é a magnitude da força aleatória. Esta força geralmente não afeta o tempo de residência pois é definida para respeitar uma condição mínima de tempo de simulação para cada trajetória.

#### 4.4.2.13 GROMACS

GROMACS é um software para simulações de dinâmica molecular projetado para estudos da dinâmica de sistemas biomoleculares, tais como proteínas, ácidos nucleicos e lipídeos. A denominação de GROMACS vem da abreviação de **G**Roningen **M**Achine for **C**hemical **S**imulation, onde o primeiro termo é referente a cidade da Universidade onde este software foi desenvolvido (SPOEL, VAN DER et al., 2005). Este programa é o resultado de um projeto de colaboração entre o Departamento de Química e o Departamento de Ciência da Computação da Universidade de Groningen na Holanda no início dos anos de 1990. O Gromacs dispõe da possibilidade de utilização de vários campos de força, entre eles o OPLS, Gromos, Charmm e Amber que são bem comuns em estudos de DM (SPOEL, VAN DER et al., 2005). Além disso, há atualizações frequentes que acompanham os avanços computacionais em termos de hardwares e softwares, possibilitando simulações mais rápidas e com boa eficiência. Este software é de livre acesso com licença geral livre GNU-LGPL, sendo mantido por diversos centros de pesquisa. A sua robustez e rapidez em

comparação com os concorrentes, além do fato de ser gratuito, contribui para que este seja um dos programas mais utilizados nesta área.

## 5 Metodologia

### 5.1

#### Preparando as estruturas tridimensionais

Primeiramente a estrutura 3D do ligante e as estruturas 3D cristalográficas dos alvos foram obtidos no formato \*.PDB. Logo, após a aquisição destes arquivos, um processo de preparo pré-simulação foi realizado, para que os arquivos iniciais estivessem em conformidade com a necessidade de cada programa de simulação e para a otimização do custo computacional. O programa Autodock foi utilizado para o ancoramento molecular e posteriormente o programa Gromacs para a simulação de dinâmica molecular.

A estrutura 3D do peptídeo antiviral ATN-161 foi obtida do banco de dados PubChem (KIM et al., 2021) com o código CID: 9960285. Os alvos macromoleculares foram adquiridos do banco de dados de proteínas, onde as integrinas  $\alpha 5 \beta 1$  de conformação fechada e aberta foram obtidas pelos códigos PDB ID: 4WK4 (XIA; SPRINGER, 2014) e PDB ID: 7NWL (SCHUMACHER et al., 2021), respectivamente. Os alvos da proteinase principal 3C-Like ( $M^{PRO}$ ) e a proteína S da variante ômicron complexada com hACE2 do SARS-CoV-2 foram adquiridos pelos códigos PDB ID: 7BQY (JIN et al., 2020) e PDB ID: 7WPB (YIN et al., 2022), respectivamente.

Os alvos macromoleculares foram visualizados e sanitizados utilizando o pacote UCSF Chimera, onde foram removidas as estruturas que não são de interesse (PETTERSEN et al., 2004). Logo, na estrutura do 4WK4 foram retirados ligantes (sacarídeos) que não são de interesse. Em 7NWL, foram removidas as cadeias C da fibronectina, e fragmentos de anticorpos artificiais das cadeias D e E, ficando apenas com as cadeias referentes as subunidades  $\alpha 5$  e  $\beta 1$ . Para a estrutura de 7BQY, apenas a cadeia C referente ao peptídeo artificial N3 foi removida. Finalmente, apenas os ligantes sacarídeos foram removidos da estrutura 7WPB.

## 5.2

### Ancoramento Molecular

Para executar as simulações de ancoramento molecular foram utilizados os pacotes Autodock4.2 e o Autodock Tools (ADT) (MORRIS et al., 2014) como citado anteriormente. A simulação de ancoramento molecular foi executada com o intuito de obter as melhores poses iniciais do ligante ATN-161 complexado aos alvos macromoleculares. O sistema inicial foi preparado com o ADT, onde foram adicionados hidrogênios polares a estrutura de todas as proteínas e foram computadas as cargas de Kollmann (MORRIS et al., 1998b). Em relação ao ligante este foi adicionado ao sistema, e também foram computadas as cargas e seus pontos de torção foram detectados, no qual foram computadas 24 ligações rotacionáveis e definidos 17 torsões para simulação.

Na etapa seguinte, uma caixa de grade com dimensões de 126x126x126 Å<sup>3</sup> e espaçamento de grade de 0,375 Å foi gerada para cada sistema de alvo-ATN-161. Cabe ressaltar que as dimensões foram escolhidas para possibilitar que o ligante pudesse interagir com todos os sítios de ligação próximos ao sítio ativo definido na literatura para cada alvo. As coordenadas do centro de cada caixa de grade foram determinadas de acordo com os artigos associados a cada estrutura (JIN et al., 2020; SCHUMACHER et al., 2021; XIA; SPRINGER, 2014; YIN et al., 2022). Sendo assim as coordenadas para 4WK4 (x= 22.286; y= 11.004; z= -19.274) e 7NWL (x= 287.143; y= 288.333; z=240.064) foram referenciadas com a coordenada do sítio de ligação RGD presente em cada arquivo. A referência da estrutura 7BQY (x= 5.218; y= -2.997; z= 20.37) baseou-se na posição do resíduo de aminoácido GLN 166 que se encontra entre os domínios de ligação I e II. Finalmente, a referência da estrutura 7WPB (x= 196.913; y= 175.471; z= 248.689) foi colocada no domínio de ligação RGD.

Os sistemas foram submetidos a simulação de ancoragem molecular após a preparação de cada sistema. Vale ressaltar que a metodologia utilizada neste estudo considera as proteínas como estruturas rígidas e o ligante flexível. A simulação de ancoragem molecular consiste basicamente em duas etapas: a busca conformacional e o ranqueamento de energia para cada conformação encontrada. A busca conformacional foi realizada utilizando o algoritmo

genético Lamarckiano (LGA), onde todos os sistemas foram submetidos a 100 execuções independentes cada, com 150 indivíduos em cada população, perfazendo 25 milhões de avaliações de energia e gerando 27 mil gerações.

A partir dos resultados obtidos, 10 configurações por ranqueamento de energia ( $E_{\text{bind}}$ ) foram selecionadas dentre as 100 geradas para cada complexo do ligante ATN-161 com os alvos. Para escolher a configuração inicial para as simulações de dinâmica molecular, outros critérios foram também usados tais como a região que acomodou o maior número de configurações do ATN-161 entre as 10 primeiras poses e a interação com resíduos de aminoácidos específicos do sítio ativo de cada alvo.

### 5.3

#### Dinâmica molecular clássica

As simulações de DM para cada complexo alvo-ligante foram executadas conforme o seguinte fluxo de trabalho: 1) geração das topologias do ligante e dos alvos, 2) criação da caixa de simulação, 3) solvatação e adição de íons para neutralizar a carga do sistema, 4) minimização de energia do sistema, 5) equilíbrio do sistema, 6) simulação de produção das trajetórias e 7) análises das trajetórias.

As simulações de DM foram realizadas com o pacote GROMACS versão 2021.5 (SPOEL, VAN DER et al., 2005) usando o campo de força atômico OPLS-AA (Jorgensen et al., 1996). A plataforma PolyParGen (YABE et al., 2019) foi utilizada para gerar a topologia do ligante ATN-161. A função `pdb2gmx` do Gromacs foi usada para gerar as topologias dos alvos proteicos. Para a construção do sistema inicial, uma caixa dodecaédrica foi criada com função `gmx_editconf`. As dimensões da caixa de simulação foram definidas posicionando o complexo a uma distância de 1 nm das bordas da caixa.

Os volumes dos sistemas 4WK4, 7NWL, 7WPB e 7BQY possuem aproximadamente 2077 nm<sup>3</sup>, 3436 nm<sup>3</sup>, 1742 nm<sup>3</sup> e 661 nm<sup>3</sup>, respectivamente. Após a geração das caixas de simulação, moléculas de água foram adicionadas para solvatar o sistema usando a função `gmx_solvate`, onde o modelo de água TIP4P foi definido (GONZÁLEZ, Miguel A.; ABASCAL, 2011). Para neutralizar as cargas do sistema, a função `gmx_genion` foi utilizada para adicionar contra-íons ao sistema. Na etapa seguinte, uma minimização de energia foi realizada



aplicando o método de descida mais íngreme com um critério de convergência de  $10 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ .

A equilibração do sistema foi realizada em duas etapas. Primeiramente, 100 ps de simulação com o ensemble canônica (NVT) foi executada, onde o número de partículas, volume e temperatura são mantidos constantes. Em seguida, uma simulação de 100 ps com o ensemble isotérmico-isobárico NPT foi executada, onde foram mantidos constantes o número de partículas, a pressão e a temperatura. Desta forma, os sistemas foram otimizados em termos de energia e equilibrados na temperatura de aproximadamente 310K e pressão de 1 bar.

A partir do sistema equilibrado, executou-se uma simulação de DM de produção de trajetórias de 200 ns com temperatura constante de 310 K e pressão de 1 bar, usando o termostato reescalamento de velocidade com constante de acoplamento  $\tau$  de 0,1 ps (BUSSI; DONADIO; PARRINELLO, 2007) e o barostato de Parrinello-Rahman com uma constante de acoplamento  $\tau$  de 2,0 ps (PARRINELLO; RAHMAN, 1981). Todas as simulações usaram um intervalo integração de tempo (dt) de 2 fs. Além disso, o algoritmo LINCS foi aplicado como algoritmo de restrição e as interações eletrostáticas de longo alcance foram calculadas usando o método Particle-Mesh-Ewald (PME) (PETERSEN, 1995).

As análises das trajetórias foram feitas utilizando as funções `gmx_rms` para gerar o gráfico de raiz quadrada do desvio quadrático médio (RMSD) e a raiz quadrada da flutuação quadrática média (RMSF). O cálculo do RMSD determina o quanto a posição dos átomos da estrutura estudada se desvia em relação a uma estrutura de referência ao longo do tempo de simulação. Sendo assim, os RMSDs para a cadeia principal do ligante e para espinha dorsal das proteínas foram calculados em relação às suas estruturas iniciais de referência. De forma semelhante ao RMSD, o cálculo do RMSF determina o grau de flutuação das posições de aminoácidos específicos. Então, o RMSF foi calculado para as proteínas que mostraram RMSD acima de  $3 \text{ \AA}$ , com a finalidade de visualizar quais regiões contribuíram para o aumento do RMSD.

Além disso, o software Visual Molecular Dynamics (VMD) (HUMPHREY; DALKE; SCHULTEN, 1996) foi utilizado para visualizar as trajetórias dos

sistemas e o pacote UCSF Chimera (PETTERSEN et al., 2004) foi usado para gerar os mapas de contato de interações importantes entre o ligante e os alvos proteicos. Finalmente, a função gmx\_hbond foi utilizada para determinar a quantidade de ligações de hidrogênio que são formadas ao longo do tempo de simulação para cada sistema.

## 5.4

### Amostragem guarda-chuva

A amostragem guarda-chuva (KÄSTNER, 2011; TORRIE; VALLEAU, 1977) foi realizada com o pacote Gromacs para calcular a energia livre de ligação entre o peptídeo ATN-161 e os alvos proteicos. Os sistemas simulados anteriormente foram usados para criar um sistema para simulação de dinâmica molecular direcionada e amostragem guarda-chuva.

O solvente e os íons foram removidos para preparar o novo sistema. Uma nova caixa foi gerada com a geometria cúbica, considerando a condição em que a distância 5 nm entre a proteína e as bordas da caixa. No passo seguinte, as dimensões das caixas do sistema foram dobradas na direção de arraste do ligante na dinâmica molecular direcionada. Com isso, tem-se as seguintes dimensões das caixas de simulação para cada complexo ATN-161-alvo: 18.8 x 20.1 x 34.2 nm<sup>3</sup> para 4WK4, 23.4 x 18.9 x 44.7 nm<sup>3</sup> para 7NWL, 21.3 x 42.7 x 21.3 nm<sup>3</sup> para 7WPB e 30.8 x 11.9 x 12.5 nm<sup>3</sup> para 7BQY.

Com os as caixas geradas nas dimensões adequadas para simulação de DM direcionada, os sistemas foram novamente solvatados usando o modelo de água TIP4P (GONZÁLEZ, Miguel A.; ABASCAL, 2011) e a carga do sistema foi neutralizada pela adição de íons. Depois, foi necessário indicar a orientação para a força aplicada na simulação de dinâmica molecular direcionada. Logo, foi definido uma direção de arraste do ligante na dinâmica molecular direcionada para cada sistema considerando a melhor posição para retirada do ligante do sítio de ligação. As direções para cada sistema 4WK4 e 7NWL foram definidas no eixo Z, para o sistema 7WPB no eixo Y e para o sistema 7BQY no eixo X. Sendo assim o centro de massa do ATN-161 foi puxado ao longo da coordenada de reação ( $\xi$ ) para fora do sítio de ligação, utilizando um potencial

harmônico externo de  $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  e velocidade de arraste de  $0,0075 \text{ nm ps}^{-1}$  na dinâmica molecular direcionada.

As configurações iniciais para a simulação de amostragem guarda-chuva foram obtidas por meio das trajetórias da simulação de DM direcionada. Os quadros que foram amostrados foram gerados com um intervalo (ou janela) de  $0,1 \text{ nm}$  de distância entre si, ou seja, a configuração de um quadro foi adquirida cada vez que o ligante se moveu  $0,1 \text{ nm}$ . Os sistemas obtidos foram minimizados pelo método de descida mais íngreme com o critério de convergência de  $10 \text{ kJ mol}^{-1} \text{ nm}^{-1}$  e equilibrados por simulação NVT de  $100 \text{ ps}$  na temperatura de  $310\text{K}$  e pressão de  $1 \text{ bar}$ .

Finalmente, cada janela gerada foi equilibrada por uma simulação NVT adicional de  $500 \text{ ps}$ . Em seguida, uma simulação de produção de trajetórias de  $1 \text{ ns}$  foi realizada para cada janela usando uma constante de força de  $2000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ . Os resultados obtidos na amostragem foram analisados pelo método de análise de histograma ponderado (WHAM) (KUMAR, S. et al., 1992) onde o PMF foi calculado para cada sistema (LEMKUL; BEVAN, 2010). Desta forma, a energia livre de ligação de Gibbs ( $\Delta G_{\text{bind}}$ ) foi calculada para cada complexo usando a diferença de energia livre entre os estados ligado (mínimo) e dissociado (máximo) no perfil de energia.

## 5.5

### Dinâmica molecular de aceleração randômica ( $\tau$ -RAMD)

A simulação de  $\tau$ -RAMD calcula o tempo de residência relativo de um ligante em um sítio de ligação de um alvo proteico. Esta etapa foi realizada apenas para o complexo formado pelo ATN-161 com  $M^{\text{PRO}}$  e um complexo cristalográfico de referência. O complexo de referência utilizado é formado pelo ligante GC376 complexado na  $M^{\text{PRO}}$ . Esta referência foi obtida no PDB ID 7CBT. O complexo do ATN-161 com a  $M^{\text{PRO}}$  foi usado a partir das coordenadas iniciais utilizadas no cálculo de energia livre de ligação.

Primeiramente, o ligante foi separado da proteína usando o programa Pymol. Os átomos de hidrogênio foram adicionados ao ligante e as moléculas de água cristalográficas distantes em até  $5 \text{ \AA}$  foram mantidas e as demais removidas. Posteriormente, o complexo foi novamente formado, adicionando o

ligante e as moléculas de água nas mesmas coordenadas cristalográficas originais.

A ferramenta AmberTools 22 foi utilizada para gerar os parâmetros de topologia para os complexos. O método de AM1 foi utilizado para determinar as cargas parciais do ligante. O campo de força para proteínas ff14SB (do termo em inglês force field 14 Stony Brook) (MAIER, J. A. et al., 2015) e o campo de força GAFF (do termo em inglês Generalized Amber Force Field) (WANG, J. et al., 2004) para moléculas ligantes foram utilizados, bem como o modelo de água TIP3P (PRICE; BROOKS, 2004).

A caixa de simulação foi construída com a proteína a uma distância de 1 nm das bordas da caixa e o sistema foi solvatado com o modelo de água TIP3P (PRICE; BROOKS, 2004). Finalmente, íons foram adicionados para neutralizar a carga líquida do sistema. Com o sistema preparado, uma simulação de minimização de com 500 passos foi realizada pelo método de descida mais íngreme, seguida por 1000 passos de minimização pelo método de gradiente conjugado.

O sistema foi gradualmente aquecido até a temperatura de 300K em 1 ns, utilizando restrições harmônicas nos átomos diferentes de hidrogênio com uma constante de força de  $50 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ . Em seguida, a etapa de equilíbrio do sistema foi realizada, onde as restrições foram retiradas gradualmente diminuindo a constante de força em 50, 10, 5, 2  $\text{kcal mol}^{-1} \text{ \AA}^{-2}$ . Cada etapa foi simulada por 1ns. Posteriormente, uma simulação de DM de equilíbrio de 10 ns foi realizada sem restrições. Durante a equilíbrio, o termostato de Langevin foi utilizado para manter a temperatura em 300 K (DAVIDCHACK; HANDEL; TRETYAKOV, 2009) e o barostato de Berendsen (BERENDSEN et al., 1984) foi utilizado para manter a pressão em 1 bar.

Um corte de 10 Å foi definido para interações Coulombianas e Lennard-Jones. As condições de contorno periódicas foram aplicadas com um tratamento para interações eletrostáticas de longo alcance usando o método Particle Mesh Ewald (PME) (PETERSEN, 1995). Além disso, um intervalo de tempo de 2 fs foi empregado com ligações a átomos de hidrogênio restritas usando o algoritmo SHAKE (ANDERSEN, 1983). Com o sistema equilibrado, as coordenadas iniciais geradas pelo AMBER foram empregadas para o cálculo

do tempo de residência do ligante no alvo proteico usando o  $\tau$ -RAMD com o GROMACS.

Para o cálculo do tempo de residência do ligante no alvo proteico usando o  $\tau$ -RAMD, o sistema foi inicialmente aquecido em ensemble NVT por 10 ns, usando o termostato de Berendsen ( $\tau = 1$  ps). Em seguida, uma equilibrção foi realizada por 10 ns utilizando o termostato Nose-Hoover (EVANS; HOLIAN, 1985) e barostato de Parrinello-Rahman (PARRINELLO; RAHMAN, 1981) com constante de acoplamento de 0,8 e 2 ps, respectivamente. Esta etapa consiste em gerar a replica inicial para a simulação de  $\tau$ -RAMD. Logo, esta etapa foi repetida para gerar 5 replicadas. Cabe dizer que foram geradas coordenadas atômicas com diferentes velocidades para serem utilizadas como ponto de partida com o intuito de gerar simulações independentes e diferentes.

A simulação de  $\tau$ -RAMD foi constituída por 5 réplicas com um conjunto de 15 trajetórias foram geradas para cada replica. Em cada sistema, foi aplicada uma força de  $7 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$  com a direção da força definida aleatoriamente, após cada ponto de verificação no qual a saída do ligante fica retida. Esta força é redirecionada. As simulações foram encerradas quando o centro de massa do ligante moveu-se a uma distância superior a  $30 \text{ \AA}$  da sua posição de referência. Por fim, o método *bootstrap* foi utilizado para estimar o tempo de residência relativo, o desvio padrão e a distribuição média de cada réplica.

## 6 Resultados e Discussão

### 6.1

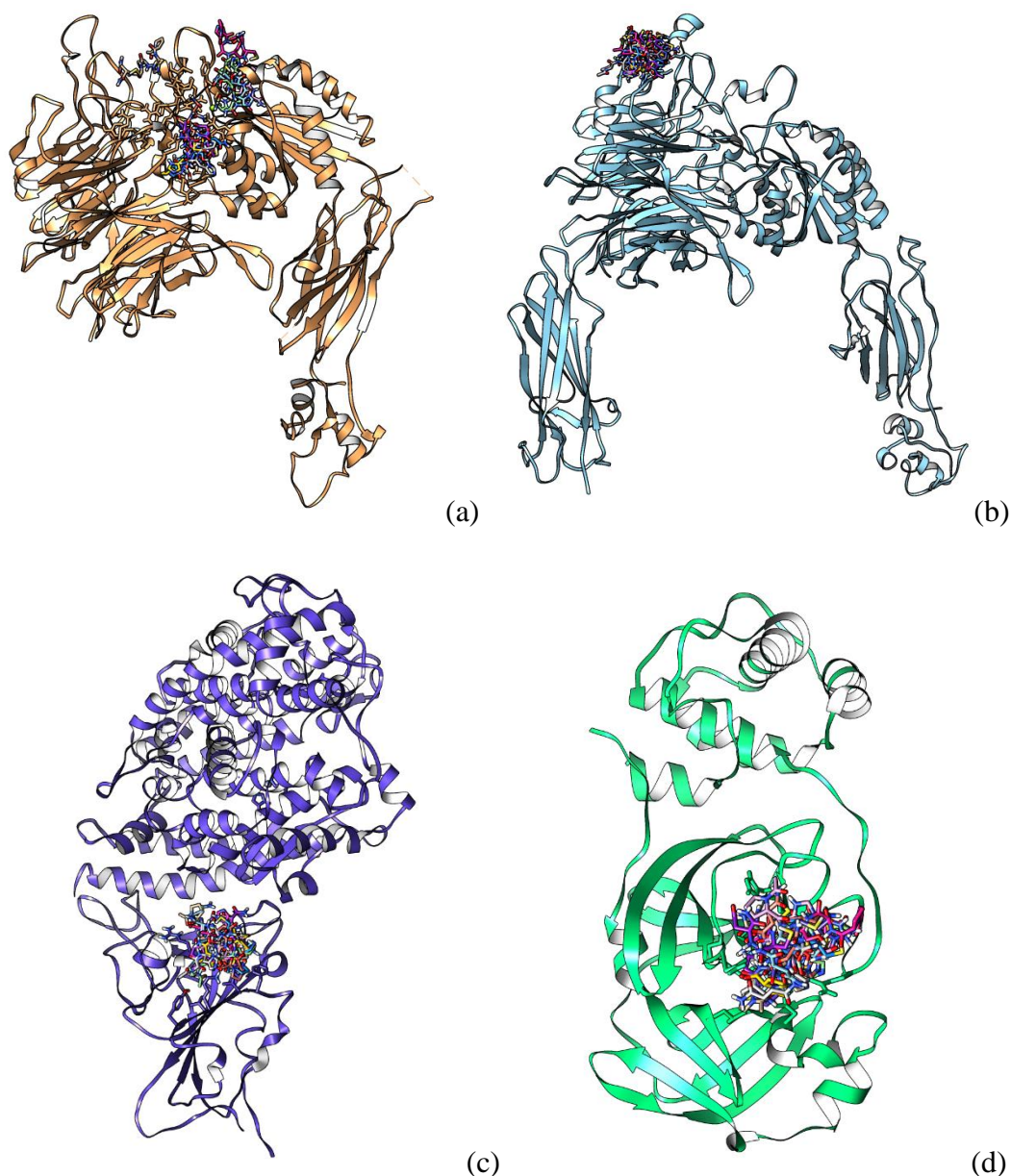
#### Ancoramento molecular

As simulações de ancoramento molecular tiveram como resultado a geração de 100 configurações alvo-ligante para cada macromolécula. Para realizar análises posteriores, estas configurações foram delimitadas as 10 primeiras de acordo com o ranqueamento de energia de ligação. Na Tabela 1 as 10 poses e suas respectivas energias ( $E_{\text{bind}}$ ) são apresentadas.

**Tabela 1.** Pontuação de energia de ligação ( $E_{\text{bind}}$ ) para a ancoragem do ATN-161 com as conformações fechada (código PDB: 4WK4) e aberta (código PDB: 7NWL) da integrina  $\alpha 5\beta 1$ , M<sup>PRO</sup> (código PDB: 7BQY) e proteína S da variante ômicron complexada ao hACE2 (código PDB: 7WPB) obtida por ancoragem molecular.

Pose	$E_{\text{bind}}$ (kcal mol <sup>-1</sup> )			
	4WK4	7NWL	7WPB	7BQY
1	-6.32	-6.33	-4.84	-7.27
2	-5.80	-5.71	-4.35	-7.20
3	-5.77	-5.36	-4.18	-7.03
4	-5.56	-5.21	-4.15	-6.99
5	-5.46	-5.18	-4.11	-6.79
6	-4.84	-5.12	-4.07	-6.75
7	-4.71	-5.01	-4.03	-6.60
8	-4.62	-4.84	-3.83	-6.22
9	-4.32	-4.76	-3.71	-6.19
10	-4.24	-4.60	-3.64	-6.18

A Tabela 1 mostra que a pontuação entre as 10 primeiras configurações para cada sistema é bastante próxima. Logo, para selecionar apenas uma configuração como estrutura de partida para as simulações de dinâmica molecular, fez-se necessário utilizar outros critérios. Os critérios adotados para selecionar as poses para integrina consideraram a energia de ligação ( $E_{\text{bind}}$ ), a região que acomodou o maior número de configurações entre as 10 primeiras já selecionadas, e em particular, os aminoácidos que interagem no complexo alvo-ligante. Os primeiros 10 complexos gerados pela simulação de ancoragem molecular para cada alvo são mostrados na Figura 10.

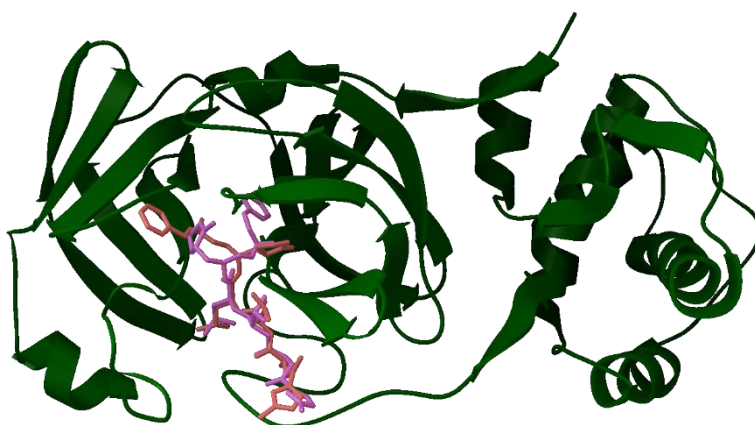


**Figura 10.** As primeiras 10 poses geradas pelo acoplamento molecular do ATN-161 na (a) integrina conformação fechada (código PDB: 4WK4) e na (b) conformação aberta (código PDB: 7NWL), na (c) proteína S da variante ômicron/hACE2 (código PDB: 7WPB) e na (d)  $M^{\text{PRO}}$  do SARS-CoV-2 (código PDB: 7BQY).

A Figura 10 mostra que há uma região de ocupação em que o ATN-161 se liga em cada sistema. Como em quase todos os sistemas, o maior número de poses foi em apenas uma região, como mostra a Figura 10b, c e d. A configuração de menor  $E_{\text{bind}}$  foi selecionada. Para a Figura 10a, também foram observados os aminoácidos que interagem, e ao comparar com a integrina de

conformação aberta, as poses de menor energia tinham interações semelhantes. Logo, foi selecionada aquela configuração de menor energia no ranqueamento. Portanto, as configurações selecionadas para o cálculo de dinâmica molecular foram aquelas de menor energia em cada sistema.

Como validação do ancoramento molecular, a simulação de reancoragem foi realizada. No entanto, as estruturas cristalográficas dos alvos 4WK4, 7NWL e 7WPB não possuem ligantes cristalográficos nos respectivos sítios de ligação estudados. Sendo assim, a reancoragem foi executada apenas para 7BQY, pois o ligante N3 está presente no sítio de ligação. O resultado para sobreposição é mostrado na Figura 11.



**Figura 11.** Reancoragem molecular do ligante N3 no sítio de ligação M<sup>PRO</sup> para os dados cristalográficos encontrados no código PDB: 7BQY. A pose cristalográfica de referência é mostrada em roxo e a pose do encaixe é mostrada em laranja.

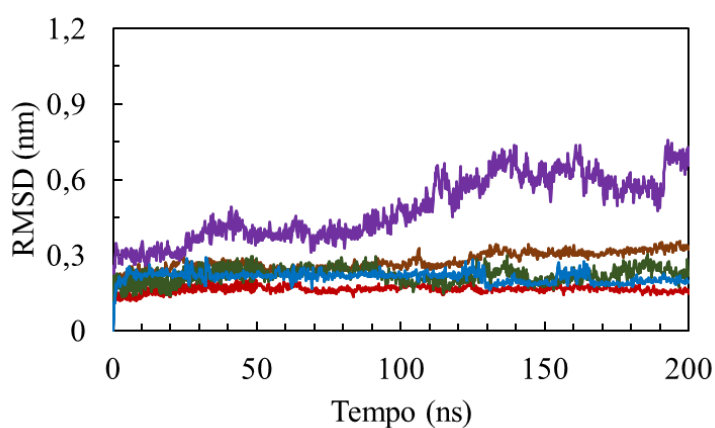
Os resultados deste reancoragem molecular mostram uma grande sobreposição das poses, onde foi calculado um RMSD de 1,388 Å, que é considerado dentro de uma margem de tolerância (geralmente RMSD < 3 Å). Além disso, foi também calculado um  $E_{\text{bind}}$  igual a -8,95 kcal mol<sup>-1</sup> para a melhor pose da ancoragem molecular (ligante em laranja). Para os demais sistemas, aminoácidos que interagem já foram reportados em outros estudos de modelagem molecular (LAN et al., 2020; SCHUMACHER et al., 2021; XIA; SPRINGER, 2014; YIN et al., 2022).

## 6.2

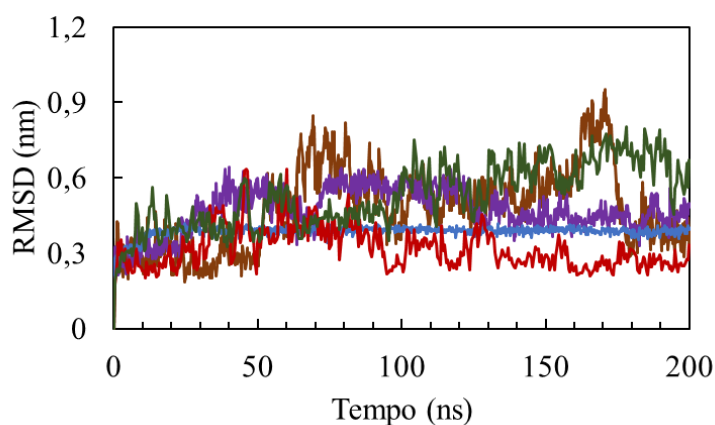


## Dinâmica molecular

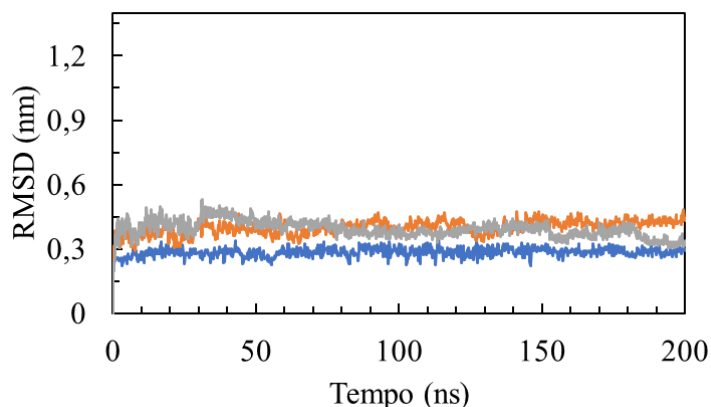
As estruturas obtidas das simulações de ancoramento molecular foram submetidas à solvatação, neutralização da carga líquida por adição de íons, minimização de energia e equilíbrio do sistema. Por fim, foi realizada uma simulação de produção de trajetória de 200 ns onde a estabilidade do sistema foi analisada, avaliando a estabilidade de seus componentes ao longo do tempo de simulação utilizando o RMSD das posições atômicas. Os resultados da análise de RMSD são mostrados na Figura 12, onde a variação das posições atômicas da cadeia principal das proteínas e do ligante em cada complexo, e das proteínas não complexadas foram avaliados.



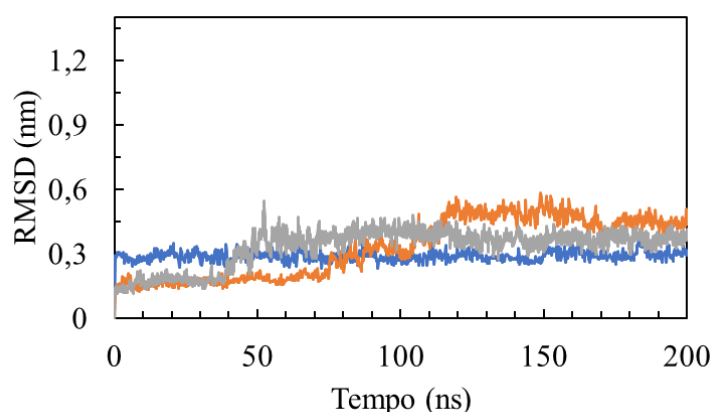
(a)



(b)



(c)



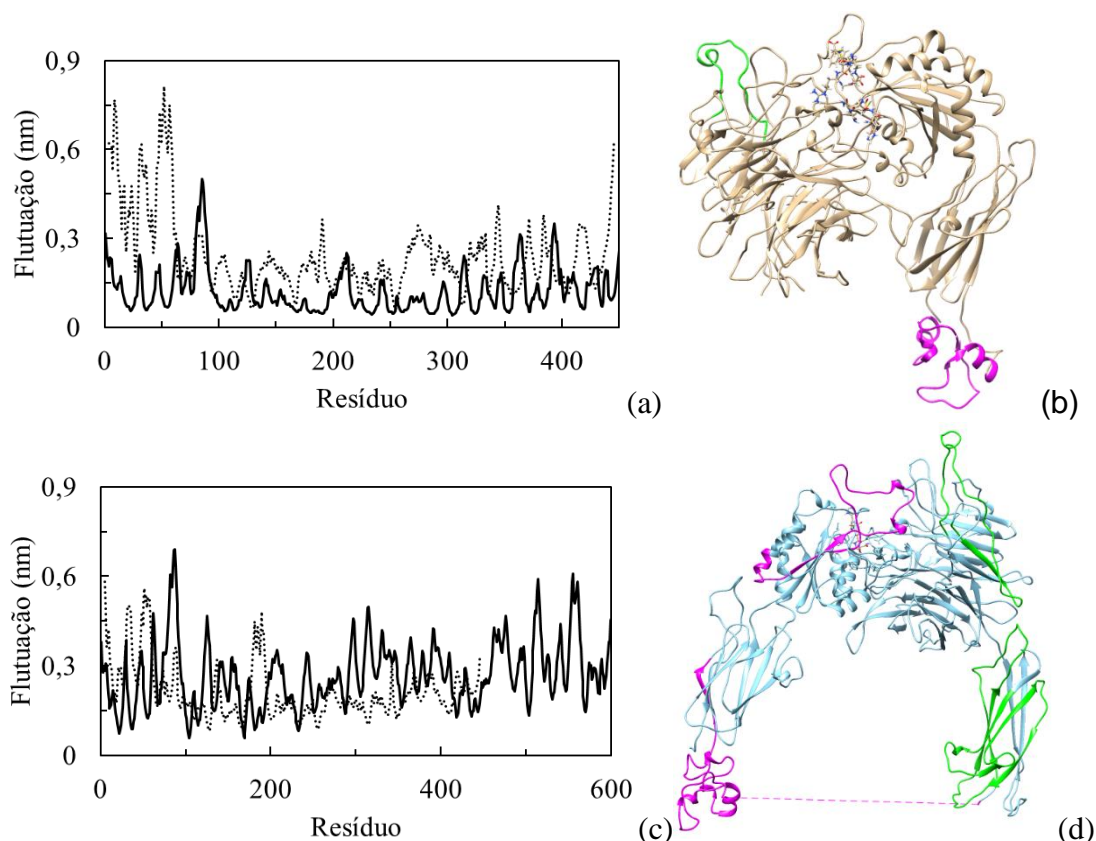
(d)

**Figura 12.** Raiz quadrada do desvio quadrático médio (RMSD) da posição atômica da cadeia principal da proteína complexada com ligante, cadeia principal da proteína sem ligante e átomos pesados de ATN-161 (azul). (a) conformação fechada e (b) conformação aberta da integrina  $\alpha 5\beta 1$ , no qual a cadeia A da proteína complexada está em marrom e cadeia B em roxo, a cadeia A da proteína sem ligante está em vermelho e a cadeia B em verde. (c) proteína S da variante ômicron/hACE2 e (d) complexo da  $M^{PRO}$  do SARS-CoV-2, proteína complexada em laranja e proteína sem ligante em cinza.

A Figura 12 mostra como o RMSD variou ao longo do tempo de simulação de 200 ns. Em geral, nota-se que o ligante ATN-161 não apresenta grande mobilidade no sítio de ligação, tendo assim apenas pequenas variações de RMSD. Análises visuais da trajetória de simulação pelo software VMD, indicam que o ligante se acomoda em uma determinada posição no sítio de ligação nos primeiros picosegundos da simulação. Então, após 5 ns de simulação não há variações acima de 0,1 nm para o ATN-161. Já para as

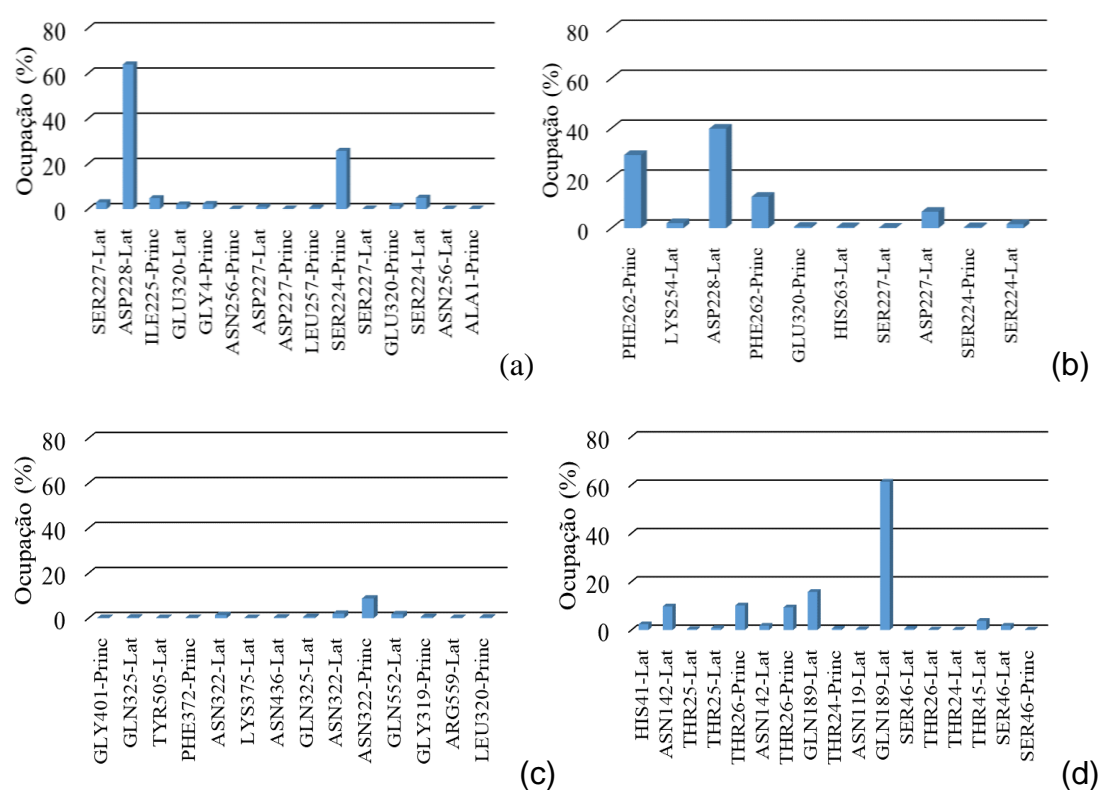
macromoléculas é notoriamente visível que há uma variação maior em relação ao ATN-161, o que é considerado normal, levando em conta que o número de átomos é muito maior nas proteínas em comparação com o ligante.

Analisando individualmente cada sistema, sugere-se fortemente que há estabilidade dos ligantes nos sítios de ligação de cada alvo proteico. Apesar disso, as maiores variações de RMSD nas proteínas (aproximadamente 0,3 nm) são mostradas nas Figuras 3a e 3b. Portanto, para esses casos em particular, a RMSF foi calculada e os resultados são mostrados na Figura 13.



**Figura 13.** Raiz quadrada da flutuação quadrática média (RMSF) da posição atômica por resíduo de aminoácido. As regiões que mostraram variação de RMSF maior que 0,3 nm foram destacadas. (a) RMSF da integrina de conformação fechada, (b) representação das regiões de maior variação na integrina conformação fechada. (c) RMSF da integrina de conformação aberta, (b) representação das regiões de maior variação na integrina conformação aberta. Em (a) e (c) as linhas sólidas referem-se à cadeia A e as linhas pontilhadas referem-se à cadeia B, em (b) e (d) as marcações de cor verde representam as maiores variações na cadeia A e em magenta na cadeia B.

A Figura 13 mostra que as regiões as quais causam o aumento significativo no RMSD, tendem a não alterar a estabilidade do sítio de ligação em que o ATN-161 liga-se. Logo, é demonstrado que os aminoácidos que participam do sítio de ligação e que interagem com o ATN-161 não apresentam variações maiores que 0,3 nm, indicando que o sistema está em equilíbrio. Além da variação das posições atômicas, a formação das ligações de hidrogênio é verificada durante a simulação em termos de ocupação relativa conforme apresentado na Figura 14. A distribuição de ocupação de ligações de hidrogênio indica o quanto o ligante permaneceu interagindo ao longo da simulação com o respectivo resíduo, assim sendo um dos indicadores de quão forte o ligante está interagindo sítio de ligação.

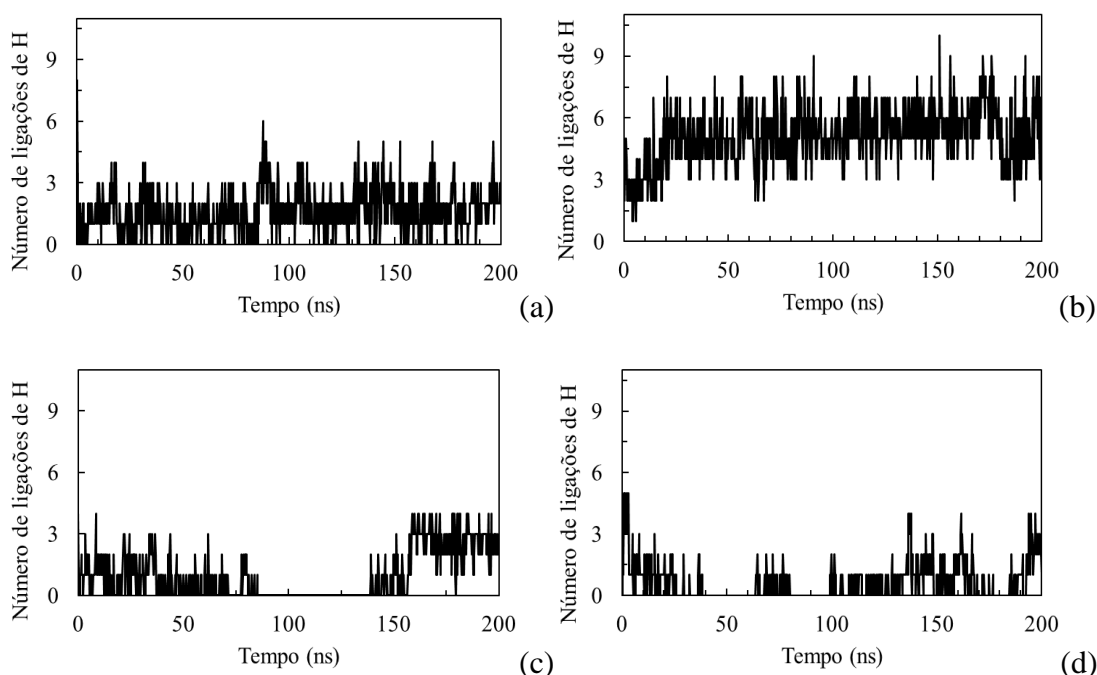


**Figura 14.** Distribuição de ligações de hidrogênio em termos de ocupância relativa entre o ATN-161 e os alvos. (a) integrina  $\alpha_5\beta_1$  de conformação fechada e (b) aberta, (c) proteína S da variante ômicron/hACE2. (d) M<sup>PRO</sup> do SARS-CoV-2. “Princ” refere-se a cadeia principal e “Lat” a cadeia lateral.

A análise para os sistemas com a integrina  $\alpha_5\beta_1$  mostrado nas Figuras 14a e 14b indicam que o ATN-161 interage fortemente com a cadeia lateral do resíduo ASP228, computando uma ocupação em torno de 60% no complexo

com conformação fechada e quase 40% em sua conformação aberta. Observa-se que os resíduos SER224 para a conformação fechada e PHE262 para a conformação aberta apresentam ocupação significativa (em torno de 25% e 30%). O sistema com a proteína S da variante da ômicron/hACE2 presente na Figura 14c foi aquele que apresentou a menor ocupação, não computando nenhuma ocupação acima de 10%, o que pode indicar que não há afinidade suficiente entre o ligante e os resíduos do sítio de ligação. O complexo com a M<sup>PRO</sup> mostrado na Figura 14d e a conformação fechada de integrina Figura 4a são aqueles que apresentaram a maior ocupação de ligação de hidrogênio entre todos os sistemas. O sistema com M<sup>PRO</sup> apresentou a maior ocupação de ligações de hidrogênio, sendo o valor em torno de 60% de ocupação entre o ATN-161 e GLN189.

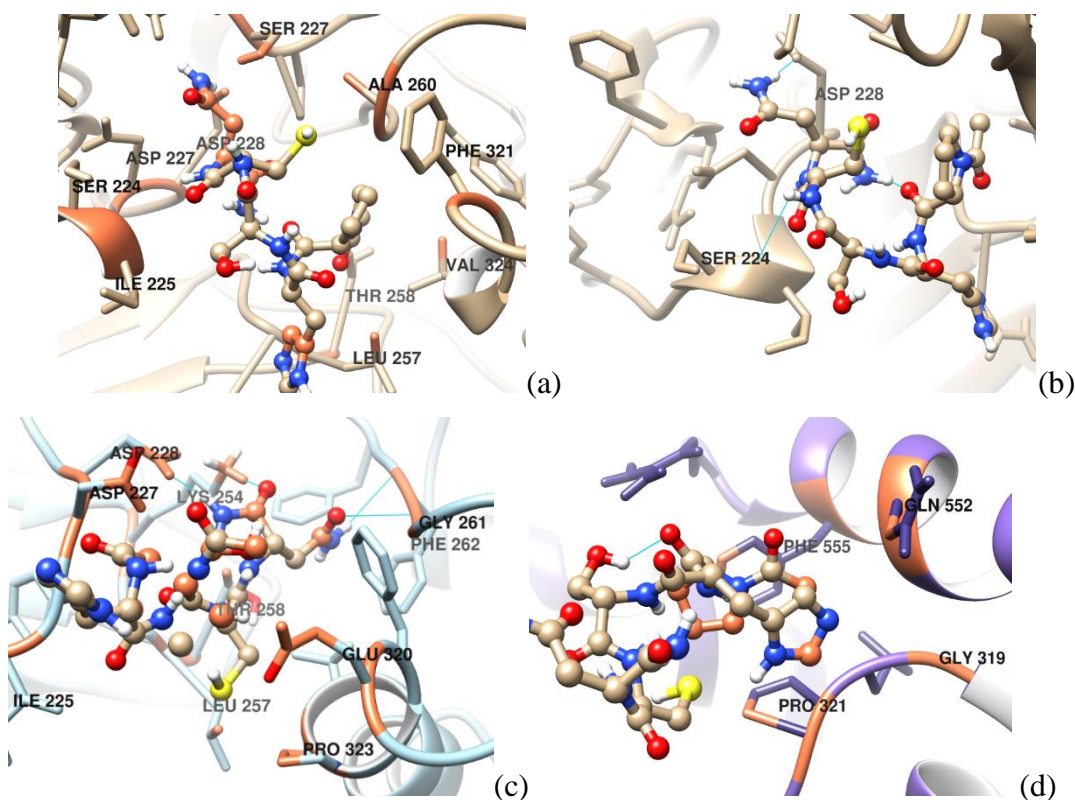
Para ter uma visão geral de como foram geradas as interações por ligações de hidrogênio, o número de ligações de hidrogênio formadas ao longo do tempo de simulação (200 ns) foram computadas para cada sistema estudado e estão presentes na Figura 15.



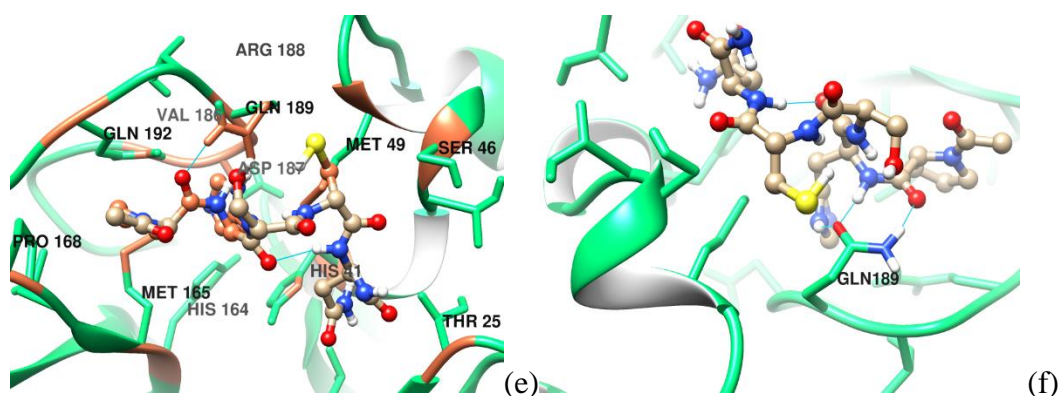
**Figura 15.** Número de ligações de hidrogênio dos sistemas de ATN-161 e alvos (a) integrina  $\alpha_5\beta_1$  com conformação fechada e (b) aberta, (c) proteína S da variante ômicron/hACE2. (d) M<sup>PRO</sup> do SARS-CoV-2.

A Figura 15 mostra que o ATN-161 forma várias ligações de hidrogênio no decorrer da simulação. Logo, isto indica que o ligante está estável no sítio de ligação. De forma particular, a Figura 15b referente ao ATN-161 com a conformação aberta da integrina é aquela que apresenta um maior valor médio de ligações de hidrogênio formadas. Comparando com a conformação fechada (Figura 15a), a quantidade média de ligações de hidrogênio computadas é quase duas vezes menor, o que pode resultar em uma maior afinidade por esta conformação. Os sistemas representados nas Figuras 15c e 15d apresentaram praticamente o mesmo valor médio de ligações de hidrogênio formadas, e é notável que em alguns períodos não são computadas as ligações de hidrogênio. Cabe ainda ressaltar que as ligações de hidrogênio têm uma contribuição importante para afinidade de ligação, porém outros fatores podem contribuir para uma maior ou menor afinidade de ligação.

Após analisar as ligações de hidrogênio, uma configuração do sistema foi escolhida para o cálculo de energia livre de ligação. Esta configuração também foi utilizada para calcular as demais interações. Estas interações intermoleculares entre o ATN-161 e as proteínas são apresentadas na Figura 16.







**Figura 16.** Interações intermoleculares entre ATN-161 e alvos. (a) interações da integrina  $\alpha_5\beta_1$  com conformação fechada, (b) ligações de hidrogênio da integrina  $\alpha_5\beta_1$  com conformação fechada, (c) interações da integrina  $\alpha_5\beta_1$  com conformação aberta, (d) interações da proteína S da variante ômicron/hACE2, (d) interações da M<sup>PRO</sup> do SARS-CoV-2 e (e) ligações de hidrogênio da M<sup>PRO</sup> do SARS-CoV-2.

A Figura 16a apresenta os resíduos que interagem com o ligante no sistema com a integrina com conformação fechada. Neste complexo, as interações com resíduos VAL324, PHE321, ALA260, THR258, LEU257, ASP228, ASP227, SER227, ILE225 e SER224 foram computadas. A Figura 16b dá um destaque as interações por ligações de hidrogênio, onde os resíduos ASP228 e SER224 foram mostrados. Pode-se notar que este resultado está acordo com os dados apresentados na Figura 14a. O sistema da conformação aberta da integrina é apresentado na Figura 16c. Para este complexo, as interações com os resíduos GLU320, PRO323, PHE262, GLY261, THR258, LEU257, LYS254, ASP228, ASP227 e ILE225 foram contabilizadas. Cabe ainda comentar que, dentre as interações registradas, 4 foram por ligações de hidrogênio, sendo interações com os resíduos PHE262, LYS254 e ASP228. Em particular, o resíduo PHE262 possui duas interações por ligação de hidrogênio.

Uma busca foi realizada com a finalidade de encontrar trabalhos que relatam o sítio de ligação do ATN-161 na integrina  $\alpha_5\beta_1$ . No entanto, não foram encontrados estudos com esse dado, sendo apenas sugerido que este peptídeo se liga ao sítio de sinergia, devido este ser inspirado pelo fragmento de fibronectina que interage neste sítio. Para analisar este fato primeiramente, é importante notar que o ATN-161 tem sequência Ac-PHSCN-NH<sub>2</sub>, e a região

de sinergismo da fibronectina tem sequência PHSRN, no qual é possível notar semelhanças e diferenças entre as estruturas.

A sequência PHSRN interage exclusivamente com o sítio na região N-terminal da subunidade  $\alpha 5$  integrina (KHALILI et al., 2006; SCHAFFNER; RAY; DONTENWILL, 2013). Porém, a substituição do resíduo ARG (R) no fragmento de fibronectina por CYS (C) no ATN-161 e inserção dos terminais acetil e amina fornecem novas possibilidades de interação como é mostrado nas Figuras 16a, 16b e 16c. Os resultados obtidos para os sistemas com a integrina mostram que o ATN-161 se liga entre o sítio de sinergismo e o bolso de ligação (RGD) para ambas as conformações estudadas.

A análise visual do sistema mostra as interações de ATN-161 com GLU320 e ASP227 no complexo  $\alpha 5 \beta 1$ /ATN-161 de conformação aberta. Essas interações são consideradas importantes, pois esses resíduos fazem parte do domínio de ligação RGD. Além disso, outros resíduos já citados anteriormente ficam próximos dos sítios de sinergismo e RGD em ambos os sistemas com integrinas (SCHUMACHER et al., 2021; XIA; SPRINGER, 2014). Cabe destacar que o resíduo ASP228 interage por ligação de hidrogênio e os resíduos TRH258, LEU257, ASP227 e ILE225 interagem por interações hidrofóbicas em ambos os complexos (conformações aberta e fechada).

O sistema formado pela proteína S da variante do ômicron/hACE2 é apresentado na Figura 16d. A análise deste complexo contabilizou as interações com os resíduos PHE555, GLN552, PRO321 e GLY319, sendo o sistema que apresentou o menor número de interações. Além disso, não foram computadas ligações de hidrogênio neste complexo. O resultado para o sistema com a M<sup>PRO</sup> é mostrado nas Figuras 16e e 16f, no qual foram computadas as interações com os resíduos GLN192, GLN189, ARG188, ASP187, VAL186, PRO168, MET165, HIS164, MET49, SER46, HIS41 e THR26. A Figura 16f destaca as ligações de hidrogênio formadas neste complexo, onde é possível notar que foram formadas 3 ligações, sendo todas com o resíduo GLN189.

Dentre as interações mostradas para o sistema com M<sup>PRO</sup>, cabe destacar a alta afinidade do ATN-161 pelo resíduo GLN189, que é um aminoácido importante para o domínio de ligação dessa proteína (ANTONOPOULOU et al., 2022; JIN et al., 2020). Além dessa interação, existem as interações com os



resíduos GLN192, ARG188, ASP187, VAL186, MET165, MET49 e HIS41 que fazem parte do sítio de ligação S2. A interação com o resíduo HIS41 é mais um indicativo da ação inibitória do ATN-161, considerando que HIS41 é um componente da díade catalítica dessa enzima (ANAND et al., 2003; NGUYEN et al., 2020; OWEN et al., 2021).

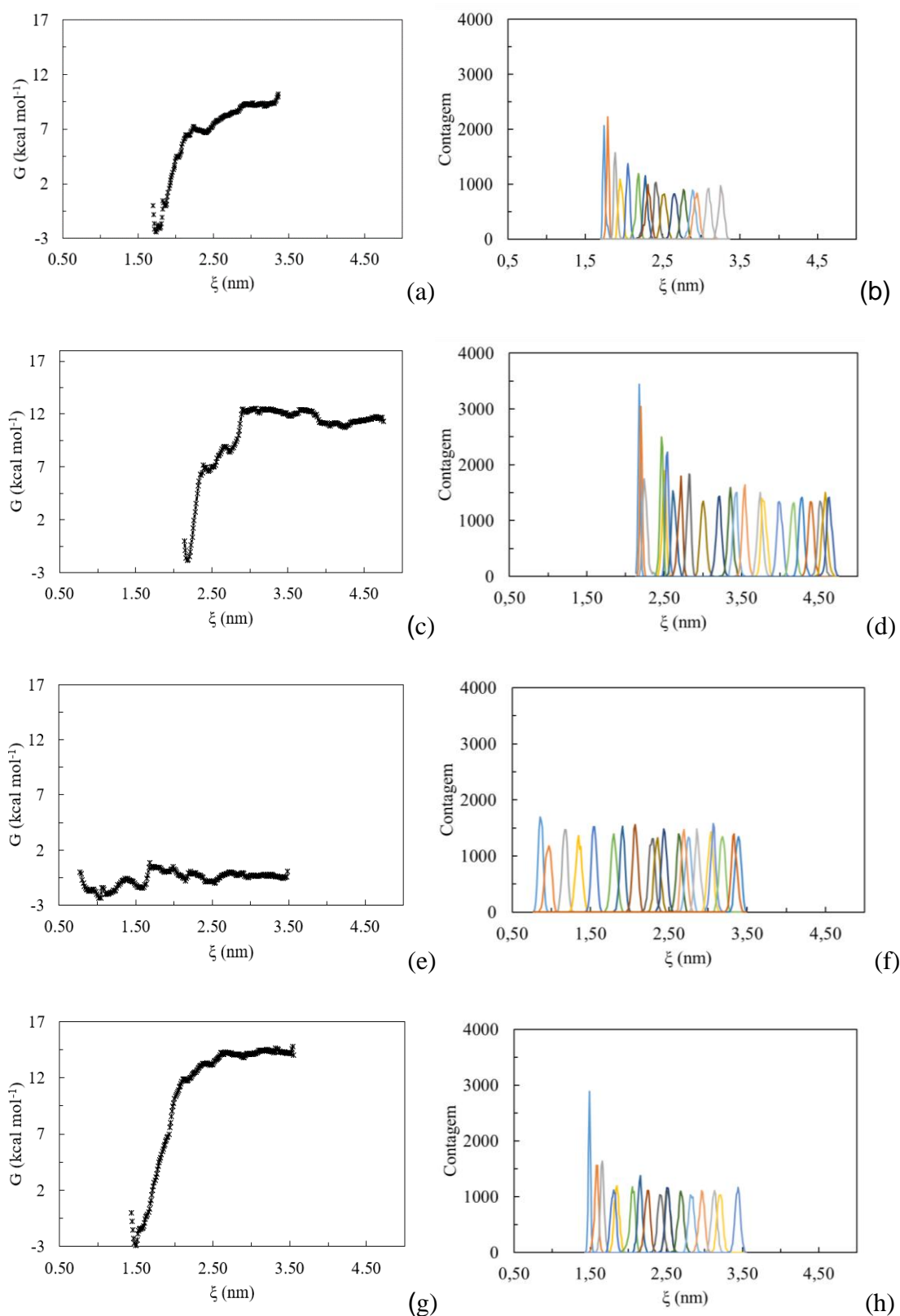
Sobre a interação do ATN-161 com M<sup>PRO</sup>, a estabilidade do ligante é mediada pela interação com aminoácidos importantes para o funcionamento dessa enzima, indicando que o peptídeo estudado tem grande potencial para inibir esse alvo e consequentemente inibir a replicação viral decorrente do SARS-CoV-2.

De forma geral, as Figuras 14, 15 e 16 apresentam dados que concordam entre si. Comparando os dados da Figura 15, nota-se que o valor médio do número de ligações de hidrogênio está próximo do apresentado para cada sistema na Figura 16. Os resíduos que interagem por ligação de hidrogênio são aqueles que apresentaram maiores ocupações na Figura 14, com exceção da LYS254 na Figura 14c.

### 6.3

#### Método de Amostragem Guarda-Chuva

O método de amostragem guarda-chuva foi utilizado para calcular a energia livre de ligação ( $\Delta G_{\text{bind}}$ ) para cada sistema. A partir da  $\Delta G_{\text{bind}}$  é possível aferir o quão forte o ATN-161 liga-se a cada alvo, bem como calcular a constante de inibição aproximada. O cálculo da  $\Delta G_{\text{bind}}$  é feito através da diferença de energia (G) entre os estados ligado (mínimo) e dissociado (máximo). O potencial de força média ao longo da coordenada de reação é utilizado para calcular o  $\Delta G_{\text{bind}}$  como mostrados na Figuras 17a, 17c, 17e e 17g. Os histogramas para cada sistema são mostrados na Figuras 17b, 17d, 17f e 17h.



**Figura 17.** Perfis de energia livre de Gibbs e histogramas obtidos pelo método de amostragem guarda-chuva e análise de histograma ponderado (WHAM) para ATN-161/Alvos. (a) potencial de força média e (b) histograma para integrina  $\alpha_5\beta_1$  conformação fechada, (c) potencial de força média e (d)

histograma para integrina  $\alpha_5\beta_1$  conformação aberta, (e) Potencial de força média e (f) histograma para a proteína S da variante de ômicron/hACE2, (g) potencial de força média e (h) histograma para a M<sup>PRO</sup> do SARS-CoV-2.

A Figura 17 apresenta os PMF's e os respectivos histogramas para cada sistema como citado anteriormente. A energia livre é determinada a partir da análise do potencial de força média. Os histogramas servem para validar estatisticamente os valores obtidos na simulação. É possível notar que todos os histogramas apresentados na Figura 17 estão dentro de uma margem tolerável de sobreposição, conferindo confiabilidade as energias livres calculada. Um resumo das  $\Delta G_{\text{bind}}$  calculadas é mostrado na Tabela 2.

**Tabela 2.** Energia livre de ligação ( $\Delta G_{\text{bind}}$ ) do ATN-161 com as conformações fechada (código PDB: 4WK4) e aberta (código PDB: 7NWL) da integrina  $\alpha_5\beta_1$ , da proteína M<sup>PRO</sup> (código PDB: 7BQY) e da proteína S da variante do ômicron/hACE2 (código PDB: 7WPB), calculada por amostragem guarda-chuva.

<b>Proteína Alvo</b>	<b><math>\Delta G_{\text{bind}}</math> (kcal mol<sup>-1</sup>)</b>
<b>4WK4</b>	-12.6±0.2
<b>7NWL</b>	-14.4±0.5
<b>7WPB</b>	-3.2±0.3
<b>7BQY</b>	-17.7±0.2

Os resultados para  $\Delta G_{\text{bind}}$ , de todos os ligantes mostram que os complexos são formados espontaneamente, pois apresentam energia livre de ligação negativas. Porém, é possível notar que o ATN-161 apresenta maior afinidade para o sistema 7BQY. Os resultados para a integrina em conformação aberta (7NWL) e fechada (4WK4) mostram afinidades de ligação termodinamicamente semelhantes, ou seja, não há uma diferença significativa. Entretanto, o sistema 7WPB apresenta uma baixa afinidade de ligação.

Os sistemas 4WK4 e 7NWL apresentam  $\Delta G_{\text{bind}}$  de  $-12,6 \pm 0,2$  e  $-14,4 \pm 0,5$  kcal mol<sup>-1</sup> respectivamente, com a conformação fechada tendo uma energia ligeiramente inferior, conforme mostrado na Tabela 2. De forma geral, a energia

livre de ligação indica que o ATN-161 se liga mais facilmente à conformação aberta da integrina  $\alpha_1\beta_5$ , possivelmente devido ao impedimento estérico na conformação fechada (NAGAE et al., 2012; TAKAGI et al., 2003). No entanto, é muito provável que o ATN-161 se ligue a essa proteína mesmo na forma inativa, como já foi demonstrado para a fibronectina (SCHUMACHER et al., 2021).

Os resultados para os alvos do SARS-CoV-2 indicam que o ATN-161 se liga fortemente à proteína M<sup>PRO</sup> (7BQY), com uma  $\Delta G_{\text{bind}}$  de aproximadamente  $-17,7 \pm 0,2 \text{ kcal mol}^{-1}$  (Tabela 2), sendo o sistema que apresentou a maior afinidade de ligação. No entanto, o ATN-161 se liga fracamente a proteína S da variante ômicron/hACE2 (7WPB), com um  $\Delta G_{\text{bind}}$  de  $-3,2 \pm 0,3 \text{ kcal mol}^{-1}$  (Tabela 2). A explicação para esse resultado pode estar relacionada à baixa ocupação nas ligações de hidrogênio, bem como nas poucas interações intermoleculares em geral (Figura 4).

A energia de livre de ligação pode ser comparada a dados experimentais relatados na literatura se a energia livre de ligação experimental ( $\Delta G_{\text{bind,exp}}$ ) for calculada a partir da concentração inibitória máxima ( $IC_{50}$ , molar) ou constante de inibição ( $K_i$ ) pela seguinte fórmula:

$$\Delta G_{\text{bind.exp}} \approx RT \ln IC_{50} \approx RT \ln K_i \quad \text{Eq. 52}$$

onde R é a constante geral dos gases e T é a temperatura. Portanto, para efeito de comparação, dados experimentais foram encontrados para cada sistema, quando possível, para alguns ligantes com atividade comprovada em cada alvo.

O  $IC_{50}$  experimental para a inibição da integrina  $\alpha_5\beta_1$  com ATN-161 é de 15 nM segundo dados da literatura (YAO; VEINE; LIVANT, 2016). Logo, a sua  $\Delta G_{\text{bind,exp}}$  é de  $-10,7 \text{ kcal mol}^{-1}$ , onde pode-se considerar termodinamicamente de acordo com os valores encontrados pelas simulações de dinâmica molecular,  $-12,6 \pm 0,2$  e  $-14,4 \pm 0,5 \text{ kcal mol}^{-1}$  para a integrina  $\alpha_5\beta_1$  nas conformações fechada (4WK4) e aberta (7NWL), respectivamente. A ligação do ATN-161 com a integrina  $\alpha_1\beta_5$  na sua forma inativa é um resultado importante. Logo, sugere-se que esse inibidor tem atividade contra a ativação de dentro para fora que ocorre por interação de elementos do citoesqueleto (DHALIWAL; SHEPHERD, 2022; MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021). A ativação de dentro pra fora ocasiona a vulnerabilidade do receptor à ligação por proteínas virais e proteínas de adesão. Portanto, a sua inibição

também irá inibir a sinalização de fora para dentro por ligantes extracelulares como a fibronectina que ativam esse receptor (DHALIWAL; SHEPHERD, 2022; MAKOWSKI; OLSON-SIDFORD; WEISEL, 2021).

A busca realizada para o sistema 7WPB mostrou que até o momento, não há estudos experimentais de ATN-161 com a proteína S da variante ômicron. Porém estudos *in vitro* e *in vivo* com o vírus nativo comprovaram que esse peptídeo atua diminuindo a carga viral e a replicação do vírus dentro das células (AMRUTA et al., 2021; BEDDINGFIELD et al., 2021). O experimento de inibição de Beddingfield et al (2021) obteve um IC<sub>50</sub> de 3,16 µM, o que resulta em um  $\Delta G_{\text{bind,exp}}$  de -7,50 kcal mol<sup>-1</sup>. No entanto, o resultado da simulação do ATN-161 com a proteína S da variante do ômicron apresenta quase que a metade do  $\Delta G_{\text{bind,exp}}$  para a proteína nativa do vírus, o que pode ser explicado pela mutação presente no domínio de ligação.

Assim como para o sistema ATN-161/7WPB, o sistema com ATN-161/7BQY também não apresenta dados experimentais, o que é compreensível dado que o SARS-CoV-2 é uma doença recente. Portanto, foram pesquisadas moléculas aprovadas para o uso contra o SARS-CoV-2 na literatura (ANTONOPOULOU et al., 2022; FU, L. et al., 2020). Estas moléculas apresentam estudos clínicos e *in vivo* comprovando a atividade inibitória contra a M<sup>PRO</sup>. Os ligantes Lufotrelvir, Nirmatrelvir, S-217622 e GC376 foram selecionados como referência, que indicam IC<sub>50</sub> de 0,27, 77,9, 13 e 150 nM, respectivamente. Estes valores resultam em  $\Delta G_{\text{bind,exp}}$  de -13,03, -9,70, -10,76 e -9,31 kcal mol<sup>-1</sup> para Lufotrelvir, Nirmatrelvir, S-217622 e GC376, respectivamente (ANTONOPOULOU et al., 2022; FU, L. et al., 2020). Assim sugere-se que a energia de ligação no complexo ATN-161/M<sup>PRO</sup> encontrada pela simulação de dinâmica molecular apresenta afinidade de ligação comparável com esses inibidores da M<sup>PRO</sup>.

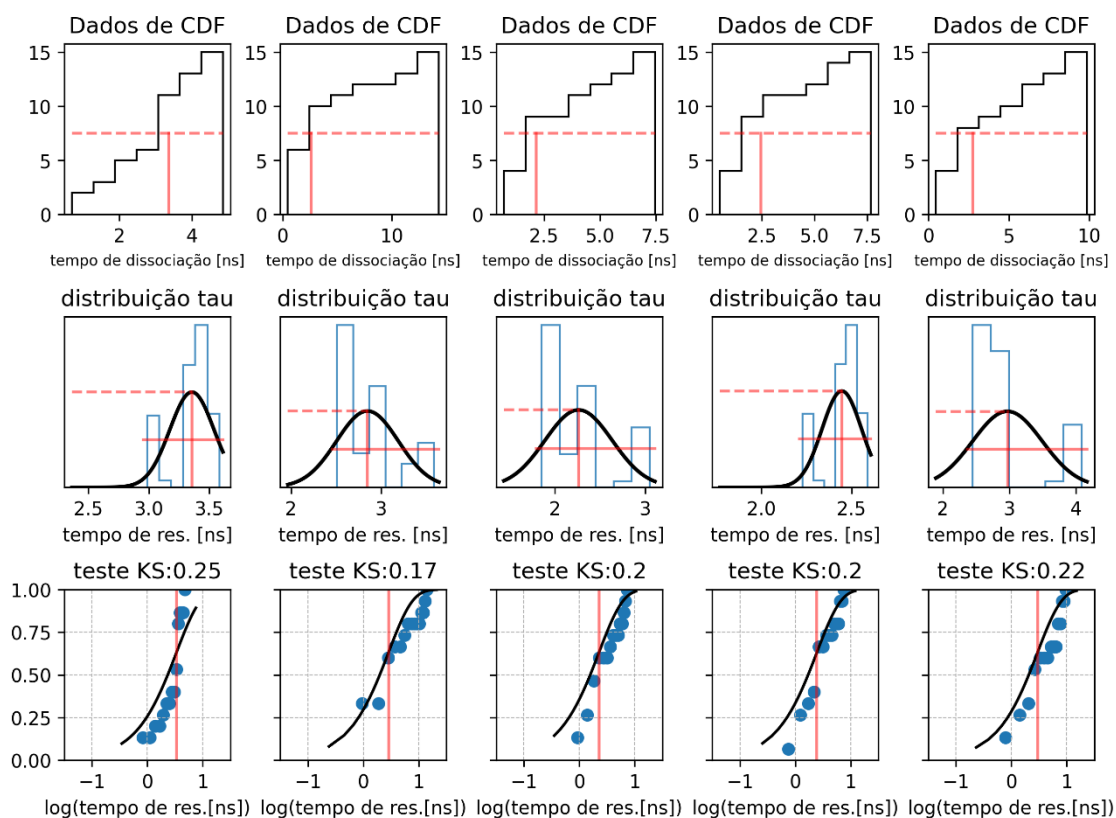
A forte interação do ATN-161 com uma das enzimas mais importantes para a replicação viral do SARS-CoV-2 é um resultado considerado promissor. Além disso, esse sistema apresenta menor  $\Delta G_{\text{bind}}$  do que os complexos  $\alpha 5\beta 1$ /ATN-161. Portanto, os resultados encontrados neste trabalho, indicam que o ATN-161 apresenta grande potencial como candidato a inibidor da M<sup>PRO</sup> do SARS-CoV-2. Também, a diminuição da carga viral apresentada em estudos

experimentais pode ser reflexo da inibição concomitante da entrada alternativa por integrinas, bem como pela inibição da protease principal M<sup>PRO</sup>.

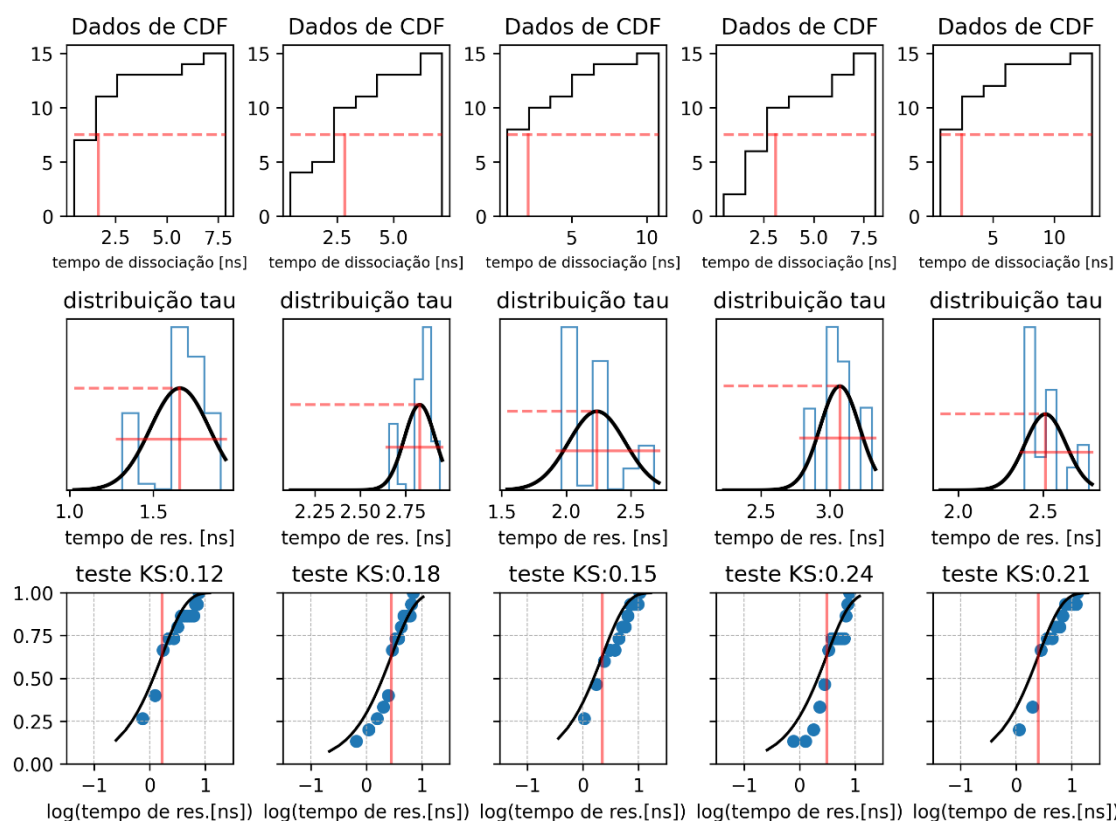
### 6.3

#### Dinâmica molecular de aceleração randômica ( $\tau$ -RAMD)

O tempo de residência relativo calculado por  $\tau$ -RAMD para cada sistema, é estimado através de cada replica, sendo que uma única replica gera 15 trajetórias de dissociação do ligante. A validação da metodologia é feita através do método *bootstrap* que gera gráficos de função de distribuição cumulativa (CDF), função de distribuição de tempo residência efetivo e de teste de Kolmogorov-Smirnov (KS). Os resultados estatísticos para as simulações dos complexos ATN-161/M<sup>PRO</sup> e GC376/M<sup>PRO</sup> são mostrados nas Figuras 18 e 19.



(a)

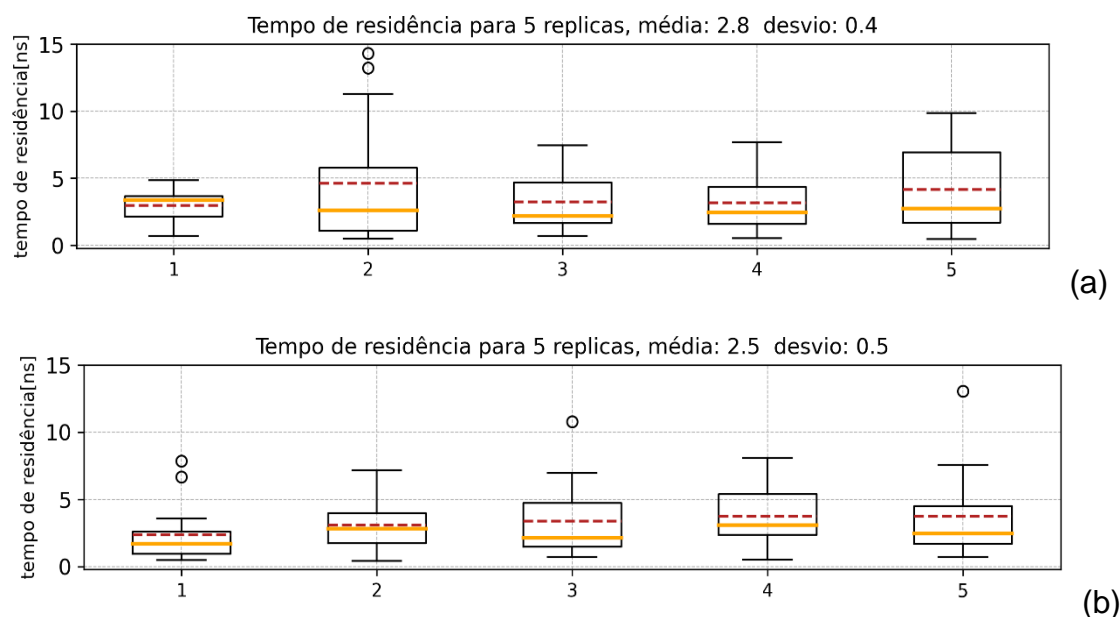


(b)

**Figura 18.** Resultados estatísticos para o complexo (a) ATN-161 com  $M^{\text{PRO}}$  e (b) GC376 com  $M^{\text{PRO}}$ : Função de distribuição cumulativa (CDF), função de distribuição de tempo residência efetivo e teste de Kolmogorov-Smirnov (KS).

A função de distribuição cumulativa (CDF) para os dois complexos mostra que o tempo de residência dos ligantes no sítio de ligação da  $M^{\text{PRO}}$  que foi observado em metade das trajetórias é entre 1,5 e 3,5 ns, sendo indicado pela linha sólida vermelha. A segunda linha das Figuras 18a e 18b mostra a distribuição gaussiana (linha sólida preta) em relação aos tempos de dissociação e as linhas vermelhas indicam o tempo médio para cada replica, indicando o mesmo tempo de dissociação já observado. A terceira linha das Figuras 18a e 18b mostra o teste de Kolmogorov-Smirnov (KS) que apresenta a função de distribuição cumulativa de Poisson (PCDF) na linha sólida preta. Esta função é comparada com a função de densidade cumulativa empírica (ECDF) representada pelos símbolos azuis obtidos dos dados brutos da simulação.

O teste de KS indica que há um comportamento semelhante entre as distribuições. Estes dados são determinados pela distância entre a distribuição de Poisson e os CDFs empíricos. Portanto, os valores de KS indicam que as quantidades de réplicas utilizadas são suficientes para que tenha uma boa amostragem do tempo de residência para tais sistemas. Sendo assim, os tempos de residência relativos são apropriados. A variação do tempo de residência relativo para cada replica é mostrado na Figura 19.



**Figura 19.** Variação do tempo de residência médio para cada replica. Complexo (a) ATN-161 – M<sup>PRO</sup> e (b) GC376 – M<sup>PRO</sup>. A caixa representa o intervalo dos valores médios entre as trajetórias e os valores discrepantes são representados por círculos. A barra mostra o desvio dos valores entre as trajetórias de cada replica. A média é mostrada pelas linhas vermelhas tracejadas e a mediana pela linha sólida laranja.

Os valores dos tempos de residência médios para cada réplica apresentam variações aceitáveis considerando que o método *bootstrap* realiza um tratamento de dados. O método *bootstrap* retornou valores de tempos de residência relativos de  $2,8 \pm 0,4$  e  $2,5 \pm 0,5$  ns para os sistemas de M<sup>PRO</sup> com ATN-161 e GC376, respectivamente. Apesar de não se tratar de um dado absoluto, ou seja, não é um dado comparável com dados experimentais, esta metodologia mostrou-se eficaz na classificação de tempos de residências para diversos fármacos (KOKH et al., 2018). Além disso, simulações com dados



absolutos precisam de uma extensiva amostragem, considerando que esses processos podem levar de segundos a horas em sistemas com moléculas ativas, o que requer um alto custo computacional.

Os resultados para ATN-161 são promissores quando comparado ao GC376, já que apresenta um tempo de residência similar a esta molécula que apresenta diversos estudos que comprovam a sua atividade inibitória contra a M<sup>PRO</sup>. Estudos recentes relatam que o tempo de residência e a afinidade de ligação apresentam boas correlações com a eficácia de uma molécula ativa (COPELAND, 2016; 2021; FOLMER, 2018). Portanto, estes dados são mais um indicativo que o ATN-161 tem grande potencial para ser um fármaco inibidor da replicação viral do SARS-CoV-2.

## 7 Implicações

As implicações deste trabalho estão relacionadas ao entendimento do mecanismo de ação do ATN-161 em alguns alvos proteicos importantes do SARS-CoV-2 e a possibilidade de aplicação do ATN-161 em um tratamento terapêutico contra o SARS-CoV-2. Como este peptídeo antiviral apresenta estudos de atividade contra outras patogenicidades, como acidente vascular cerebral isquêmico (EDWARDS et al., 2020), colite (DANESE et al., 2007), câncer de mama (KHALILI et al., 2006), metástase hepática (STOELTZING et al., 2003), encefalomielite hemaglutinante suína (LV et al., 2019) e degeneração macular relacionada à idade (WANG, W. Q. et al., 2016), entre outros, existe um amplo conhecimento da comunidade científica, o estudo apresentado nesta dissertação é mais amplo do que a área de estudo do SARS-CoV-2. Além disso, estudos comprovam que o ATN-161 apresenta baixa toxicidade, e por ser um peptídeo antimicrobiano tem baixa capacidade de desenvolver resistência, o que potencializa a importância e implicações desta dissertação em uma área de grande ameaça e importância para a humanidade.

## Conclusões

Neste trabalho, foi demonstrado com sucesso que o pentapeptídeo ATN-161 forma complexos estáveis espontaneamente com os alvos da integrina  $\alpha_5\beta_1$  (conformações aberta e fechada), da proteína S variante do ômicron com complexo hACE2 e da protease principal ( $M^{PRO}$ ). As configurações iniciais foram geradas por ancoramento molecular, e a estabilidade dos complexos foi verificada por dinâmica molecular. Os resultados obtidos por estudos de dinâmica molecular indicam que o ATN-161 se liga fortemente à integrina  $\alpha_5\beta_1$  (tanto nas conformações abertas quanto nas fechadas) com  $\Delta G_{bind}$  de  $-14.4 \pm 0.5$  e  $-12.6 \pm 0.2$  kcal mol<sup>-1</sup>, respectivamente. Também, o ATN-161 se liga a  $M^{PRO}$  com  $\Delta G_{bind}$  de  $-17.7 \pm 0.2$  kcal mol<sup>-1</sup>. Os resíduos que contribuem para ligação do ATN-161 com sítio de ligação de cada alvo também foram apresentados. Além disso, descobriu-se que o ATN-161 se liga fracamente à proteína S variante ômicron complexada a hACE2 com uma  $\Delta G_{bind}$  de  $-3.2 \pm 0.3$  kcal mol<sup>-1</sup>.

Portanto, verifica-se que o ATN-161 interagiu com a integrina  $\alpha_5\beta_1$  em uma região que fica entre o domínio RGD e o local de sinergismo. De acordo com a literatura, o ATN-161 atua ligando-se fora do bolso RGD. Os resultados demonstram que o ATN-161 pode se ligar tanto a forma ativa como inativa da integrina  $\alpha_5\beta_1$ , inibindo a ativação de fora para dentro e de dentro para fora, sugerindo o grande potencial de inibição que é observado experimentalmente. Os resultados para a interação de ATN-161 com a integrina  $\alpha_5\beta_1$  são importantes porque podem contribuir para o desenvolvimento de novos inibidores de integrinas não dependentes do sítio RGD, pois as integrinas são alvos de várias patogenicidades, incluindo SARS-CoV-2 que foi recentemente sugerido por ensaios in vitro e in vivo na literatura.

Os resultados para a ligação de ATN-161 com o complexo hACE2-complexo de proteína S variante ômicron mostram baixa afinidade de ligação, tendo poucas interações intermoleculares. Os resultados para ATN-161 com  $M^{PRO}$  destacam-se por apresentar a maior afinidade de ligação entre todos os sistemas. Entre os resultados apresentados para este sistema pode-se também salientar a interação com os aminoácidos que fazem parte do sítio de ligação

dessa enzima, como o resíduo GLN189 que computou 3 pontes de hidrogênio e o HIS41, que faz parte da díade catalítica. Os resultados de ATN-161 com M<sup>PRO</sup> levaram a realização de estudos da cinética de ligação deste complexo. A simulação da cinética de dissociação teve como resultado para ATN-161 um tempo de residência relativo maior ou no mínimo igual ao ligante de referência utilizado, no qual foi computado um tempo de  $2,8 \pm 0,4$  para o ATN-161 e  $2,5 \pm 0,5$  ns GC376.

Os resultados obtidos neste trabalho corroboram que o ATN-161 é um ótimo candidato a novo fármaco contra SARS-CoV-2, tendo em vista o surgimento de novas variantes e as opções limitadas de fármacos para o tratamento desta doença. Uma das principais vantagens do uso de peptídeos é o fato de que geralmente estes não desenvolvem resistência á agentes microbianos. Portanto, os resultados deste trabalho contribuem para a sociedade científica indicando uma molécula com grande potencial de inibição da replicação viral da SARS-CoV-2. Além disso, ele contempla dados que podem justificar resultados obtidos em *in vivo* e *in vitro* relatados na literatura, como a diminuição da infecção viral quando aplicado concentrações de ATN-161, que segundo os resultados obtidos pode ser devido a inibição concomitante da M<sup>PRO</sup> e da via alternativa de entrada por integrinas. Além de indicar o possível sitio de ligação do ATN-161 com a integrina  $\alpha_5\beta_1$ , que é um resultado que não foi encontrado na literatura. Porém, mesmo com esses excelentes resultados preliminares, mais estudos experimentais são necessários para uma melhor validação e descrição destes sistemas.

## 9

### Perspectivas Futuras

Para uma possível continuação deste trabalho indica-se a realização de estudos experimentais como forma de validação dos resultados adquiridos por simulações moleculares. Além de testar a afinidade de ligação do ATN-161 com demais alvos do SARS-CoV-2. Atualmente um estudo experimental do sistema ATN-161 com M<sup>PRO</sup> está sendo desenvolvido em parceria com o Prof. Gregory J. Bix da Tulane University em Nova Orleans nos EUA, o que motivou a realização de estudos da cinética de ligação para este mesmo complexo. É importante também estudar a cinética de ligação do ATN-161 nos sítios de ligação dos demais alvos proteicos do SARS-CoV-2.

## Referências Bibliográficas

ABRAHAM, Mark J.; GREASY, J. E. **Optimization of parameters for molecular dynamics simulation using smooth particle-mesh Ewald in GROMACS 4.5.** *Journal of Computational Chemistry*, [s.l.], v. 32, nº 9, p. 2031–2040, 2011. ISSN: 01928651, DOI: 10.1002/jcc.21773.

ABRAHAM, Mark James et al. **Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers.** *SoftwareX*, [s.l.], v. 1–2, p. 19–25, 2015. ISSN: 23527110, DOI: 10.1016/j.softx.2015.06.001.

ABRAMOV, Y. A.; SUN, G.; ZENG, Q. **Emerging Landscape of Computational Modeling in Pharmaceutical Development.** *Journal of Chemical Information and Modeling*, [s.l.], v. 62, nº 5, p. 1160–1171, 2022. ISSN: 1549-9596, DOI: 10.1021/acs.jcim.1c01580.

ADELUSI, T. I. et al. **Molecular modeling in drug discovery.** *Informatics in Medicine Unlocked*, [s.l.], v. 29, p. 100880, 2022. ISSN: 23529148, DOI: 10.1016/j.imu.2022.100880.

AGARWAL, G.; GABRANI, R. **Antiviral Peptides: Identification and Validation.** *International Journal of Peptide Research and Therapeutics*, [s.l.], v. 27, nº 1, p. 149–168, 2021. ISBN: 0123456789, ISSN: 15733904, DOI: 10.1007/s10989-020-10072-0.

ALLEN.P, M. **Introduction to Molecular Dynamics Simulation.** *Neumann Institute for Computing*, [s.l.], v. 23, p. 1–28, 2004.

AMRUTA, N. et al. **In Vivo protection from SARS-CoV-2 infection by ATN-161 in k18-hACE2 transgenic mice.** *Life Sciences*, [s.l.], v. 284, 2021. ISSN: 18790631, DOI: 10.1016/j.lfs.2021.119881.

ANAND, K. et al. **Coronavirus main proteinase (3CLpro) Structure: Basis for design of anti-SARS drugs.** *Science*, [s.l.], v. 300, nº 5626, p. 1763–1767, 2003. ISSN: 00368075, DOI: 10.1126/science.1085658.

ANDERSEN, H. C. **Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations.** *Journal of Computational Physics*, [s.l.], v. 52, nº 1, p. 24–34, 1983. ISSN: 00219991, DOI: 10.1016/0021-9991(83)90014-1.

ANTONOPOULOU, I. et al. **Inhibition of the main protease of SARS-CoV-2 (Mpro) by repurposing/designing drug-like substances and utilizing nature’s toolbox of bioactive compounds.** *Computational and Structural Biotechnology Journal*, [s.l.], v. 20, p. 1306–1344, 2022. ISSN: 20010370, DOI: 10.1016/j.csbj.2022.03.009.

AOTA, S.; NOMIZU, M.; YAMADA, K. M. **The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function.** *Journal of Biological Chemistry*, [s.l.], v. 269, nº 40, p. 24756–24761, 1994. ISSN: 00219258, DOI: 10.1016/S0021-9258(17)31456-4.

AOUDJIT, F.; VUORI, K. **Integrin Signaling in Cancer Cell Survival and**

**Chemoresistance.** *Chemotherapy Research and Practice*, [s.l.], v. 2012, p. 1–16, 2012. ISSN: 2090-2107, DOI: 10.1155/2012/283181.

AROOJ, M. et al. **Computational insights into binding mechanism of drugs as potential inhibitors against SARS-CoV-2 targets.** *Chemical Papers*, [s.l.], v. 76, nº 1, p. 111–121, 2022. ISBN: 0123456789, ISSN: 13369075, DOI: 10.1007/s11696-021-01843-0.

ASSA-MUNT, N. et al. **Solution Structures and Integrin Binding Activities of an RGD Peptide with Two Isomers.** *Biochemistry*, [s.l.], v. 40, nº 8, p. 2373–2378, 2001. ISSN: 0006-2960, DOI: 10.1021/bi002101f.

BAHAR, A.; REN, D. **Antimicrobial Peptides.** *Pharmaceuticals*, [s.l.], v. 6, nº 12, p. 1543–1575, 2013. ISSN: 1424-8247, DOI: 10.3390/ph6121543.

BAI, Q. et al. **Application advances of deep learning methods for de novo drug design and molecular dynamics simulation.** *WIREs Computational Molecular Science*, [s.l.], v. 12, nº 3, 2022. ISSN: 1759-0876, DOI: 10.1002/wcms.1581.

BARKER, J. A.; WATTS, R. O. **Structure of water; A Monte Carlo calculation.** *Chemical Physics Letters*, [s.l.], v. 3, nº 3, p. 144–145, 1969. ISSN: 00092614, DOI: 10.1016/0009-2614(69)80119-3.

BEDDINGFIELD, B. J. et al. **The Integrin Binding Peptide, ATN-161, as a Novel Therapy for SARS-CoV-2 Infection.** *JACC: Basic to Translational Science*, [s.l.], v. 6, nº 1, p. 1–8, 2021. ISSN: 2452302X, DOI: 10.1016/j.jacbts.2020.10.003.

BELOUZARD, S. et al. **Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein.** *Viruses*, [s.l.], v. 4, nº 6, p. 1011–1033, 2012. ISSN: 1999-4915, DOI: 10.3390/v4061011.

BENVENUTO, D. et al. **The 2019-new coronavirus epidemic: Evidence for virus evolution.** *Journal of Medical Virology*, [s.l.], v. 92, nº 4, p. 455–459, 2020. ISSN: 10969071, DOI: 10.1002/jmv.25688.

BERENDSEN, H. J. C. et al. **Molecular dynamics with coupling to an external bath.** *The Journal of Chemical Physics*, [s.l.], v. 81, nº 8, p. 3684–3690, 1984. ISSN: 0021-9606, DOI: 10.1063/1.448118.

BROOIJMANS, N.; KUNTZ, I. D. **Molecular Recognition and Docking Algorithms.** *Annual Review of Biophysics and Biomolecular Structure*, [s.l.], v. 32, nº 1, p. 335–373, 2003. ISSN: 1056-8700, DOI: 10.1146/annurev.biophys.32.110601.142532.

BUSSI, G.; DONADIO, D.; PARRINELLO, M. **Canonical sampling through velocity rescaling.** *Journal of Chemical Physics*, [s.l.], v. 126, nº 1, 2007. ISSN: 00219606, DOI: 10.1063/1.2408420.

CARVACHO, I.; PIESCHE, M. **RGD-binding integrins and TGF- $\beta$  in SARS-CoV-2 infections – novel targets to treat COVID-19 patients?** *Clinical & Translational Immunology*, [s.l.], v. 10, nº 3, 2021. ISSN: 2050-0068, DOI: 10.1002/cti2.1240.

CHEN, W. et al. **Detectable 2019-nCoV viral RNA in blood is a strong**

**indicator for the further clinical severity.** *Emerging Microbes & Infections*, [s.l.], v. 9, nº 1, p. 469–473, 2020. ISSN: 2222-1751, DOI: 10.1080/22221751.2020.1732837.

CHOWDHURY, S. M. et al. **Antiviral Peptides as Promising Therapeutics against SARS-CoV-2.** *Journal of Physical Chemistry B*, [s.l.], v. 124, nº 44, p. 9785–9792, 2020. ISSN: 15205207, DOI: 10.1021/acs.jpcc.0c05621.

CIANFROCCA, M. E. et al. **Phase 1 trial of the antiangiogenic peptide ATN-161 (Ac-PHSCN-NH 2), a beta integrin antagonist, in patients with solid tumours.** *British Journal of Cancer*, [s.l.], v. 94, nº 11, p. 1621–1626, 2006. ISSN: 00070920, DOI: 10.1038/sj.bjc.6603171.

CLEGG, D. O. **Integrins in the development function and dysfunction of the nervous system.** *Frontiers in Bioscience*, [s.l.], v. 8, nº 4, p. 1020, 2003. ISSN: 10939946, DOI: 10.2741/1020.

COPELAND, R. A. **The drug–target residence time model: a 10-year retrospective.** *Nature Reviews Drug Discovery*, [s.l.], v. 15, nº 2, p. 87–95, 2016. ISSN: 1474-1776, DOI: 10.1038/nrd.2015.18.

COPELAND, R. A. **Evolution of the drug-target residence time model.** *Expert Opinion on Drug Discovery*, [s.l.], v. 16, nº 12, p. 1441–1451, 2021. ISSN: 1746-0441, DOI: 10.1080/17460441.2021.1948997.

CORNELL, W. D. et al. **A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules.** *Journal of the American Chemical Society*, [s.l.], v. 117, nº 19, p. 5179–5197, 1995. ISSN: 0002-7863, DOI: 10.1021/ja00124a002.

COSTA, F. et al. **Covalent immobilization of antimicrobial peptides (AMPs) onto biomaterial surfaces.** *Acta Biomaterialia*, [s.l.], v. 7, nº 4, p. 1431–1440, 2011. ISSN: 17427061, DOI: 10.1016/j.actbio.2010.11.005.

COSTA, R. K. M. et al. **The reconciliation between the experimental and calculated octanol-water partition coefficient of 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine using atomistic molecular dynamics: an open question.** *Journal of Biomolecular Structure and Dynamics*, [s.l.], v. 1, 2023.

DANESE, S. et al. **Angiogenesis blockade as a new therapeutic approach to experimental colitis.** *Gut*, [s.l.], v. 56, nº 6, p. 855–862, 2007. ISSN: 0017-5749, DOI: 10.1136/gut.2006.114314.

DANHIER, F.; BRETON, A. LE; PRÉAT, V. **RGD-Based Strategies To Target Alpha(v) Beta(3) Integrin in Cancer Therapy and Diagnosis.** *Molecular Pharmaceutics*, [s.l.], v. 9, nº 11, p. 2961–2973, 2012. ISSN: 1543-8384, DOI: 10.1021/mp3002733.

DARRIBÈRE, T. et al. **Integrins: Regulators of embryogenesis.** *Biology of the Cell*, [s.l.], v. 92, nº 1, p. 5–25, 2000. ISSN: 02484900, DOI: 10.1016/S0248-4900(00)88760-2.

DAVIDCHACK, R. L.; HANDEL, R.; TRETYAKOV, M. V. **Langevin thermostat for rigid body dynamics.** *The Journal of Chemical Physics*, [s.l.],



v. 130, nº 23, p. 234101, 2009. ISSN: 0021-9606, DOI: 10.1063/1.3149788.

DESGROSELLIER, J. S.; CHERESH, D. A. **Integrins in cancer: biological implications and therapeutic opportunities.** *Nature Reviews Cancer*, [s.l.], v. 10, nº 1, p. 9–22, 2010. ISSN: 1474-175X, DOI: 10.1038/nrc2748.

DHALIWAL, D.; SHEPHERD, T. G. **Molecular and cellular mechanisms controlling integrin-mediated cell adhesion and tumor progression in ovarian cancer metastasis: a review.** *Clinical and Experimental Metastasis*, [s.l.], v. 39, nº 2, p. 291–301, 2022. ISBN: 0123456789, ISSN: 15737276, DOI: 10.1007/s10585-021-10136-5.

EDWARDS, D. N. et al. **Integrin  $\alpha 5 \beta 1$  inhibition by ATN-161 reduces neuroinflammation and is neuroprotective in ischemic stroke.** *Journal of Cerebral Blood Flow and Metabolism*, [s.l.], v. 40, nº 8, p. 1695–1708, 2020. ISSN: 15597016, DOI: 10.1177/0271678X19880161.

EL-DIN ABUO-RAHMA, G. A. et al. **Potential repurposed SARS-CoV-2 (COVID-19) infection drugs.** *RSC Advances*, [s.l.], v. 10, nº 45, p. 26895–26916, 2020. ISSN: 20462069, DOI: 10.1039/d0ra05821a.

ELANGBAM, C. S.; QUALLS, C. W.; DAHLGREN, R. R. **Cell Adhesion Molecules—Update.** *Veterinary Pathology*, [s.l.], v. 34, nº 1, p. 61–73, 1997. ISSN: 0300-9858, DOI: 10.1177/030098589703400113.

ELBER, R.; RUYMGAART, A. P.; HESS, B. **SHAKE parallelization.** *The European Physical Journal Special Topics*, [s.l.], v. 200, nº 1, p. 211–223, 2011. ISSN: 1951-6355, DOI: 10.1140/epjst/e2011-01525-9.

ESSMANN, U. et al. **A smooth particle mesh Ewald method.** *The Journal of Chemical Physics*, [s.l.], v. 103, nº 19, p. 8577–8593, 1995. ISSN: 0021-9606, DOI: 10.1063/1.470117.

EVANS, D. J.; HOLIAN, B. L. **The Nose–Hoover thermostat.** *The Journal of Chemical Physics*, [s.l.], v. 83, nº 8, p. 4069–4074, 1985. ISSN: 0021-9606, DOI: 10.1063/1.449071.

FOLEY, B. L.; TESSIER, M. B.; WOODS, R. J. **Carbohydrate force fields.** *Wiley Interdisciplinary Reviews: Computational Molecular Science*, [s.l.], v. 2, nº 4, p. 652–697, 2012. ISSN: 17590876, DOI: 10.1002/wcms.89.

FOLMER, R. H. A. **Drug target residence time: a misleading concept.** *Drug Discovery Today*, [s.l.], v. 23, nº 1, p. 12–16, 2018. ISSN: 13596446, DOI: 10.1016/j.drudis.2017.07.016.

FU, H. et al. **Accurate determination of protein:ligand standard binding free energies from molecular dynamics simulations.** *Nature Protocols*, [s.l.], v. 17, nº 4, p. 1114–1141, 2022. ISSN: 1754-2189, DOI: 10.1038/s41596-021-00676-1.

FU, L. et al. **Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease.** *Nature Communications*, [s.l.], v. 11, nº 1, p. 4417, 2020. ISSN: 2041-1723, DOI: 10.1038/s41467-020-18233-x.

GAHMBERG, C. G.; GRÖNHOLM, M.; MADHAVAN, S. **Regulation of Dynamic Cell Adhesion by Integrin-Integrin Crosstalk.** *Cells*, [s.l.], v. 11, nº

10, p. 1685, 2022. ISSN: 2073-4409, DOI: 10.3390/cells11101685.

GAN, B. H. et al. **The multifaceted nature of antimicrobial peptides: current synthetic chemistry approaches and future directions.** *Chemical Society Reviews*, [s.l.], v. 50, nº 13, p. 7820–7880, 2021. ISSN: 0306-0012, DOI: 10.1039/D0CS00729C.

GARRETT M, M. et al. **AutoDock Version 4.2.** *Citeseer*, [s.l.], p. 1–66, 2012.

GILBERT, J. C.; NOCEDAL, J. **Global Convergence Properties of Conjugate Gradient Methods for Optimization.** *SIAM Journal on Optimization*, [s.l.], v. 2, nº 1, p. 21–42, 1992. ISSN: 1052-6234, DOI: 10.1137/0802003.

GONZÁLEZ, M.A. **Force fields and molecular dynamics simulations.** *École thématique de la Société Française de la Neutronique*, [s.l.], v. 12, p. 169–200, 2011. ISSN: 2107-7223, DOI: 10.1051/sfn/201112009.

GONZÁLEZ, Miguel A.; ABASCAL, J. L. F. **A flexible model for water based on TIP4P/2005.** *The Journal of Chemical Physics*, [s.l.], v. 135, nº 22, p. 224516, 2011. ISSN: 0021-9606, DOI: 10.1063/1.3663219.

GUEDES, I. A.; MAGALHÃES, C. S. DE; DARDENNE, L. E. **Receptor–ligand molecular docking.** *Biophysical Reviews*, [s.l.], v. 6, nº 1, p. 75–87, 2014. ISSN: 1867-2450, DOI: 10.1007/s12551-013-0130-2.

GUNSTEREN, W. F. VAN; BERENDSEN, H. J. C. **A Leap-frog Algorithm for Stochastic Dynamics.** *Molecular Simulation*, [s.l.], v. 1, nº 3, p. 173–185, 1988. ISSN: 0892-7022, DOI: 10.1080/08927028808080941.

HARBOTTLE, R. P. et al. **An RGD–Oligolysine Peptide: A Prototype Construct for Integrin-Mediated Gene Delivery.** *Human Gene Therapy*, [s.l.], v. 9, nº 7, p. 1037–1047, 1998. ISSN: 1043-0342, DOI: 10.1089/hum.1998.9.7-1037.

HERWIG, R. **Computational modeling of drug response with applications to neuroscience.** *Dialogues in Clinical Neuroscience*, [s.l.], v. 16, nº 4, p. 465–477, 2014. ISSN: 1958-5969, DOI: 10.31887/DCNS.2014.16.4/rherwig.

HESS, B. et al. **LINCS: A linear constraint solver for molecular simulations.** *Journal of Computational Chemistry*, [s.l.], v. 18, nº 12, p. 1463–1472, 1997. ISSN: 0192-8651, DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

HOLLINGSWORTH, S. A.; DROR, R. O. **Molecular Dynamics Simulation for All.** *Neuron*, [s.l.], v. 99, nº 6, p. 1129–1143, 2018. ISSN: 08966273, DOI: 10.1016/j.neuron.2018.08.011.

HOU, J. et al. **The Roles of Integrin  $\alpha 5 \beta 1$  in Human Cancer.** *OncoTargets and Therapy*, [s.l.], v. Volume 13, p. 13329–13344, 2020. ISSN: 1178-6930, DOI: 10.2147/OTT.S273803.

HUAN, Y. et al. **Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields.** *Frontiers in Microbiology*, [s.l.], v. 11, 2020. ISSN: 1664-302X, DOI: 10.3389/fmicb.2020.582779.

HUANG, J.; MACKERELL, A. D. **CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data.** *Journal of Computational Chemistry*, [s.l.], v. 34, nº 25, p. 2135–2145, 2013. ISSN: 01928651, DOI: 10.1002/jcc.23354.

HUEY, R. et al. **Software News and Update A Semiempirical Free Energy Force Field with Charge-Based Desolvation.** *Journal of computational chemistry*, [s.l.], v. 32, p. 174–182, 2007. ISBN: 1096-987X, ISSN: 1096-987X, DOI: 10.1002/jcc.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. **VMD: Visual Molecular Dynamics.** *Journal of molecular graphics*, [s.l.], v. 14, p. 33–38, 1996. ISBN: 0263-7855, ISSN: 0263-7855.

HUSSEIN, H. A. M. et al. **Beyond RGD: virus interactions with integrins.** *Archives of Virology*, [s.l.], v. 160, nº 11, p. 2669–2681, 2015. ISSN: 0304-8608, DOI: 10.1007/s00705-015-2579-8.

JACKSON, C. B. et al. **Mechanisms of SARS-CoV-2 entry into cells.** *Nature Reviews Molecular Cell Biology*, [s.l.], v. 23, nº 1, p. 3–20, 2022. ISSN: 1471-0072, DOI: 10.1038/s41580-021-00418-x.

JAIN, A. N.; NICHOLLS, A. **Recommendations for evaluation of computational methods.** *Journal of Computer-Aided Molecular Design*, [s.l.], v. 22, nº 3–4, p. 133–139, 2008. ISSN: 0920-654X, DOI: 10.1007/s10822-008-9196-5.

JIN, Z. et al. **Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors.** *Nature*, [s.l.], v. 582, nº 7811, p. 289–293, 2020. ISSN: 14764687, DOI: 10.1038/s41586-020-2223-y.

JORGENSEN, W. L.; MAXWELL, D. S.; TIRADO-RIVES, J. **Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.** *Journal of the American Chemical Society*, [s.l.], v. 118, nº 45, p. 11225–11236, 1996. ISSN: 00027863, DOI: 10.1021/ja9621760.

KARPLUS, M.; MCCAMMON, J. A. **Molecular dynamics simulations of biomolecules.** *Nature Structural Biology*, [s.l.], v. 9, nº 9, p. 646–652, 2002. ISSN: 10728368, DOI: 10.1038/nsb0902-646.

KÄSTNER, J. **Umbrella sampling.** *Wiley Interdisciplinary Reviews: Computational Molecular Science*, [s.l.], v. 1, nº 6, p. 932–942, 2011. ISSN: 17590884, DOI: 10.1002/wcms.66.

KE, Q. et al. **Effects of thermostats/barostats on physical properties of liquids by molecular dynamics simulations.** *Journal of Molecular Liquids*, [s.l.], v. 365, p. 120116, 2022. ISSN: 01677322, DOI: 10.1016/j.molliq.2022.120116.

KĘSKA, P.; GUSTAW, W.; STADNIK, J. **Trends in In Silico Approaches to the Prediction of Biologically Active Peptides in Meat and Meat Products as an Important Factor for Preventing Food-Related Chronic Diseases.** *Applied Sciences*, [s.l.], v. 11, nº 23, p. 11236, 2021. ISSN: 2076-3417, DOI: 10.3390/app112311236.

KHALILI, P. et al. **A non-RGD-based integrin binding peptide (ATN-161) blocks breast cancer growth and metastasis in vivo.** *Molecular Cancer Therapeutics*, [s.l.], v. 5, nº 9, p. 2271–2280, 2006. ISSN: 15357163, DOI: 10.1158/1535-7163.MCT-06-0100.

KHAN, T. et al. **Computational Drug Designing and Prediction Of Important Parameters Using in silico Methods- A Review.** *Current Computer-Aided Drug Design*, [s.l.], v. 15, nº 5, p. 384–397, 2019. ISSN: 15734099, DOI: 10.2174/1573399815666190326120006.

KIM, S. et al. **PubChem in 2021: New data content and improved web interfaces.** *Nucleic Acids Research*, [s.l.], v. 49, nº D1, p. D1388–D1395, 2021. ISSN: 13624962, DOI: 10.1093/nar/gkaa971.

KOKH, D. B. et al. **Estimation of Drug-Target Residence Times by  $\tau$ -Random Acceleration Molecular Dynamics Simulations.** *Journal of Chemical Theory and Computation*, [s.l.], v. 14, nº 7, p. 3859–3869, 2018. ISSN: 1549-9618, DOI: 10.1021/acs.jctc.8b00230.

KRAUTLER, V.; GUNSTEREN, W. F. VAN; HUNENBERGER, P. H. **A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations.** *Journal of Computational Chemistry*, [s.l.], v. 22, nº 5, p. 501–508, 2001. ISSN: 0192-8651, DOI: 10.1002/1096-987X(20010415)22:5<501::AID-JCC1021>3.0.CO;2-V.

KUMAR, H.; MAITI, P. K. **Introduction to Molecular Dynamics Simulation.** *Statistical Physics*. [s.l.]: [s.n.], 2011. p. 190.

KUMAR, S. et al. **The weighted histogram analysis method for free-energy calculations on biomolecules.** *Journal of computational chemistry*, [s.l.], v. 13, nº 8, p. 1011–1021, 1992.

LAN, J. et al. **Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor.** *Nature*, [s.l.], v. 581, nº 7807, p. 215–220, 2020. ISSN: 14764687, DOI: 10.1038/s41586-020-2180-5.

LEMKUL, J. A.; BEVAN, D. R. **Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics.** *Journal of Physical Chemistry B*, [s.l.], v. 114, nº 4, p. 1652–1660, 2010. ISSN: 15206106, DOI: 10.1021/jp9110794.

LETKO, M.; MARZI, A.; MUNSTER, V. **Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses.** *Nature Microbiology*, [s.l.], v. 5, nº 4, p. 562–569, 2020. ISSN: 2058-5276, DOI: 10.1038/s41564-020-0688-y.

LI, T. et al. **Characterization and antimicrobial mechanism of CF-14, a new antimicrobial peptide from the epidermal mucus of catfish.** *Fish & Shellfish Immunology*, [s.l.], v. 92, p. 881–888, 2019. ISSN: 10504648, DOI: 10.1016/j.fsi.2019.07.015.

LIMA, M. et al. **A QUÍMICA DOS SANEANTES EM TEMPOS DE COVID-19: VOCÊ SABE COMO ISSO FUNCIONA?** *Química Nova*, [s.l.], v. 43, p. 668–678, 2020. ISSN: 01004042, DOI: 10.21577/0100-4042.20170552.

LINDAHL, E. R. **Molecular Dynamics Simulations**. In: KUKOL, A. (Org.). *Molecular Modeling of Proteins*. Totowa: Humana Press, 2008. p. 3–23. DOI: 10.1007/978-1-59745-177-2\_1.

LIU, J. et al. **Integrin mediates cell entry of the SARS-CoV-2 virus independent of cellular receptor ACE2**. *Journal of Biological Chemistry*, [s.l.], v. 298, nº 3, p. 101710, 2022. ISSN: 1083351X, DOI: 10.1016/j.jbc.2022.101710.

LIU, Z. et al. **The assessment of transmission efficiency and latent infection period in asymptomatic carriers of SARS-CoV-2 infection**. *International Journal of Infectious Diseases*, [s.l.], v. 99, p. 325–327, 2020. ISSN: 12019712, DOI: 10.1016/j.ijid.2020.06.036.

LUO, B.-H.; SPRINGER, T. A. **Integrin structures and conformational signaling**. *Current Opinion in Cell Biology*, [s.l.], v. 18, nº 5, p. 579–586, 2006. ISSN: 09550674, DOI: 10.1016/j.ceb.2006.08.005.

LUO, Y.; SONG, Y. **Mechanism of Antimicrobial Peptides: Antimicrobial, Anti-Inflammatory and Antibiofilm Activities**. *International Journal of Molecular Sciences*, [s.l.], v. 22, nº 21, p. 11401, 2021. ISSN: 1422-0067, DOI: 10.3390/ijms222111401.

LUSTIG, R. **Statistical thermodynamics in the classical molecular dynamics ensemble. I. Fundamentals**. *The Journal of Chemical Physics*, [s.l.], v. 100, nº 4, p. 3048–3059, 1994. ISSN: 0021-9606, DOI: 10.1063/1.466446.

LV, X. et al. **ATN-161 reduces virus proliferation in PHEV-infected mice by inhibiting the integrin  $\alpha 5 \beta 1$ -FAK signaling pathway**. *Veterinary Microbiology*, [s.l.], v. 233, p. 147–153, 2019. ISSN: 03781135, DOI: 10.1016/j.vetmic.2019.04.029.

LYUBARTSEV, A. P.; RABINOVICH, A. L. **Force Field Development for Lipid Membrane Simulations**. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, [s.l.], v. 1858, nº 10, p. 2483–2497, 2016. ISSN: 00052736, DOI: 10.1016/j.bbamem.2015.12.033.

MACKERELL, A. D. et al. **All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins**. *The Journal of Physical Chemistry B*, [s.l.], v. 102, nº 18, p. 3586–3616, 1998. ISSN: 1520-6106, DOI: 10.1021/jp973084f.

MACKERELL, A. D. **Empirical force fields for biological macromolecules: Overview and issues**. *Journal of Computational Chemistry*, [s.l.], v. 25, nº 13, p. 1584–1604, 2004. ISSN: 0192-8651, DOI: 10.1002/jcc.20082.

MADEWELL, Z. J. et al. **Household Transmission of SARS-CoV-2**. *JAMA Network Open*, [s.l.], v. 3, nº 12, p. e2031756, 2020. ISSN: 2574-3805, DOI: 10.1001/jamanetworkopen.2020.31756.

MAIER, H. J.; BICKERTON, E.; BRITTON, P. **Coronaviruses**. In: MAIER, H. J.; BICKERTON, E.; BRITTON, P. (Orgs.). New York, NY: Springer New York, 2015. v. 1282, 1–23 p. ISBN: 978-1-4939-2437-0, DOI: 10.1007/978-1-4939-2438-7.

MAIER, J. A. et al. **ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB.** *Journal of Chemical Theory and Computation*, [s.l.], v. 11, nº 8, p. 3696–3713, 2015. ISSN: 1549-9618, DOI: 10.1021/acs.jctc.5b00255.

MAKOWSKI, L.; OLSON-SIDFORD, W.; WEISEL, J. W. **Biological and clinical consequences of integrin binding via a rogue rgd motif in the sars cov-2 spike protein.** *Viruses*, [s.l.], v. 13, nº 2, 2021. ISSN: 19994915, DOI: 10.3390/v13020146.

MARTYNA, G. J.; KLEIN, M. L.; TUCKERMAN, M. **Nosé–Hoover chains: The canonical ensemble via continuous dynamics.** *The Journal of Chemical Physics*, [s.l.], v. 97, nº 4, p. 2635–2643, 1992. ISSN: 0021-9606, DOI: 10.1063/1.463940.

MASERAS, F.; MOROKUMA, K. **IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states.** *Journal of Computational Chemistry*, [s.l.], v. 16, nº 9, p. 1170–1179, 1995. ISBN: 1096-987X, ISSN: 0192-8651, DOI: 10.1002/jcc.540160911.

MCEVER, R. P. **Selectins.** *Current Opinion in Immunology*, [s.l.], v. 6, nº 1, p. 75–84, 1994. ISSN: 09527915, DOI: 10.1016/0952-7915(94)90037-X.

MENACHERY, V. D. et al. **SARS-like WIV1-CoV poised for human emergence.** *Proceedings of the National Academy of Sciences of the United States of America*, [s.l.], v. 113, nº 11, p. 3048–3053, 2016. ISSN: 10916490, DOI: 10.1073/pnas.1517719113.

MEYER, A. et al. **Targeting RGD Recognizing Integrins: Drug Development, Biomaterial Research, Tumor Imaging and Targeting.** *Current Pharmaceutical Design*, [s.l.], v. 12, nº 22, p. 2723–2747, 2006. ISSN: 13816128, DOI: 10.2174/138161206777947740.

MEZA, J. C. **Steepest descent.** *WIREs Computational Statistics*, [s.l.], v. 2, nº 6, p. 719–722, 2010. ISSN: 1939-5108, DOI: 10.1002/wics.117.

MILLS, M.; ANDRICIOAEI, I. **An experimentally guided umbrella sampling protocol for biomolecules.** *The Journal of Chemical Physics*, [s.l.], v. 129, nº 11, p. 114101, 2008. ISSN: 0021-9606, DOI: 10.1063/1.2976440.

MIRZAEI, M. **Science and engineering in antiquity.** *Advanced journal of science and engineering*, [s.l.], v. 1, nº 1, p. 1–2, 2020. ISSN: 00280836, DOI: 10.1038/142822a0.

MONTICELLI, L.; TIELEMAN, D. P. **Force Fields for Classical Molecular Dynamics.** [s.l.]: [s.n.], 2013. p. 197–213. DOI: 10.1007/978-1-62703-017-5\_8.

MORADI, M. et al. **A contemporary review on the important role of in silico approaches for managing different aspects of COVID-19 crisis.** *Informatics in Medicine Unlocked*, [s.l.], v. 28, p. 100862, 2022. ISSN: 23529148, DOI: 10.1016/j.imu.2022.100862.

MOROKUMA, K. et al. **Model studies of the structures, reactivities, and**

**reaction mechanisms of metalloenzymes.** *IBM Journal of Research and Development*, [s.l.], v. 45, nº 3.4, p. 367–395, 2001. ISBN: 0018-8646, ISSN: 0018-8646, DOI: 10.1147/rd.453.0367.

MORRIS, G. M. et al. **Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.** *Journal of Computational Chemistry*, [s.l.], v. 19, nº 14, p. 1639–1662, 1998a. ISSN: 01928651, DOI: 10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.

\_\_\_\_\_. **Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.** *Journal of Computational Chemistry*, [s.l.], v. 19, nº 14, p. 1639–1662, 1998b. ISBN: 0192-8651, ISSN: 01928651, DOI: 10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.

\_\_\_\_\_. **AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.** *Journal of Computational Chemistry*, [s.l.], v. 30, nº 16, p. 2785–2791, 2009. ISSN: 01928651, DOI: 10.1002/jcc.21256.

MORRIS, G. M. et al. **AutoDock Version 4.2: Updated for version 4.2.6.** *User Guide*, [s.l.], p. 1–49, 2014.

MOTHAY, D.; RAMESH, K. V. **Binding site analysis of potential protease inhibitors of COVID-19 using AutoDock.** *VirusDisease*, [s.l.], v. 31, nº 2, p. 194–199, 2020. ISBN: 1333702000, ISSN: 23473517, DOI: 10.1007/s13337-020-00585-z.

MYERS, S.; BAKER, A. **Drug discovery—an operating model for a new era.** *Nature Biotechnology*, [s.l.], v. 19, nº 8, p. 727–730, 2001. ISSN: 1087-0156, DOI: 10.1038/90765.

NAGAE, M. et al. **Crystal structure of  $\alpha 5 \beta 1$  integrin ectodomain: Atomic details of the fibronectin receptor.** *Journal of Cell Biology*, [s.l.], v. 197, nº 1, p. 131–140, 2012. ISSN: 00219525, DOI: 10.1083/jcb.201111077.

NAMBA, A. M.; SILVA, V. B. Da; SILVA, C. H. T. P. Da. **Dinâmica molecular: teoria e aplicações em planejamento de fármacos.** *Eclética Química*, [s.l.], v. 33, nº 4, p. 13–24, 2008. ISSN: 0100-4670, DOI: 10.1590/S0100-46702008000400002.

NGUYEN, H. L. et al. **Remdesivir strongly binds to both RNA-dependent RNA polymerase and main protease of SARS-COV-2: Evidence from molecular simulations.** *Journal of Physical Chemistry B*, [s.l.], v. 124, nº 50, p. 11337–11348, 2020. ISSN: 15205207, DOI: 10.1021/ACS.JPCB.0C07312/SUPPL\_FILE/JP0C07312\_SI\_001.PDF.

OBARA, M.; KANG, M. S.; YAMADA, K. M. **Site-directed mutagenesis of the cell-binding domain of human fibronectin: Separable, synergistic sites mediate adhesive function.** *Cell*, [s.l.], v. 53, nº 4, p. 649–657, 1988. ISSN: 00928674, DOI: 10.1016/0092-8674(88)90580-6.

OKEGAWA, T. et al. **The role of cell adhesion molecule in cancer progression and its application in cancer therapy.** *Acta Biochimica Polonica*, [s.l.], v. 51, nº 2, p. 445–457, 2004. ISSN: 1734-154X, DOI:

10.18388/abp.2004\_3583.

OOSTENBRINK, C. et al. **A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6.** *Journal of Computational Chemistry*, [s.l.], v. 25, nº 13, p. 1656–1676, 2004. ISSN: 0192-8651, DOI: 10.1002/jcc.20090.

OTHMAN, H. et al. **SARS-CoV-2 Spike Protein Unlikely to Bind to Integrins via the Arg-Gly-Asp (RGD) Motif of the Receptor Binding Domain: Evidence From Structural Analysis and Microscale Accelerated Molecular Dynamics.** *Frontiers in Molecular Biosciences*, [s.l.], v. 9, 2022. ISSN: 2296889X, DOI: 10.3389/fmolb.2022.834857.

OWEN, D. R. et al. **An oral SARS-CoV-2 MPro inhibitor clinical candidate for the treatment of COVID-19.** *Science*, [s.l.], v. 3, p. 1586–1593, 2021.

PAGADALA, N. S.; SYED, K.; TUSZYNSKI, J. **Software for molecular docking: a review.** *Biophysical Reviews*, [s.l.], v. 9, nº 2, p. 91–102, 2017. ISSN: 1867-2450, DOI: 10.1007/s12551-016-0247-1.

PARK, E. J. et al. **The spike glycoprotein of sars-cov-2 binds to  $\beta 1$  integrins expressed on the surface of lung epithelial cells.** *Viruses*, [s.l.], v. 13, nº 4, 2021. ISSN: 19994915, DOI: 10.3390/v13040645.

PARK, S. et al. **Structural study of novel antimicrobial peptides, nigrocins, isolated from *Rana nigromaculata*.** *FEBS Letters*, [s.l.], v. 507, nº 1, p. 95–100, 2001. ISSN: 00145793, DOI: 10.1016/S0014-5793(01)02956-8.

PARRINELLO, M.; RAHMAN, A. **Polymorphic transitions in single crystals: A new molecular dynamics method.** *Journal of Applied Physics*, [s.l.], v. 52, nº 12, p. 7182–7190, 1981. ISSN: 00218979, DOI: 10.1063/1.328693.

PENG, R. et al. **Cell entry by SARS-CoV-2.** *Trends in Biochemical Sciences*, [s.l.], v. 46, nº 10, p. 848–860, 2021. ISSN: 09680004, DOI: 10.1016/j.tibs.2021.06.001.

PETERSEN, H. G. **Accuracy and efficiency of the particle mesh Ewald method.** *The Journal of Chemical Physics*, [s.l.], v. 103, nº 9, p. 3668–3679, 1995. ISSN: 00219606, DOI: 10.1063/1.470043.

PETERSON, R. J.; KOVAL, M. **Above the Matrix: Functional Roles for Apically Localized Integrins.** *Frontiers in Cell and Developmental Biology*, [s.l.], v. 9, 2021. ISSN: 2296-634X, DOI: 10.3389/fcell.2021.699407.

PETTERSEN, E. F. et al. **UCSF Chimera - A visualization system for exploratory research and analysis.** *Journal of Computational Chemistry*, [s.l.], v. 25, nº 13, p. 1605–1612, 2004. ISSN: 01928651, DOI: 10.1002/jcc.20084.

PONDER, J. W.; CASE, D. A. **Force Fields for Protein Simulations.** *Advances in Protein Chemistry*. [s.l.]: [s.n.], 2003. p. 27–85. DOI: 10.1016/S0065-3233(03)66002-X.

PRICE, D. J.; BROOKS, C. L. **A modified TIP3P water potential for simulation with Ewald summation.** *The Journal of Chemical Physics*, [s.l.], v. 121, nº 20, p. 10096–10103, 2004. ISSN: 0021-9606, DOI:



10.1063/1.1808117.

RAMOS, R. et al. **Recombinant expression and purification of the antimicrobial peptide magainin-2.** *Biotechnology Progress*, [s.l.], v. 29, nº 1, p. 17–22, 2013. ISSN: 87567938, DOI: 10.1002/btpr.1650.

RIDER, T. H. et al. **Broad-spectrum antiviral therapeutics.** *PLoS ONE*, [s.l.], v. 6, nº 7, 2011. ISSN: 19326203, DOI: 10.1371/journal.pone.0022572.

RIVELLESE, F.; PREDILETTO, E. **ACE2 at the centre of COVID-19 from paucisymptomatic infections to severe pneumonia.** *Autoimmunity Reviews*, [s.l.], v. 19, nº 6, p. 102536, 2020. ISSN: 15689972, DOI: 10.1016/j.autrev.2020.102536.

ROY, F. VAN; BERX, G. **The cell-cell adhesion molecule E-cadherin.** *Cellular and Molecular Life Sciences*, [s.l.], v. 65, nº 23, p. 3756–3788, 2008. ISSN: 1420-682X, DOI: 10.1007/s00018-008-8281-1.

RUBIN, D. et al. **FDA Approval of Remdesivir - A Step in the Right Direction.** *New England Journal of Medicine*, [s.l.], v. 383, nº 27, p. 2595–2598, 2020. ISSN: 0028-4793, DOI: 10.1056/nejmp2030646.

RUOSLAHTI, E. **RGD AND OTHER RECOGNITION SEQUENCES FOR INTEGRINS.** *Annual Review of Cell and Developmental Biology*, [s.l.], v. 12, nº 1, p. 697–715, 1996. ISSN: 1081-0706, DOI: 10.1146/annurev.cellbio.12.1.697.

SALSBURY, A. M.; LEMKUL, J. A. **Recent developments in empirical atomistic force fields for nucleic acids and applications to studies of folding and dynamics.** *Current Opinion in Structural Biology*, [s.l.], v. 67, p. 9–17, 2021. ISSN: 0959440X, DOI: 10.1016/j.sbi.2020.08.003.

SANI, S. et al. **Biological Relevance of RGD-Integrin Subtype-Specific Ligands in Cancer.** *ChemBioChem*, [s.l.], v. 22, nº 7, p. 1151–1160, 2021. ISSN: 1439-4227, DOI: 10.1002/cbic.202000626.

SANT'ANNA, C. M. R. **Molecular modeling methods in the study and design of bioactive compounds: An introduction.** *Revista Virtual de Química*, [s.l.], v. 1, nº 1, 2009. ISSN: 1984-6835, DOI: 10.5935/1984-6835.20090007.

SCHAFFNER, F.; RAY, A.; DONTENWILL, M. **Integrin  $\alpha 5 \beta 1$ , the Fibronectin Receptor, as a Pertinent Therapeutic Target in Solid Tumors.** *Cancers*, [s.l.], v. 5, nº 4, p. 27–47, 2013. ISSN: 2072-6694, DOI: 10.3390/cancers5010027.

SCHUMACHER, S. et al. **Structural insights into integrin  $\alpha 5 \beta 1$  opening by fibronectin ligand.** *Science Advances*, [s.l.], v. 7, nº 19, p. 1–18, 2021. ISSN: 23752548, DOI: 10.1126/sciadv.abe9716.

SENN, H. M.; THIEL, W. **QM/MM methods for biomolecular systems.** *Angewandte Chemie - International Edition*, [s.l.], v. 48, nº 7, p. 1198–1229, 2009. ISBN: 1521-3773, ISSN: 14337851, DOI: 10.1002/anie.200802019.

SHEREEN, M. A. et al. **COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses.** *Journal of Advanced Research*,

[s.l.], v. 24, p. 91–98, 2020. ISSN: 20901232, DOI: 10.1016/j.jare.2020.03.005.

SHI, Z.; HU, Z. **A review of studies on animal reservoirs of the SARS coronavirus.** *Virus Research*, [s.l.], v. 133, nº 1, p. 74–87, 2008. ISSN: 01681702, DOI: 10.1016/j.virusres.2007.03.012.

SILVA, R. S. et al. **Absolute binding free energies of the antiviral peptide ATN-161 with protein targets of SARS-CoV-2.** *Journal of Biomolecular Structure and Dynamics*, [s.l.], p. 1–12, 2022. ISSN: 0739-1102, DOI: 10.1080/07391102.2022.2154848.

SINGANAYAGAM, A. et al. **Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study.** *The Lancet Infectious Diseases*, [s.l.], v. 22, nº 2, p. 183–195, 2022. ISSN: 14744457, DOI: 10.1016/S1473-3099(21)00648-4.

SOHRABY, F.; NUNES-ALVES, A. **Advances in computational methods for ligand binding kinetics.** *Trends in Biochemical Sciences*, [s.l.], 2022. ISSN: 09680004, DOI: 10.1016/j.tibs.2022.11.003.

SPERANDIO, O. et al. **Receptor-Based Computational Screening of Compound Databases: The Main Docking-Scoring Engines.** *Current protein and Peptide Science*, [s.l.], v. 7, nº 5, p. 369–393, 2006. ISSN: 13892037, DOI: 10.2174/138920306778559377.

SPOEL, D. VAN DER et al. **GROMACS: Fast, flexible, and free.** *Journal of Computational Chemistry*, [s.l.], v. 26, nº 16, p. 1701–1718, 2005. ISSN: 01928651, DOI: 10.1002/jcc.20291.

STASI, C. et al. **SARS-CoV-2 Vaccination and the Bridge between First and Fourth Dose: Where Are We?** *Vaccines*, [s.l.], v. 10, nº 3, p. 1–21, 2022. ISSN: 2076393X, DOI: 10.3390/vaccines10030444.

STOELTZING, O. et al. **Inhibition of integrin  $\alpha 5 \beta 1$  function with a small peptide (ATN-161) plus continuous 5-fu infusion reduces colorectal liver metastases and improves survival in mice.** *International Journal of Cancer*, [s.l.], v. 104, nº 4, p. 496–503, 2003. ISBN: 7137924689, ISSN: 00207136, DOI: 10.1002/ijc.10958.

SUN, H. **COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase Applications Overview with Details on Alkane and Benzene Compounds.** *The Journal of Physical Chemistry B*, [s.l.], v. 102, nº 38, p. 7338–7364, 1998. ISSN: 1520-6106, DOI: 10.1021/jp980939v.

TAKAGI, J. et al. **Structure of integrin  $\alpha 5 \beta 1$  in complex with fibronectin.** *EMBO Journal*, [s.l.], v. 22, nº 18, p. 4607–4615, 2003. ISSN: 02614189, DOI: 10.1093/emboj/cdg445.

TOBIAS, D. J.; MARTYNA, G. J.; KLEIN, M. L. **Molecular dynamics simulations of a protein in the canonical ensemble.** *The Journal of Physical Chemistry*, [s.l.], v. 97, nº 49, p. 12959–12966, 1993. ISSN: 0022-3654, DOI: 10.1021/j100151a052.

TONK, M.; RŮŽEK, D.; VILCINSKAS, A. **Compelling evidence for the**

**activity of antiviral peptides against SARS-CoV-2.** *Viruses*, [s.l.], v. 13, nº 5, 2021. ISSN: 19994915, DOI: 10.3390/v13050912.

TONNER, R. **Molecular Modeling Basics.** By Jan H. Hensen. *ChemPhysChem*, [s.l.], v. 12, nº 12, p. 2352–2352, 2011. ISSN: 14394235, DOI: 10.1002/cphc.201000744.

TORRIE, G. M.; VALLEAU, J. P. **Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling.** *Journal of Computational Physics*, [s.l.], v. 23, nº 2, p. 187–199, 1977. ISSN: 00219991, DOI: 10.1016/0021-9991(77)90121-8.

TRABOCCHI, A.; GUARNA, A. **Peptidomimetics in Organic and Medicinal Chemistry.** Chichester, UK: John Wiley & Sons, Ltd, 2014. ISBN: 9781118683033, DOI: 10.1002/9781118683033.

TRZPIS, M. et al. **Epithelial Cell Adhesion Molecule.** *The American Journal of Pathology*, [s.l.], v. 171, nº 2, p. 386–395, 2007. ISSN: 00029440, DOI: 10.2353/ajpath.2007.070152.

V'KOVSKI, P. et al. **Coronavirus biology and replication: implications for SARS-CoV-2.** *Nature Reviews Microbiology*, [s.l.], v. 19, nº 3, p. 155–170, 2021. ISSN: 1740-1526, DOI: 10.1038/s41579-020-00468-6.

VEGA, C. et al. **What ice can teach us about water interactions: a critical comparison of the performance of different water models.** *Faraday Discuss.*, [s.l.], v. 141, p. 251–276, 2009. ISSN: 1359-6640, DOI: 10.1039/B805531A.

VERLET, L. **Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules.** *Physical Review*, [s.l.], v. 159, nº 1, p. 98–103, 1967. ISSN: 0031-899X, DOI: 10.1103/PhysRev.159.98.

WANG, J. et al. **Development and testing of a general amber force field.** *Journal of Computational Chemistry*, [s.l.], v. 25, nº 9, p. 1157–1174, 2004. ISSN: 0192-8651, DOI: 10.1002/jcc.20035.

WANG, Q. et al. **Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2.** *Cell*, [s.l.], v. 181, nº 4, p. 894-904.e9, 2020. ISSN: 00928674, DOI: 10.1016/j.cell.2020.03.045.

WANG, W. Q. et al. **Joint Antiangiogenic Effect of ATN-161 and Anti-VEGF Antibody in a Rat Model of Early Wet Age-Related Macular Degeneration.** *Molecular Pharmaceutics*, [s.l.], v. 13, nº 9, p. 2881–2890, 2016. ISSN: 1543-8384, DOI: 10.1021/acs.molpharmaceut.6b00056.

WHO. **Therapeutics and COVID-19: Living guideline, 22 April 2022.** *World Health Organization*, [s.l.], v. 4, nº July, p. 1–142, 2022.

WIERSINGA, W. J. et al. **Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19).** *JAMA*, [s.l.], v. 324, nº 8, p. 782, 2020. ISSN: 0098-7484, DOI: 10.1001/jama.2020.12839.

WILLIAMS, A. F. **A year in the life of the immunoglobulin superfamily.** *Immunology Today*, [s.l.], v. 8, nº 10, p. 298–303, 1987. ISSN: 01675699, DOI:

10.1016/0167-5699(87)90016-8.

WU, F. et al. **A new coronavirus associated with human respiratory disease in China.** *Nature*, [s.l.], v. 579, nº 7798, p. 265–269, 2020. ISSN: 0028-0836, DOI: 10.1038/s41586-020-2008-3.

XIA, W.; SPRINGER, T. A. **Metal ion and ligand binding of Integrin  $\alpha 5\beta 1$ .** *Proceedings of the National Academy of Sciences of the United States of America*, [s.l.], v. 111, nº 50, p. 17863–17868, 2014. ISSN: 10916490, DOI: 10.1073/pnas.1420645111.

XIU, S. et al. **Inhibitors of SARS-CoV-2 Entry: Current and Future Opportunities.** *Journal of Medicinal Chemistry*, [s.l.], v. 63, nº 21, p. 12256–12274, 2020. ISSN: 15204804, DOI: 10.1021/acs.jmedchem.0c00502.

XU, R. et al. **Epidemiologic Clues to SARS Origin in China.** *Emerging Infectious Diseases*, [s.l.], v. 10, nº 6, p. 0–1, 2004.

YABE, M. et al. **Development of PolyParGen Software to Facilitate the Determination of Molecular Dynamics Simulation Parameters for Polymers.** *Journal of Computer Chemistry, Japan -International Edition*, [s.l.], v. 5, nº 0, p. n/a, 2019. DOI: 10.2477/jccjie.2018-0034.

YANG, H.; YANG, J. **A review of the latest research on M pro targeting SARS-COV inhibitors.** *RSC Medicinal Chemistry*, [s.l.], v. 12, nº 7, p. 1026–1036, 2021. ISSN: 2632-8682, DOI: 10.1039/D1MD00066G.

YAO, H.; VEINE, D. M.; LIVANT, D. L. **Therapeutic inhibition of breast cancer bone metastasis progression and lung colonization: breaking the vicious cycle by targeting  $\alpha 5\beta 1$  integrin.** *Breast Cancer Research and Treatment*, [s.l.], v. 157, nº 3, p. 489–501, 2016. ISSN: 15737217, DOI: 10.1007/s10549-016-3844-6.

YIN, W. et al. **Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody.** *Science*, [s.l.], v. 375, nº 6584, p. 1048–1053, 2022. ISSN: 10959203, DOI: 10.1126/science.abn8863.

ZASLOFF, M. **Antimicrobial peptides of multicellular organisms.** *Nature*, [s.l.], v. 415, nº 6870, p. 389–395, 2002. ISSN: 0028-0836, DOI: 10.1038/415389a.

ZHOU, P. et al. **A pneumonia outbreak associated with a new coronavirus of probable bat origin.** *Nature*, [s.l.], v. 579, nº 7798, p. 270–273, 2020. ISBN: 4158602020127, ISSN: 14764687, DOI: 10.1038/s41586-020-2012-7.

ZHU, N. et al. **A Novel Coronavirus from Patients with Pneumonia in China, 2019.** *New England Journal of Medicine*, [s.l.], v. 382, nº 8, p. 727–733, 2020. ISSN: 0028-4793, DOI: 10.1056/nejmoa2001017.

ZUMLA, A.; HUI, D. S.; PERLMAN, S. **Middle East respiratory syndrome.** *The Lancet*, [s.l.], v. 386, nº 9997, p. 995–1007, 2015. ISSN: 01406736, DOI: 10.1016/S0140-6736(15)60454-8.

## Anexos

### Arquivos de entrada ancoramento molecular.

#### ❖ Parâmetros de grade (GPF)

```
npts 126 126 126          # num.grid points in xyz
gridfld 4wk4_Ca_Mg.maps.fld  # grid_data_file
spacing 0.375              # spacing(A)
receptor_types A C Ca H HD Mg N NA OA SA # receptor atom types
ligand_types A C HD N NA OA SA # ligand atom types
receptor 4wk4_Ca_Mg.pdbqt   # macromolecule
gridcenter 22.286 11.004 -19.274 # xyz-coordinates or auto
smooth 0.5                  # store minimum energy w/in rad(A)
map 4wk4_Ca_Mg.A.map        # atom-specific affinity map
map 4wk4_Ca_Mg.C.map        # atom-specific affinity map
map 4wk4_Ca_Mg.HD.map       # atom-specific affinity map
map 4wk4_Ca_Mg.N.map        # atom-specific affinity map
map 4wk4_Ca_Mg.NA.map       # atom-specific affinity map
map 4wk4_Ca_Mg.OA.map       # atom-specific affinity map
map 4wk4_Ca_Mg.SA.map       # atom-specific affinity map
elecmap 4wk4_Ca_Mg.e.map    # electrostatic potential map
dsolvmap 4wk4_Ca_Mg.d.map   # desolvation potential map
dielectric -0.1465          # <0, AD4 distance-dep.diel;>0, constant
```

#### ❖ Parâmetros de ancoragem (DPF)

```
autodock_parameter_version 4.2 # used by autodock to validate parameter set
parameter_file AD4.1_bound.dat # parameter library filename
intelec # calculate internal electrostatics
seed pid time # seeds for random generator
ligand_types A C NA OA N SA HD # atoms types in ligand
fld 4wk4_Ca_Mg.maps.fld # grid_data_file
map 4wk4_Ca_Mg.A.map # atom-specific affinity map
map 4wk4_Ca_Mg.C.map # atom-specific affinity map
map 4wk4_Ca_Mg.NA.map # atom-specific affinity map
map 4wk4_Ca_Mg.OA.map # atom-specific affinity map
map 4wk4_Ca_Mg.N.map # atom-specific affinity map
map 4wk4_Ca_Mg.SA.map # atom-specific affinity map
map 4wk4_Ca_Mg.HD.map # atom-specific affinity map
elecmap 4wk4_Ca_Mg.e.map # electrostatics map
desolvmap 4wk4_Ca_Mg.d.map # desolvation map
move ATN.pdbqt # small molecule
about 1.4065 5.9535 6.4482 # small molecule center
tran0 random # initial coordinates/A or random
quaternion0 random # initial orientation
dihe0 random # initial dihedrals (relative) or random
torsdof 17 # torsional degrees of freedom
rmstol 2.0 # cluster_tolerance/A
extnrg 1000.0 # external grid energy
e0max 0.0 10000 # max initial energy; max number of retries
ga_pop_size 150 # number of individuals in population
ga_num_evals 25000000 # maximum number of energy evaluations
ga_num_generations 27000 # maximum number of generations
ga_elitism 1 # number of top individuals to survive to next generation
ga_mutation_rate 0.02 # rate of gene mutation
ga_crossover_rate 0.8 # rate of crossover
ga_window_size 10 #
```

```

ga_cauchy_alpha 0.0      # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0       # Beta parameter Cauchy distribution
set_ga                # set the above parameters for GA or LGA
sw_max_its 300           # iterations of Solis & Wets local search
sw_max_succ 4            # consecutive successes before changing rho
sw_max_fail 4            # consecutive failures before changing rho
sw_rho 1.0               # size of local search space to sample
sw_lb_rho 0.01           # lower bound on rho
ls_search_freq 0.06      # probability of performing local search on individual
set_psw1               # set the above pseudo-Solis & Wets parameters
unbound_model bound      # state of unbound ligand
ga_run 100               # do this many hybrid GA-LS runs
analysis                # perform a ranked cluster analysis

```

### Arquivos de entrada dinâmica molecular.

#### ❖ Parâmetros de dinâmica molecular - NVT.mdp

```

title                = Protein-ligand complex NVT equilibration
define               = -DPOSRES ; position restrain the protein and ligand
; Run parameters
integrator           = md        ; leap-frog integrator
nsteps               = 50000     ; 2 * 50000 = 100 ps
dt                   = 0.002     ; 2 fs
; Output control
nstenergy            = 500       ; save energies every 1.0 ps
nstlog               = 500       ; update log file every 1.0 ps
nstxout-compressed   = 500       ; save coordinates every 1.0 ps
; Bond parameters
continuation         = no        ; first dynamics run
constraint_algorithm = lincs     ; holonomic constraints
constraints           = h-bonds  ; bonds to H are constrained
lincs_iter           = 1         ; accuracy of LINCS
lincs_order          = 4         ; also related to accuracy
; Neighbor searching and vdW
cutoff-scheme        = Verlet
ns_type               = grid     ; search neighboring grid cells
nstlist              = 20        ; largely irrelevant with Verlet
rlist                 = 1.2
vdwtype               = cutoff
vdw-modifier          = force-switch
rvdw-switch           = 1.0
rvdw                  = 1.2      ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype           = PME       ; Particle Mesh Ewald for long-range electrostatics
rcoulomb              = 1.2       ; short-range electrostatic cutoff (in nm)
pme_order             = 4         ; cubic interpolation
fourierspacing        = 0.16     ; grid spacing for FFT
; Temperature coupling
tcoupl                = V-rescale ; modified Berendsen thermostat
tc-grps               = Protein_ATN Water_and_ions ; two coupling groups - more accurate
tau_t                 = 0.1 0.1   ; time constant, in ps
ref_t                  = 310 310   ; reference temperature, one for each group, in K
; Pressure coupling
pcoupl                = no        ; no pressure coupling in NVT
; Periodic boundary conditions
pbc                   = xyz       ; 3-D PBC
; Dispersion correction is not used for proteins with the C36 additive FF
DispCorr              = no
; Velocity generation

```

```

gen_vel      = yes      ; assign velocities from Maxwell distribution
gen_temp     = 310      ; temperature for Maxwell distribution
gen_seed     = -1       ; generate a random seed

```

❖ Parâmetros de dinâmica molecular - NPT.mdp

```

title        = Protein-ligand complex NPT equilibration
define       = -DPOSRES ; position restrain the protein and ligand
; Run parameters
integrator   = md       ; leap-frog integrator
nsteps       = 50000    ; 2 * 50000 = 100 ps
dt           = 0.002    ; 2 fs
; Output control
nstenergy    = 500      ; save energies every 1.0 ps
nstlog       = 500      ; update log file every 1.0 ps
nstxout-compressed = 500 ; save coordinates every 1.0 ps
; Bond parameters
continuation = yes      ; continuing from NVT
constraint_algorithm = lincs ; holonomic constraints
constraints  = h-bonds  ; bonds to H are constrained
lincs_iter   = 1        ; accuracy of LINCS
lincs_order  = 4        ; also related to accuracy
; Neighbor searching and vdW
cutoff-scheme = Verlet
ns_type       = grid    ; search neighboring grid cells
nstlist       = 20      ; largely irrelevant with Verlet
rlist         = 1.2
vdwtype       = cutoff
vdw-modifier  = force-switch
rvdw-switch   = 1.0
rvdw          = 1.2     ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype   = PME      ; Particle Mesh Ewald for long-range electrostatics
rcoulomb      = 1.2
pme_order     = 4        ; cubic interpolation
fourierspacing = 0.16   ; grid spacing for FFT
; Temperature coupling
tcoupl        = V-rescale ; modified Berendsen thermostat
tc-grps       = Protein_ATN Water_and_ions ; two coupling groups - more accurate
tau_t         = 0.1 0.1 ; time constant, in ps
ref_t         = 310 310 ; reference temperature, one for each group, in K
; Pressure coupling
pcoupl        = Berendsen ; pressure coupling is on for NPT
pcoupltype    = isotropic ; uniform scaling of box vectors
tau_p         = 2.0      ; time constant, in ps
ref_p         = 1.0      ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc           = xyz      ; 3-D PBC
; Dispersion correction is not used for proteins with the C36 additive FF
DispCorr      = no
; Velocity generation
gen_vel       = no      ; velocity generation off after NVT

```

❖ Parâmetros de dinâmica molecular MD\_prod.mdp.

```

title        = Protein-ligand complex MD simulation

```

```

; Run parameters
integrator      = md      ; leap-frog integrator
nsteps         = 100000000 ; 2 * 100000000 = 200000 ps (200 ns)
dt             = 0.002    ; 2 fs
; Output control
nstenergy      = 50000    ; save energies every 500.0 ps
nstlog         = 50000    ; update log file every 500 ps
nstxout-compressed = 100000 ; save coordinates every 100.0 ps
; Bond parameters
continuation    = yes     ; continuing from NPT
constraint_algorithm = lincs ; holonomic constraints
constraints     = h-bonds ; bonds to H are constrained
lincs_iter      = 1       ; accuracy of LINCS
lincs_order     = 4       ; also related to accuracy
; Neighbor searching and vdW
cutoff-scheme   = Verlet
ns_type         = grid    ; search neighboring grid cells
nstlist         = 20      ; largely irrelevant with Verlet
rlist           = 1.2
vdwtype         = cutoff
vdw-modifier     = force-switch
rvdw-switch     = 1.0
rvdw            = 1.2     ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype     = PME     ; Particle Mesh Ewald for long-range electrostatics
rcoulomb        = 1.2
pme_order       = 4       ; cubic interpolation
fourierspacing  = 0.16   ; grid spacing for FFT
; Temperature coupling
tcoupl          = V-rescale ; modified Berendsen thermostat
tc-grps         = Protein_ATN Water_and_ions ; two coupling groups - more accurate
tau_t           = 0.1 0.1 ; time constant, in ps
ref_t           = 310 310 ; reference temperature, one for each group, in K
; Pressure coupling
pcoupl          = Parrinello-Rahman ; pressure coupling is on for NPT
pcoupltype      = isotropic ; uniform scaling of box vectors
tau_p           = 2.0     ; time constant, in ps
ref_p           = 1.0     ; reference pressure, in bar
compressibility  = 4.5e-5 ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc             = xyz     ; 3-D PBC
; Dispersion correction is not used for proteins with the C36 additive FF
DispCorr        = no
; Velocity generation
gen_vel         = no      ; continuing from NPT equilibration

```

### Arquivos de entrada Amostragem guarda chuva

❖ Parâmetros de dinâmica molecular- MD-pull.mdp

```

title      = Umbrella pulling simulation
define     = -DPOSRES_P
integrator = md
dt         = 0.002
tinit      = 0
nsteps     = 500000 ; 1 ns
nstcomm    = 10
; Output parameters
nstxout    = 5000 ; every 10 ps
; nstvout   = 5000

```



```

;nstfout   = 500
nstxtcout  = 500   ; every 1 ps
;nstenergy = 500
; Bond parameters
constraint_algorithm = lincs
constraints          = h-bonds
continuation        = yes   ; continuing from NPT
; Single-range cutoff scheme
cutoff-scheme = Verlet
nstlist          = 20
ns_type          = grid
rlist            = 1.4
rcoulomb         = 1.4
rvdw             = 1.4
; PME electrostatics parameters
coulombtype      = PME
fourierspacing   = 0.12
fourier_nx       = 0
fourier_ny       = 0
fourier_nz       = 0
pme_order        = 4
ewald_rtol       = 1e-5
optimize_fft     = yes
; Berendsen temperature coupling is on in two groups
Tcoupl          = V-rescale
tc_grps         = Protein Non-Protein
tau_t           = 0.1 0.1
ref_t           = 310 310
; Pressure coupling is on
Pcoupl          = Parrinello-Rahman
pcoupltype      = isotropic
tau_p           = 2.0
compressibility  = 4.5e-5
ref_p           = 1.0
refcoord_scaling = com
; Generate velocities is off
gen_vel         = no
; Periodic boundary conditions are on in all directions
pbc             = xyz
; Long-range dispersion correction
DispCorr        = EnerPres
; Pull code
pull            = yes
pull_ncoords     = 1   ; only one reaction coordinate
pull_ngroups     = 2   ; two groups defining one reaction coordinate
pull_group1_name = K36
pull_group2_name = Protein
pull_coord1_type  = umbrella ; harmonic potential
pull_coord1_geometry = distance ; simple distance increase
pull_coord1_dim   = N N Y
pull_coord1_groups = 1 2
pull_coord1_start = yes   ; define initial COM distance > 0
pull_coord1_rate  = 0.0075 ; 0.0075 nm per ps = 7.5 nm per ns
pull_coord1_k     = 1000  ; kJ mol-1 nm-2

```

❖ Parâmetros de dinâmica molecular - NPT- umbrella.mdp

```

title      = Umbrella pulling simulation
define     = -DPOSRES_P
; Run parameters
integrator = md
dt         = 0.002
tinit      = 0
nsteps     = 250000 ; 500 ps
nstcomm    = 10
; Output parameters
nstvout     = 25000 ; every 50 ps
nstfout     = 25000
nstxout-compressed = 25000
nstenergy   = 25000
; Bond parameters
constraint_algorithm = lincs
constraints   = h-bonds
continuation  = no
; Single-range cutoff scheme
cutoff-scheme = Verlet
nstlist      = 20
ns_type      = grid
rlist        = 1.4
rcoulomb     = 1.4
rvdw         = 1.4
; PME electrostatics parameters
coulombtype  = PME
fourierspacing = 0.12
fourier_nx   = 0
fourier_ny   = 0
fourier_nz   = 0
pme_order    = 4
ewald_rtol   = 1e-5
optimize_fft = yes
; Berendsen temperature coupling is on in two groups
Tcoupl       = Berendsen
tc_grps      = Protein Non-Protein
tau_t        = 0.5 0.5
ref_t        = 310 310
; Pressure coupling is on
Pcoupl       = Berendsen
pcoupltype   = isotropic
tau_p        = 1.0
compressibility = 4.5e-5
ref_p        = 1.0
refcoord_scaling = com
; Generate velocities is on
gen_vel      = yes
gen_temp     = 310
; Periodic boundary conditions are on in all directions
pbc         = xyz
; Long-range dispersion correction
DispCorr     = EnerPres
; Pull code
pull         = yes
pull_ncoords = 1 ; only one reaction coordinate
pull_ngroups = 2 ; two groups defining one reaction coordinate
pull_group1_name = ATN
pull_group2_name = Protein
pull_coord1_type = umbrella ; harmonic potential
pull_coord1_geometry = distance ; simple distance increase

```

```

pull_coord1_dim      = N N Y
pull_coord1_groups   = 1 2
pull_coord1_start    = yes      ; define initial COM distance > 0
pull_coord1_rate     = 0.0      ; restrain in place
pull_coord1_k        = 2000     ; kJ mol-1 nm-2

```

❖ Parâmetros de dinâmica molecular - md-umbrella.mdp

```

title      = Umbrella pulling simulation
define     = -DPOSRES_P
; Run parameters
integrator = md
dt         = 0.002
tinit      = 0
nsteps     = 500000 ; 1 ns
nstcomm    = 10
; Output parameters
nstxout-compressed = 500 ; every 1 ps
nstenergy   = 500
; Bond parameters
constraint_algorithm = lincs
constraints    = h-bonds
continuation   = yes
; Single-range cutoff scheme
cutoff-scheme = Verlet
nstlist       = 20
ns_type       = grid
rlist         = 1.4
rcoulomb      = 1.4
rvdw          = 1.4
; PME electrostatics parameters
coulombtype   = PME
fourierspacing = 0.12
fourier_nx    = 0
fourier_ny    = 0
fourier_nz    = 0
pme_order     = 4
ewald_rtol    = 1e-5
optimize_fft  = yes
; Berendsen temperature coupling is on in two groups
Tcoupl       = V-Rescale
tc_grps      = Protein Non-Protein
tau_t        = 0.1 0.1
ref_t        = 310 310
; Pressure coupling is on
Pcoupl       = Parrinello-Rahman
pcoupltype    = isotropic
tau_p        = 2.0
compressibility = 4.5e-5
ref_p        = 1.0
refcoord_scaling = com
; Generate velocities is off
gen_vel      = no
; Periodic boundary conditions are on in all directions
pbc         = xyz
; Long-range dispersion correction
DispCorr    = EnerPres
; Pull code
pull         = yes
pull_ncoords = 1 ; only one reaction coordinate
pull_ngroups = 2 ; two groups defining one reaction coordinate

```

```

pull_group1_name      = ATN
pull_group2_name      = Protein
pull_coord1_type      = umbrella ; harmonic potential
pull_coord1_geometry  = distance ; simple distance increase
pull_coord1_dim       = N N Y
pull_coord1_groups    = 1 2
pull_coord1_start     = yes      ; define initial COM distance > 0
pull_coord1_rate      = 0.0     ; restrain in place
pull_coord1_k         = 2000    ; kJ mol-1 nm-2

```

### Arquivos de Dinâmica molecular de aceleração randômica

❖ Parâmetros de dinâmica molecular - RAMD.mdp

```

define                =
integrator            = md
comm-mode             = Linear
nstcomm              = 100
comm_grps            = System
dt                   = 0.002
nsteps               = 15000000 ;(30ns)
nstxout              = 5000
nstvout              = 5000
nstlog               = 5000
nstenergy            = 5000
nstxtcout            = 5000
nstfout              = 5000
compressed-x-precision = 1000
xtc_grps             = SYSTEM
pbc                  = xyz
rlist                = 1.2
coulombtype          = PME
rcoulomb             = 1.2
cutoff-scheme        = Verlet
nstlist              = 20
fourierspacing       = 0.12
pme_order            = 4
ewald_geometry       = 3d
ewald-rtol           = 1e-5
ewald-rtol-lj        = 1e-5
optimize_fft         = yes
vdw-type             = Cut-off
vdw-modifier         = none
rvdw                 = 1.2
tcoupl               = Nose-Hoover
tc_grps              = Protein_K36 Water_and_ions
tau_t                = 1.0 1.0
ref_t                = 300 300
Pcoupl               = Parrinello-Rahman
pcoupltype           = isotropic
tau_p                = 2.0 ; 5 before
compressibility       = 4.5e-5
ref_p                = 1.0
refcoord_scaling     = com
gen_vel              = no
gen_seed             = -1
continuation         = yes
constraints           = h-bonds
constraint_algorithm  = lincs
lincs_order          = 4
lincs_iter           = 1

```

```

lincs_warnangle      = 60
DispCorr             = EnerPres
ramd                  = yes
ramd-seed             = 98XX
ramd-ngroups         = 1
ramd-group1-force     = 292.5 ; = 14kcal/mol A
ramd-eval-freq       = 50
ramd-group1-r-min-dist = 0.0025 ; nm
ramd-force-out-freq   = 10
ramd-group1-max-dist  = 4.0 ; nm
ramd-group1-receptor  = Protein
ramd-group1-receptor-pbc-atom = 569
ramd-group1-ligand    = K36
ramd-group1-ligand-pbc-atom = 0
ramd-pbc-ref-prev-step-com = yes
ramd-old-angle-dist   = no

```