

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

INSTITUTO DE QUÍMICA

Eduardo Gabriel Amaral de Oliveira

**INVESTIGAÇÃO DE VARIAÇÕES DE TRÊS FÁRMACOS COMERCIAIS COMO
CANDIDATOS A INIBIDORES DA PROTEÍNA SPIKE DO SARS-COV-2**

Uberlândia

2023

Eduardo Gabriel Amaral de Oliveira

**INVESTIGAÇÃO DE VARIAÇÕES DE TRÊS FÁRMACOS COMERCIAIS COMO
CANDIDATOS A INIBIDORES DA PROTEÍNA SPIKE DO SARS-COV-2**

Dissertação de Defesa de Mestrado apresentado ao Instituto de Química da Universidade Federal de Uberlândia como requisito para obtenção do título de mestre em Química.

Área de concentração: Físico-Química

Orientador: Prof. Dr. Guedmiller S. de Oliveira

Uberlândia

2023

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

O48 2023	<p>Oliveira, Eduardo Gabriel Amaral de, 1997- INVESTIGAÇÃO DE VARIAÇÕES DE TRÊS FÁRMACOS COMERCIAIS COMO CANDIDATOS A INIBIDORES DA PROTEÍNA SPIKE DO SARS- COV-2 [recurso eletrônico] / Eduardo Gabriel Amaral de Oliveira. - 2023.</p> <p>Orientador: Guedmiller Souza de Oliveira. Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Química. Modo de acesso: Internet. Disponível em: http://doi.org/10.14393/ufu.di.2023.476 Inclui bibliografia.</p> <p>1. Química. I. Oliveira, Guedmiller Souza de, 1980-, (Orient.). II. Universidade Federal de Uberlândia. Pós- graduação em Química. III. Título.</p> <p style="text-align: right;">CDU: 54</p>
-------------	---

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
 Coordenação do Programa de Pós-Graduação em Química
 Av. João Naves de Ávila, 2121, Bloco 5I - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
 Telefone: (34) 3239-4385 - www.cpgquimica.iq.ufu.br - cpgquimica@ufu.br



ATA

Programa de Pós-Graduação em:	Química				
Defesa de:	Dissertação de Mestrado Acadêmico, 378, PPGQUI				
Data:	Vinte e oito de julho de dois mil e vinte e três	Hora de início:	14:00	Hora de encerramento:	17:40
Matrícula do Discente:	12122QMI004				
Nome do Discente:	Eduardo Gabriel Amaral de Oliveira				
Título do Trabalho:	"Investigação de Variações de Três Fármacos Comerciais como Candidatos a Inibidores da Proteína Spike do SARS-CoV-2"				
Área de concentração:	Química				
Linha de pesquisa:	Química Teórica e Físico-Química de Macromoléculas e Coloides				
Projeto de Pesquisa de vinculação:	Uso de Modelagem Molecular no Desenvolvimento de Nanobiossensores para Detecção de Herbicidas e de Doenças Causadas pelo Mosquito Aedes Aegypt				

Reuniu-se, por webconferência, link <https://meet.google.com/tib-axjz-ywb>, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Química, assim composta: Professores Doutores: Eduardo de Faria Franca, e Amanda Danuello Pivatto ambos do IQUFU - Universidade Federal de Uberlândia; Kalil Bernardino, da Universidade Federal de São Carlos e Guedmiller Souza de Oliveira, orientador do candidato.

Iniciando os trabalhos o presidente da mesa, Dr. Guedmiller Souza de Oliveira, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.



Documento assinado eletronicamente por **Guedmiller Souza de Oliveira, Professor(a) do Magistério Superior**, em 28/07/2023, às 17:48, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Eduardo de Faria Franca, Professor(a) do Magistério Superior**, em 28/07/2023, às 17:48, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Amanda Danuello Pivatto, Professor(a) do Magistério Superior**, em 28/07/2023, às 17:48, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Kalil Bernardino, Usuário Externo**, em 28/07/2023, às 17:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4579409** e o código CRC **903762C1**.

AGRADECIMENTOS

Agradeço aos meus pais, Paulo Constantino de Oliveira e Maria Cristina Amaral Oliveira pelo apoio constante que sempre me deram ao longo de minha trajetória acadêmica, e também por tudo o que veio antes, proporcionando-me a oportunidade de chegar até aqui.

Agradeço ao Prof. Dr. Guemiller Souza de Oliveira, meu orientador, quem tornou possível a realização deste trabalho assim como pela constante atenção dedicada a este. Aos professores Antônio Eduardo da Hora Machado, Eduardo de Faria Franca e Guilherme Fernandes de Souza Miguel por toda a ajuda no decorrer do meu período de vínculo com o Programa de Pós-Graduação em Química.

Aos discentes Richard Gabriel Freitas Guedes, William Oliveira Soté e Lourival Rodrigues de Sousa Neto, com os quais tive contato por meio do Laboratório de Cristalografia e Química Computacional ou por disciplinas da pós-graduação, pela ajuda e amizade.

A Geovanna Soldi Mendes Correia, Jessica Santos Gomes, Jean Carlos Guimarães, Bruno Viotti de Melo, Alexandre dos Santos Mello, Victor Barcelos Gontijo de Carvalho e Rafael Gustavo Amaral de Oliveira pelo suporte constante, fosse presencial ou à distância.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), pelo financiamento recebido, permitindo a realização de atividades acadêmicas em prol do desenvolvimento científico.

O presente trabalho foi realizado com apoio da CAPES – Código de Financiamento 001.

RESUMO

O presente trabalho visou a utilização de técnicas de mecânica e dinâmica molecular, assim como cálculos de estrutura eletrônica para o estabelecimento de um protocolo de descoberta de fármacos, o qual foi testado para a pesquisa de 2053 fármacos análogos com possibilidade de inibirem a proteína *spike* do SARS-CoV-2. Foram gerados candidatos a fármacos com base em três fármacos existente: ácido baloxavir, oseltamivir e zanamivir e nas estruturas de 45 ligantes orgânicos a serem adicionados a posições específicas de cada um dos fármacos. A melhor estrutura com relação à energia de afinidade e localizada próxima do sítio ativo da proteína foi submetida à dinâmica molecular, em busca de otimizar as posições atômicas do sistema de partida. Um total de 100 ns de trajetórias foram obtidos para que, então, fosse determinado o quadro de maior estabilidade do fármaco de melhor performance amostrado pela técnica para que o mesmo fosse submetido a um cálculo mecânica quântica, nível TDFDT, em busca de informações mais precisas acerca da interação do fármaco com os resíduos de aminoácidos ao seu redor. O melhor candidato selecionado foi uma variante do ácido baloxavir na qual havia sido adicionado um átomo de bromo, porém, os resultados referentes aos estados excitados indicaram baixa probabilidade de movimento eletrônico no sistema avaliado. Tais resultados enfatizam a necessidade de se conduzir experimentos mais robustos no sentido de filtrar resultados preliminares provenientes de ancoragens moleculares, pois foi possível detectar a baixa chance de sucesso da espécie análoga mencionada mesmo após seu sucesso preliminar. Portanto, levando em conta também o custo computacional de todas as técnicas, o teste do protocolo mostra que é possível eliminar falsos positivos oriundos da ancoragem molecular por intermédio de técnicas mais consistentes como a dinâmica molecular.

Palavras-chave: ancoragem molecular, dinâmica molecular, teoria da densidade funcional.

ABSTRACT

The present work aimed to use molecular mechanics, dynamics and electronic structure calculation techniques in order to establish a drug discovery protocol, which was tested for the research of 2053 analog drugs with possibility of inhibiting SARS-CoV-2's spike protein. It was generated drug candidates based on three existing ones: baloxavir acid, oseltamivir and zanamivir and the structures of 45 organic ligands added to specific positions of each drug. The best docked structure in regard to affinity energy and proximity to the protein's active site, was submitted to molecular dynamics in order to optimize system's atomic positions. A total of 100 ns trajectory was obtained so it could be passed on to a quantum mechanics calculation, TDDFT level, in search of more precise informations regarding the drug interaction with the aminoacid residues nearby. The best candidate was a variant of baloxavir acid to which was added a bromine atom, but, the excited states results showed a little probability of electronic movement for the available system. Such a result emphasizes the need of conducting more robust experiments to filter preliminary results from molecular dockings, because it was possible to detect the low success chance of the cited mutant even after its preliminary success. Hence, taking into account the computational cost of all the techniques, the protocol test shows that it is possible to eliminate false positives from molecular dockings by the use of more consistent techniques such as molecular dynamics.

Keywords: molecular docking, molecular dynamics, density functional theory.

LISTA DE FIGURAS

Figura 1. Ciclo reprodutivo do SARS-CoV-2.....	2
Figura 2. Fórmulas estruturais dos fármacos base deste estudo: (a) Oseltamivir, (b) Zanamivir e (c) Baloxavir Marboxol.....	4
Figura 3. Exemplo de correção de simetria efetuada no mRMSD.....	11
Figura 4. Grupos substituintes empregados.....	15
Figura 5. Metodologia de drug discovery empregada.....	16
Figura 6. Proteína spike do SARS-CoV-2 e seu receptor ACE2.....	17
Figura 7. Ilustração das (a) ancoragem rígida e (b) ancoragem semiflexível.....	18
Figura 8. Explicação dos termos da Tabela 1.....	21
Figura 9. Fluxograma das etapas da DM.....	25
Figura 10. Exemplo de visualização de arquivo CUBE no programa VMD.....	27
Figura 11. Modificações estruturais do fármaco zanamivir.....	29
Figura 12. Modificações estruturais do fármaco oseltamivir.....	30
Figura 13. Modificações estruturais do fármaco ácido baloxavir.....	31
Figura 14. Estruturas dos fármacos com base na indexação dos carbonos substituídos: (a) oseltamivir, (b) zanamivir e (c) ácido baloxavir.....	32
Figura 15. Distribuições de probabilidade representando as energias de afinidade da conformação mais estável dos ligantes com relação ao receptor. As curvas descrevem o comportamento dos análogos agrupados quanto aos seus respectivos fármacos base: (a) ácido baloxavir, (b) oseltamivir e (c) zanamivir.....	33
Figura 16. Mapa de calor das substituições feitas com a energia de ligação da conformação de afinidade mais estável da ancoragem molecular do grupo B.....	35
Figura 17. Mapa de calor das substituições feitas com a energia de ligação da conformação de afinidade mais estável da ancoragem molecular do grupo O.....	35
Figura 18. Mapa de calor das substituições feitas com a energia de ligação da conformação de afinidade mais estável da ancoragem molecular do grupo Z.....	36
Figura 19. Interações entre o ligante B-31-2 e o receptor.....	37
Figura 20. Interações entre o ligante O-12-1 e o receptor.....	38
Figura 21. Interações entre o ligante Z-12-1 e o receptor.....	38
Figura 22. Mapa de calor dos análogos do grupo B com a energia de ligação da conformação cuja ancoragem molecular mais se aproximou de NAG.....	39

Figura 23. Mapa de calor dos análogos do grupo O com a energia de ligação da conformação cuja ancoragem molecular mais se aproximou de NAG.....	40
Figura 24. Mapa de calor dos análogos do grupo Z com a energia de ligação da conformação cuja ancoragem molecular mais se aproximou de NAG.....	40
Figura 25. Interações entre o ligante B-18-1 e o receptor.....	41
Figura 26. Ilustração da estrutura do fármaco selecionado e dos resultados obtidos para alguns dos parâmetros calculados.....	42
Figura 27. Superfície eletrônica calculada para o recorte.....	42

LISTA DE TABELAS

Tabela 1. Pesos e termos da função de ranqueamento.....	20
Tabela 2. Fármacos comerciais e seus respectivos SMILES.....	28
Tabela 3. Energias de transição para o sistema equilibrado a um raio de corte de 5 Å do ligante.....	43

LISTA DE ABREVIATURAS E SIGLAS

- ACD** – *Antimicrobial Chemotherapeutics*
- ADME** – Absorção, Distribuição, Metabolismo e Excreção
- ALH** – Aceptores de Ligação de Hidrogênio
- AM** – Ancoragem Molecular
- API** – *Application Programming Interface*
- AS** – Ancoragem Semiflexível
- ASPT** – Área Superficial Polar Topológica
- B** – Conjunto de Fármacos Análogos do Ácido Baloxavir
- BDGS** -- Broyden-Fletcher-Goldfarb-Shanno
- CEE** – Cálculos de Estruturas Eletrônica
- CHARMM** – Chemistry at Harvard Macromolecular Mechanics
- cm⁻¹** – Número de Onda
- CMC** – *Comprehensive Medicinal Chemistry*
- Da** – Daltons
- d_{ij}** – Superfície de Distância entre os Átomos i e j
- DLH** – Doadores de Ligação de Hidrogênio
- DM** – Dinâmica Molecular
- E** – Proteína Espicular do SARS-CoV-2
- E1** – Domínio 1 da Proteína Espicular do SARS-CoV-2
- E2** – Domínio 2 da Proteína Espicular do SARS-CoV-2
- ECA2** – Enzima Conversora de Angiotensina 2
- eV** – Elétron-volts
- F** – Força
- FDA** – *Food and Drug Administration*
- f_{os}** – Força do Oscilador
- fs** – Femtossegundos
- H** – Operador Hamiltoniano
- Inter** – Intermolecular
- Intra** – Intramolecular
- K** – Kelvin
- lb** – *Lower Bound*

LI – Linearmente Interpolado

LR – Ligações Rotacionáveis

MDDR – *MDL Drug Data Report*

mlogP – Logaritmo do Coeficiente de Partição em Octano/Água segundo o Algoritmo de Moriguchi

MM – Mecânica Molecular

mRMSD – Raiz Quadrada do Desvio Médio Padrão Minimizado

NA – Número de Átomos

NAG – 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

NAMD – *Nanoscale Molecular Dynamics*

NAn – Número de Anéis

NC – Número de Carbonos

N_{rot} – Número de Ligações Rotacionáveis

n_{inc} – Número de Incrementos

NH – Número de Heteroátomos

NPT – Número de Átomos, Pressão e Temperatura

ns – Nanossegundos

O – Conjunto de Fármacos Análogos do Oseltamivir

OMS – Organização Mundial de Saúde

PDB – *Protein Data Bank*

PDBQT – *Protein Data Bank, Partial Charge & Atom Type*

PME – *Particle Mesh Ewald*

P^N – Conjunto de Momentos

R – Raio de van der Waals

R_i – Coordenadas Espaciais do Átomo *i*

r_{ij} – Distância Interatômica

R^N – Conjunto de Coordenadas Espaciais

RBD – *Receptor-Binding Domain*

RE – Regra de Egan

RG – Regra de Ghose

RL – Regra de Lipinski

RM – Regra de Muegge

RMSD – Raiz Quadrada do Desvio Médio Padrão

RV – Regra de Veber

SMILES – *Simplified Molecular-Input Line-Entry System*

TDDFT – *Time Dependant Density Functional Theory*

TDF – Teoria da Densidade Funcional

TDFDT – Teoria da Densidade Funcional Dependente do Tempo

TIP3P – *Transferable Intermolecular Potencial with 3 Points*

U – Energia Potencial

ub – *Upper Bound*

VMD – Visual Molecular Dynamics

wlogP – Logaritmo do Coeficiente de Partição em Octano/Água segundo o Algoritmo de Wildman-Crippen

xlogP – Logaritmo do Coeficiente de Partição em Octano/Água segundo o Algoritmo do Programa XLOGP.

Å – Angstrom

Z – Conjunto de Fármacos Análogos do Zanamivir

$\theta_{i,j}$ – Ângulo entre os Átomos i e j

θ_{inc} – Incremento Angular

SUMÁRIO

1	Introdução	1
2	Objetivos	6
3	Fundamentação teórica	7
3.1	Mecânica molecular.....	7
3.2	Dinâmica molecular.....	8
3.3	Cálculo de estrutura eletrônica.....	9
3.4	Análise de resultados.....	10
3.4.1	Energia livre de ligação.....	10
3.4.2	Raíz Quadrada do Desvio quadrático médio.....	10
3.4.3	Raíz Quadrada do Desvio quadrático médio minimizado.....	11
3.5	Filtros de viabilidade oral.....	12
4	Metodologia	15
4.1	Candidatos a inibidores.....	16
4.2	Preparação dos arquivos.....	16
4.2.1	Preparação da proteína.....	16
4.2.2	Preparação dos ligantes.....	17
4.3	Ancoragem molecular.....	17
4.3.1	Ancoragem semiflexível.....	19
4.3.2	AutoDock Vina.....	19
4.3.2.1	Algoritmo de busca.....	19
4.3.2.2	Algoritmo de otimização.....	21
4.3.2.3	Resultados.....	22
4.4	Dinâmica molecular.....	22
4.4.1	Preparação dos arquivos de entrada.....	23
4.4.2	Nanoscale Molecular Dynamics.....	25
4.4.3	Campo de força.....	26
4.4.4	Análise dos resultados.....	26
4.5	Cálculo de estrutura eletrônica.....	26
5	Resultados e discussão	28
5.1	Fármacos análogos.....	28
5.2	Ancoragem molecular.....	32

5.3	Filtros de viabilidade oral.....	41
5.4	Cálculo de estrutura eletrônica.....	42
6	Conclusões.....	44
	Referências Bibliográficas.....	45

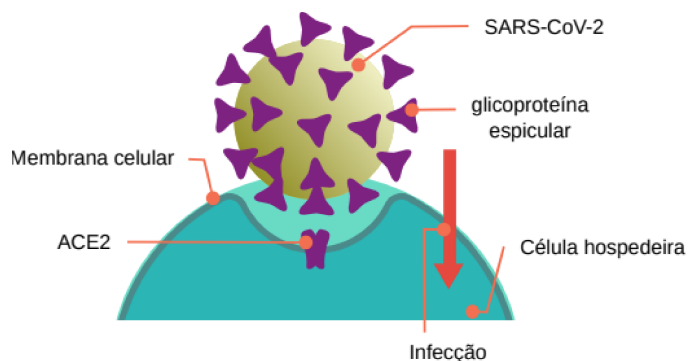
1 Introdução

Como é de conhecimento mundial, o mês de dezembro de 2019 foi marcado pela eclosão de uma doença a qual viria futuramente, em 11 de fevereiro do ano posterior, foi batizada pela Organização Mundial de Saúde (OMS) de COVID-19, causada pelo vírus Síndrome Respiratória Aguda Severa 2, SARS-CoV-2 (HE; DENG; LI, 2020; CARABELLI, 2023). A doença evoluiu para o estágio de pandemia em 11 de março de 2020, e até o dia 22/04/2023 já ultrapassava a marca de 750 milhões de infecções e 6,8 milhões de mortes, desde o seu aparecimento (CARVALHO; KRAMMER; IWASAKI, 2021; OMS, 2023). Ela é considerada um grande problema graças à sua fácil transmissão, visto que pode ser contraída via contato com plumas virais, as quais aparecem na casa dos milhares de gotículas por centímetros cúbicos em espirros e tosses secas de pessoas contaminadas, e têm a capacidade de driblar várias barreiras fisiológicas de defesa do corpo (JAYAWEERA et al, 2020).

O vírus que viria a ser denominado SARS-CoV-2 é pertencente à família Coronaviade e ao gênero Betacoronavirus. O mesmo apresenta semelhança genômica com outros membros do gênero os quais também acarretam doenças do trato respiratório humano, sendo eles o SARS-CoV, causador da Síndrome Respiratória Aguda Severa, com o qual compartilha 79% do genoma, e o MERS-CoV, causador da Síndrome Respiratória do Oriente Médio, com semelhança genética de 50% (WANG et al, 2021).

O processo de infecção começa quando a glicoproteína espicular (E), também chamada *spike*, a qual participa ativamente do ciclo reprodutivo do referido vírus, media a fusão da membrana celular viral com a da célula-alvo humana, a enzima conversora de angiotensina 2 (ECA2), como ilustra a Figura 1. Existem dois domínios que compõem a proteína E, mencionados aqui como E1 e E2. O primeiro é o responsável pela ligação com o receptor (ECA2), enquanto o segundo fica a cargo da fusão com a sua membrana celular (HE; DENG; LI, 2020; HU, 2021).

Figura 1. Ciclo reprodutivo do SARS-CoV-2.



Fonte: Autor, 2023.

A proteína *spike* desempenha um papel biológico essencial para o SARS-CoV-2, sendo um alvo terapêutico válido para estudos interativos com fármacos em busca daqueles que consigam provocar uma inibição eficaz da funcionalidade desempenhada pela macromolécula. No processo descrito, o sítio-ativo é quem desempenha o papel de reconhecer e provocar a ligação do *Receptor-Binding Domain* (RBD) da ECA2 (KALATHIYA, 2020).

Estudar ligantes que apresentem performance efetivamente contra os sítios-ativos de proteínas, regiões específicas onde ocorrem as interações moleculares e catalíticas responsáveis por suas funções biológicas, é uma etapa fundamental dos processos de reposicionamento e descoberta de fármacos (ZHAO, CAO, ZHANG, 2020; CHOUDHURY, MURUGAN, PRIYAKUMAR, 2022). É de conhecimento que as vacinas aplicadas não possuem eficácia vitalícia (DHAMELIYA; SHAH, 2021), enfatizando a necessidade por medicamentos eficazes no combate à moléstia.

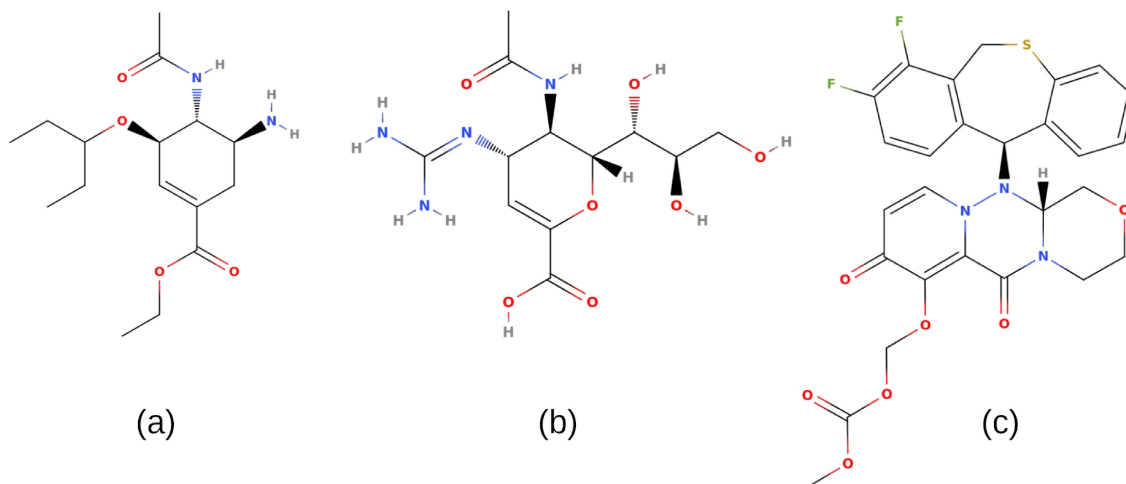
O reposicionamento ou reutilização de medicamentos, é uma abordagem que busca identificar novos alvos terapêuticos para medicamentos já existentes, sendo aproveitados aqueles que já passaram por testes de segurança e toxicidade em humanos. Isso oferece vantagens significativas no que diz respeito à quantidade de recursos (tempo e dinheiro), envolvido no processo de desenvolvimento de um novo fármaco, pois a velocidade da estratégia pode oferecer opções terapêuticas a doenças cuja cura ainda não foi descoberta, especialmente em casos urgentes como a COVID-19. Etapas iniciais do reposicionamento de medicamentos podem ser realizadas por meio de métodos computacionais, como a ancoragem molecular (PARVATHANENI, 2019; SINGH, 2020).

A descoberta de novos medicamentos se distingue do reposicionamento por envolver o desenvolvimento de compostos químicos, os quais se objetiva serem capazes de interagir de forma específica e eficaz contra alvos moleculares do patógeno. Esse processo abrange diversas etapas, desde a identificação de alvos terapêuticos potenciais até a otimização de compostos candidatos para alcançar a eficácia e a segurança desejadas. Embora existam as partes de triagem, síntese, ensaios biológicos e testes clínicos das moléculas candidatas, o processo pode ter o seu custo reduzido caso sejam empregadas técnicas computacionais em seus primórdios (GHOSH, 2021; U.S. FOOD AND DRUG ADMINISTRATION (FDA), 2018).

Nesse quesito, a disponibilidade de dados cristalográficos que descrevam sistemas contendo a proteína estudada são importantes para a validação de modelos computacionais, visto que dados obtidos experimentalmente fornecem informações precisas sobre a estrutura tridimensional de uma molécula ou material. Isso permite uma comparação com modelos computacionais previamente desenvolvidos, essencial para avaliar a qualidade e a precisão dos modelos (SLATER, KONTOYIANNI, 2019).

Este trabalho emprega a descoberta de novos fármacos. Para tanto, foram selecionados como base três fármacos disponíveis comercialmente (Figura 2): baloxavir marboxil (Xofluza), oseltamivir (Rimivat) e zanamivir (Relenza), todas ativas contra influenzas e as quais foram reposicionadas contra a proteína-alvo (BARDSLEY-ELLIOT, NOBLE, 1999; DUFRASNE, 2022; WAGHORN, 1998).

Figura 2. Fórmulas estruturais dos fármacos base deste estudo: (a) Oseltamivir, (b) Zanamivir e (c) Baloxavir Marboxol.



Fonte: Autor, 2023.

Em trabalhos puramente teóricos, no entanto, é necessário ter em mente o fato de que sistemas tão complexos quanto o corpo humano não podem ser perfeitamente representados pelos modelos atuais. As limitações computacionais contemporâneas não permitem que todas as possíveis interações dos fármacos estudados e as diversas proteínas em nossos corpos sejam computadas em rotinas eficientes (KINGHORN, 2022). Para tanto, existem artifícios computacionais que tentam simular a viabilidade das drogas em potencial, como algoritmos de aprendizado de máquina e regras, como a Regra dos Cinco, auxiliando em etapas pré-clínicas do estudo farmacológico. A Regra dos Cinco, ou Regra de Lipinski, é um princípio empregado na química medicinal para teorizar a possibilidade de um candidato a fármaco ser ministrado oralmente. De acordo com ela, o composto é mais propenso a ser ativo por via oral se atender alguns critérios: peso molecular < 500 Da, coeficiente de partição ($\log P$) < 5 , número de doadores de hidrogênio ≤ 5 e número de aceptores de hidrogênio ≤ 10 . Esses critérios foram baseados na observação empírica de que compostos essas propriedades, respeitando os referidos limites, são mais propensos a exibirem boa biodisponibilidade e permeabilidade, permitindo uma absorção e distribuição eficaz no organismo (JABEEN, RAN-GANATHAN, 2019; POLLASTRI, 2010).

Das técnicas computacionais responsáveis por descrever sistemas químicos, podem ser diferenciados dois tipos: aqueles baseadas em mecânica molecular (MM), os quais trabalham com sistemas em escala molecular (COUMAR, 2021; RENNER; SCHMIDT; DE LAS HERAS, 2021), e aqueles baseados em cálculos de estrutura eletrônica (CEE). Estes levam em conta a distribuição eletrônica de sistemas moleculares (KÜHNE, 2020).

A MM se baseia em cálculos feitos com os constituintes do sistema estando parados, sendo, portanto, independente do tempo. Essa característica a diferencia da dinâmica molecular (DM). As simulações de MM se baseiam na consideração de átomos como sendo partículas pontuais dotadas de cargas e ligações como sendo molas. (COUMAR, 2021). Essas técnicas são dependentes de campos de força que forneçam as funções de potencial para todos os átomos contidos no sistema (NERENBERG; HEAD-GORDON, 2018.), assim como a TDFDT, uma técnica CEE, é dependente de conjuntos de base que descrevem as formas dos orbitais atômicos por meio de funções matemáticas (CHAKRABORTY, CHATTARAJ, 2021).

2 Objetivos

Avaliar a interação de três fármacos comerciais modificados estruturalmente como candidatos à inibição da proteína *spike* do SARS-CoV-2, por intermédio de simulações e cálculos computacionais.

2.1 Objetivos específicos

- Estudar as interações de candidatos a fármaco e proteína por meio de anoragem molecular do tipo semiflexível.
- Utilizar os melhores resultados do tópico anterior como ponto de partida para simulações de dinâmica molecular.
- Efetuar cálculos referentes à distribuição de densidade eletrônica dos ligantes com relação aos resíduos de aminoácido mais próximos a ele.

3 Fundamentação Teórica

3.1 Mecânica molecular (MM)

A MM é um método computacional usado para simular e estudar o comportamento de moléculas a nível atômico. Ela se baseia em princípios da física clássica, como as leis do movimento de Newton e a lei de Hooke das molas, e fornece *insights* sobre as propriedades estruturais de moléculas ao calcular a energia potencial do sistema com base nas interações interatômicas. A função de energia potencial considera fatores como estiramento de ligações, flexão de ângulos, torção e interações não ligadas (forças de van der Waals e interações eletrostáticas).

$$V = \sum V_{\text{estiramento}} + \sum V_{\text{flexão}} + \sum V_{\text{torção}} + \sum V_{\text{vdW}} + \sum V_{\text{electrostática}} , \quad (1)$$

Na ancoragem molecular (AM), por exemplo, o método mencionado é empregado na otimização iterativa da geometria molecular com o fim de minimizar a energia potencial. Logo, as simulações de mecânica molecular podem prever conformações moleculares, analisar interações moleculares e fornecer informações sobre propriedades moleculares, como perfis de energia e estabilidade estrutural (AKSHAYA et al., 2023; MAHESWARI, MUTHU, SUNDIUS, 2023; MISRA, YADAV, 2022; TÜZÜN, SARIPINAR, 2020).

Na abordagem clássica, as moléculas são representadas como coleções de átomos conectados por ligações, e cada átomo é atribuído a um conjunto de parâmetros chamados campos de força os quais descrevem átomos como pontos dotados de massa os quais interagem por meio de interações não-ligantes e de valência. Esses parâmetros são tipicamente derivados de dados experimentais ou cálculos quânticos. Ao aplicar campos de força que contêm esses parâmetros, as simulações de mecânica molecular podem calcular as forças atuando em cada átomo e o movimento resultante das moléculas ao longo de passos iterativos. (VISSCHER, GEERKE, 2019; WANG et al., 2022).

A técnica encontrou amplas aplicações em várias áreas da química, bioquímica e ciência dos materiais. Naquela, contribuiu para a descoberta de medicamentos ao prever a afinidade de ligação de pequenas moléculas a proteínas-alvo, auxiliar na triagem virtual e no

design de novos candidatos a fármacos. Também desempenha um papel crucial na compreensão do comportamento de biomoléculas, como as proteínas, fornecendo informações sobre sua energia, mudanças conformacionais e interações com outras moléculas, como os ligantes terapêuticos (BUEHLER, 2022; DE PAULA JUNIOR et al., 2022).

Evidentemente, partindo de idealizações, dentre as quais algumas foram mencionadas, a MM possui limitações intrínsecas. Ela se baseia em simplificações e aproximações, assumindo que as moléculas se comportam de maneira clássica e negligenciando efeitos quânticos. Portanto, é incapaz de produzir resultados diretos acerca de como ligações químicas são feitas e desfeitas. Além disso, os campos de força usados na mecânica molecular possuem limitações inerentes devido à natureza empírica de seus parâmetros, levando à imprecisões em certos casos (HEHRE, 2003).

3.2 Dinâmica molecular

A dinâmica molecular é uma técnica computacional poderosa usada para investigar o comportamento dinâmico e a evolução temporal de sistemas moleculares. Essa abordagem simula o movimento das partículas constituintes, como átomos e moléculas, ao longo do tempo, com base nas forças interatômicas e nas leis clássicas do movimento. A dinâmica molecular permite estudar uma ampla gama de fenômenos, desde a flexão de moléculas individuais até a interação complexa de grandes sistemas biológicos, proporcionando informações detalhadas sobre a estrutura, a estabilidade, as propriedades termodinâmicas e os processos cinéticos das moléculas.

Na dinâmica molecular, as partículas são representadas por seus parâmetros de força, incluindo massas, cargas e potenciais de interação. Esses parâmetros são usados para calcular as forças intermoleculares, como ligações químicas, interações de van der Waals e interações eletrostáticas. Ao utilizar algoritmos numéricos, como o método de Verlet, as equações de movimento são integradas ao longo do tempo, permitindo que as partículas se movam de acordo com as forças aplicadas. A simulação da dinâmica molecular ocorre em pequenos incrementos de tempo, permitindo a observação do comportamento das partículas em uma escala temporal adequada (HANSSON, OOSTENBRINK, VAN GUNSTEREN, 2002).

No campo da químiainformática, ela pode desempenhar um papel crucial na previsão de propriedades termodinâmicas, como a energia livre de ligação. Assim como também é usada para estudar a relação de proteínas, ácidos nucleicos e membranas biológicas, fornecendo informações acerca da função biológica e da interação com ligantes (ALVES et al., 2018).

Uma das principais vantagens da dinâmica molecular é sua capacidade de simular sistemas em escala molecular com detalhes ocorrendo, usualmente, numa ordem que vai de picosegundos até microssegundos, possibilitando a observação de mudanças conformacionais e interações moleculares simuladas. A dinâmica molecular também pode fornecer informações sobre propriedades macroscópicas, como a viscosidade de um solvente ou as propriedades elásticas de um material (SEWELL et al. , 2003; ZHANG, OTANI, MAGINN, 2015).

Como limitação do método, devido ao alto custo computacional, a DM é frequentemente aplicada a sistemas relativamente pequenos e com limitações de tempo de simulação (SCHÜTTE, KLUS, HARTMANN, 2023), sendo eles equivalentes à ordem limite da casa dos milhares, usualmente, e a 100 ns na maioria dos estudos contemporâneos, sendo 1 ms já visto em alguns mais recentes.

3.3 Cálculo de estrutura eletrônica

A TFDDT é um método capaz de calcular as propriedades eletrônicas de sistemas moleculares excitados. Enquanto a TDF tradicional é amplamente utilizada para calcular as propriedades eletrônicas de sistemas no estado fundamental, a TFDDT estende essa abordagem para estudar as propriedades em estados eletronicamente excitados (CASIDA, HUIX-ROTLANT, 2012).

Na TFDDT, as propriedades eletrônicas dos sistemas excitados são calculadas a partir de uma equação de movimento acoplada que descreve a evolução temporal da densidade eletrônica. Essa equação é resolvida iterativamente, levando em conta a interação eletrônica com o potencial externo e a autoconsistência do potencial de Kohn-Sham. Ao incorporar a dependência temporal, é permitido o estudo das respostas eletrônicas a estímulos externos, como radiação eletromagnética. Além disso, pode ser usada para calcular espectros de absorção, emissão e ressonância Raman, fornecendo informações valiosas sobre a estrutura eletrônica e as transições de energia do sistema (BAERENDS et al., 2002).

Este tipo de cálculo tem sido aplicada em uma ampla variedade de sistemas, desde pequenas moléculas até sistemas bioquímicos. Ela desempenha um papel crucial na investigação de processos de transferência de carga em sistemas moleculares (MANDAL, MANNA, VENKATRAMANI, 2019).

3.4 Análise de resultados

3.4.1 Energia livre de ligação

Na ancoragem molecular, por exemplo, o conceito de energia livre desempenha um papel crucial na previsão da afinidade de ligação entre um receptor e um ligante. Ela representa a diferença de energia potencial entre os estados ligado e não ligado do complexo proteína-ligante. Nesse âmbito, é possível se fazer deduções acerca da estabilidade e da probabilidade de ligação. Desempenha, portanto, um papel fundamental na triagem virtual, abordagem computacional usada para identificar candidatos a medicamentos por meio da priorização e seleção de moléculas com maior chance de tratarem a enfermidade (CHOUDHARY, et al., 2020, YANG et al., 2022).

Nesse contexto, prever com precisão a energia livre de ligação continua sendo uma tarefa desafiadora devido à complexidade das interações receptor-ligante. Efeitos do solvente, flexibilidade e mudanças conformacionais durante a ligação adicionam complexidade significativa ao cálculo. Além disso, a AM busca conseguir, com precisão, as melhores conformações do ligante em relação à proteína (STANZIONE, GIANGRECO, COLE, 2021).

3.4.2 Raiz Quadrada do Desvio Quadrático Médio (RMSD)

O RMSD é uma métrica comumente usada para quantificar a diferença entre uma estrutura advinda de dados cristalográficos ou de referência e uma estrutura modelada ou prevista. Ele calcula a raiz quadrada da média dos quadrados das diferenças nas posições atômicas entre as duas estruturas, fornecendo uma medida de similaridade estrutural.

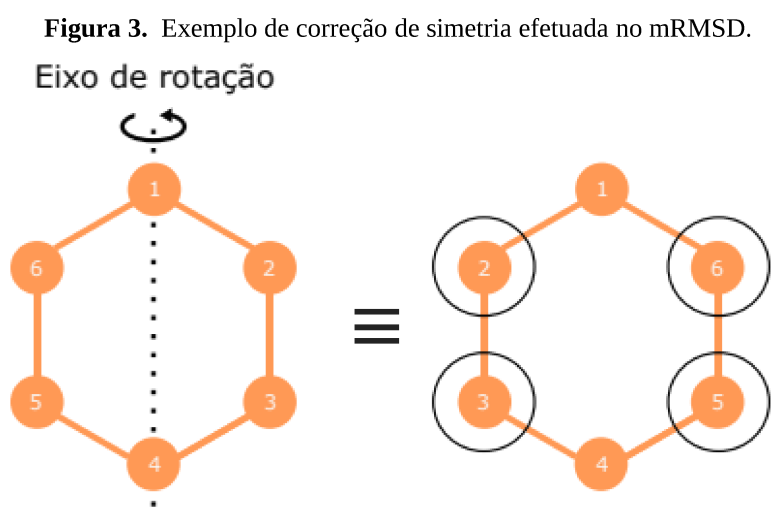
$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N d_i^2}, \quad (2)$$

onde N é a quantidade de átomos e d a diferença entre as suas respectivas posições em ambas as estruturas.

No caso de validação de docagens moleculares, é feita uma superposição das estruturas proteicas para efeitos de cálculo da diferença posicional dos átomos do ligante predito com relação ao nativo de dados experimentais (BELL, ZHANG, 2019). No caso de dinâmicas moleculares, pode-se avaliar o quanto as moléculas de interesse variam espacialmente ao longo do tempo (SARGSYAN, GRAUFFEL, LIM, 2017).

3.4.3 Raiz Quadrada do Desvio Quadrático Médio Minimizado (mRMSD)

O RMSD é uma métrica comumente empregada em análises de resultados de distintas conformações de um mesmo ligante, como é o caso nas saídas de docagens moleculares. No entanto, é fato que a simetria das moléculas não é considerada em cálculos comuns da referida métrica, aumentando, erroneamente, o seu valor. Os átomos são tratados de forma completamente discriminatória entre si. O mRMSD, nesse contexto, é uma técnica que corrige a falta de atribuição simétrica do RMSD por meio de rotações de quaternions polinomiais (MELI, BIGGIN, 2020).



3.5 Filtros de viabilidade oral

Uma das vias pelas quais é possível medicar o ser humano é por meio da ingestão oral, sendo a mais utilizada. Um fármaco empregado pelo sistema digestório irá, eventualmente, ser absorvido no trato intestinal, sendo a capacidade do fármaco de ser assimilado por essa rota um pré-requisito essencial para qualquer droga que vise tal forma de ministragem. A eficiência do transporte de substâncias pelas rotas dos tecidos epiteliais presentes no intestino dependem das suas propriedades físico-químicas, como tamanho, lipofilicidade, seu potencial de formar ligações de hidrogênio e também da sua afinidade com proteínas de transporte. A título de exemplo, Bohets et al. (2001) afirma, ainda, que compostos hidrofóbicos são transportados, principalmente, de forma passiva pelas membranas celulares de enterócitos, as células do intestino delgado responsáveis pela absorção de nutrientes. Os hidrofílicos, por sua vez, são transportados de forma passiva paracelular entre enterócitos. A fim de se identificar compostos com características de fármacos capazes de apontar candidatos viáveis de um ponto de vista relacionado à absorção, distribuição, metabolismo e excreção (ADME), houve um grande esforço por parte de pesquisadores em desenvolver filtros capazes distinguir compostos medicamentosos de não medicamentosos (MUEGGE; HEALD; BRITTELLI, 2001).

Em seu trabalho, Lipinski et al. (2001) estudou empiricamente propriedades físico-químicas de mais de 2000 compostos empregados como fármacos, todos os quais haviam entrado em fase clínica II sendo ministrados oralmente, a fim de determinar características comuns entre eles visando auxiliar no processo de estudo farmacológico. Ele e sua equipe analisaram parâmetros como massa molar (*MM*), coeficiente de partição (*logP*) octanol/água (o/a) e a quantidade de doadores (*DLH*) e aceptores de ligações de hidrogênio (*ALH*) que cada fármaco possuía. Esses quatro parâmetros passaram a compor o que ficou conhecida como Regra dos Cinco, Regra de Lipinski (RL) ou ainda Regra Pfizer, sendo que o *logP* empregado é o *mlogP*, produto do trabalho de Moriguchi et al (1992; 1994). A regra estabelece que qualquer fármaco o qual apresente mais de duas falhas quanto aos condicionais: $MM \leq 500$, $mlogP \leq 4,15$, $ALH \leq 10$ e $DLH \leq 5$ não é um bom candidato a fármaco quanto à via oral.

Entretanto, é fato hoje conhecido que a RL não é o bastante para que um candidato a fármaco seja descartado de forma efetiva, como provam os trabalhos de Egbert et al. (2019) no que diz respeito ao aumento de seletividade ligante-proteína em função do aumento da

massa do ligante, e de Zhang e Wilkinson (2007) quando eles mencionam que apenas 51% das drogas aprovadas pela Food and Drug Administration (FDA) na época passavam nos critérios estabelecidos pelo pesquisador. Para tanto, foi nomeado o termo Além da Regra dos Cinco, o qual serve para descrever alternativas a serem levadas em conta mesmo caso a regra dos cinco rejeite algum medicamento em potencial, como o emprego da técnica de MD (ABU-SALEH; YADAV; POIRIER, 2023; KRÄMER et al., 2016; L'EXACT et al., 2023).

Um esforço similar foi feito por Ghose, Viswanadhan e Wendoloski (1999), os quais estudaram de forma parecida a relação de drogas com a refratibilidade molar (RM), o número de átomos (NA), novamente a massa molecular e o $wlogP$, método de cálculo do coeficiente de partição oriundo do trabalho de Wildman e Crippen (1999). Esses parâmetros ficaram conhecidos como Regra de Ghose (RG), ou Filtro de Ghose, e quaisquer moléculas que possuam mais de uma falha quanto às condições: $160 \leq MM \leq 480$ Da; $-0,4 \leq wlogP \leq 5,6$; $40 \leq RM \leq 130$ e $20 \leq NA \leq 70$ são considerada uma falha do ponto de vista dessa regra. Isso significa que, tomando como base nas informações extraídas das amostras de fármacos estudadas pelos pesquisadores, compostos que violem as condições especificadas não são bons candidatos a fármacos orais. Entretanto, é necessário levar em consideração que as amostras, tanto as que deram origem aos parâmetros da RG quanto as progenitoras de qualquer outro filtro, não confizem com toda a população de fármacos existentes, portanto, um resultado que viole algum dos filtros não necessariamente determina que um fármaco é oralmente tóxico ou ineficaz. Os filtros de viabilidade servem para nortear a descoberta de fármacos com o objetivo de minimizar desperdícios de tempo e dinheiro investidos em aqueles com baixa chance de sucesso.

Alguns anos depois, Egan, Merz e Baldwin (2000) buscaram padrões estatísticos levando em conta compostos os quais apresentavam resultados mistos quanto à absorção passiva pelo intestino humano. Os autores pontuam que a ASPT e $wlogP$ sozinhos são parâmetros bons o bastante para atestar a viabilidade teórica de um fármaco no que diz respeito à sua capacidade de ser assimilado por via gástrica, enfatizando, por vez, que a MM é um parâmetro supérfluo a ser levado em conta como filtro, uma vez que o valor influencia nas medidas de ASPT e $wlogP$. A escolha destas é embasada pela física envolvida na permeabilidade membranar. Os valores permitidos de cada métrica são: $ASPT \leq 131,6$ Å e $wlogP \leq 5,88$, este último se aproximando bastante do valor previsto por Ghose et al. A regra discutida

é conhecida popularmente como Regra de Egan (RE).

Veber et al. (2002) concluiu, com base em um estudo *in vivo* envolvendo mais de 1100 fármacos, que é importante levar em conta a quantidade de ligações rotacionáveis, pertencentes a diedros próprios, presentes nas moléculas de estudo assim como baixa área superficial polar topológica (ASPT), as quais eram características comuns aos compostos empregados. A Regra de Veber (RV) estabelece que: $LR \leq 10$ e $ASPT \leq 140 \text{ \AA}$. Não existe uma quantidade de falhas aceitável descrita na publicação. O valor limite de ASPT se aproxima bastante do mensurado por Egan e colaboradores.

Um quinto filtro empregado na descoberta de fármacos é a Regra de Muegge (RM), ou Regra Bayer. Em seu estudo, Muegge, Heald e Brittelli (2001) discutem a importância de existirem filtros farmacofóricos independentes de redes neurais, justificando que estes métodos possuem a desvantagem de não possuírem o que chamaram de “sabedoria química”. A expressão é explicada logo em seguida pela menção do enviesamento à base de treinamento aos quais os filtros não estão submetidos. Seu filtro leva em conta a MM, ASPT, número de anéis (NAn), número de carbonos (NC), número de heteroátomos (NH), NR, ALH, DLH e o coeficiente de partição o/a $xlogP$, baseado no algoritmo de Cheng et al. (2007). Além de levar em conta um maior número de descritores moleculares, a abordagem de Muegge e colaboradores introduz alguns parâmetros desconsiderados pelos filtros anteriores. As condições estabelecidas pela RM são: $200 \leq MM \leq 600$; $ASTP \leq 150$; $NAn \leq 7$; $NC \leq 4$; $NH \leq 1$; $NR \leq 15$; $ALH \leq 10$; $DLH \leq 5$ e $2 \leq xlogP \leq 5$. Os valores de MM, ASTP, ALH e DLH são iguais ou próximos dos valores amostrados em outros filtros.

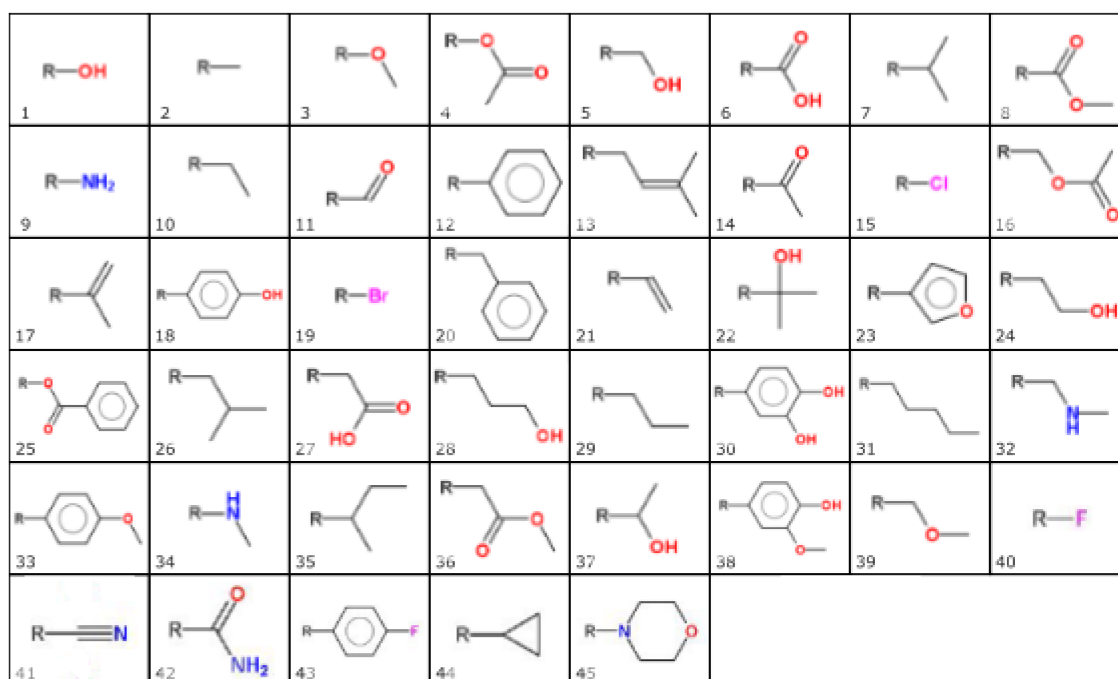
É mencionado no trabalho de Muegge e colaboradores, também, o fato de que o filtro desenvolvido conseguiu se adequar a 66-69% dos *subsets* de compostos provenientes à base de dados MDL Drug Data Report (MDDR), 61-68% dos *subsets* oriundos da base Comprehensive Medicinal Chemistry (CMC) e apenas 36% dos fármacos contidos em Antimicrobial Chemotherapeutics (ACD), o que não distoa do fato mencionado por Zhang e Wilkinson (2007).

4 Metodologia

A primeira etapa do procedimento, dado o fato de que a COVID-19 é uma doença ocasionada por um vírus, foi a determinação de fármacos disponíveis comercialmente e que apresentavam atividade comprovada contra outras moléstias também de origem viral, as quais exibiam semelhança sintomática com a mesma. Desta abordagem, foram selecionados os medicamentos Ácido Baloxavir, cujo precursor, o Baloxavir Marboxil, Oseltamivir, e Zanamivir.

A fim de averiguar quais tipos de mutações produziriam candidatos a fármacos mais eficientes na inibição da atividade protéica, foram preparadas as adições dos grupos substituintes contidos na Figura 4, selecionados a partir de levantamentos de grupos substituintes comumente empregados na química orgânica (ERTL, 2020; 2022).

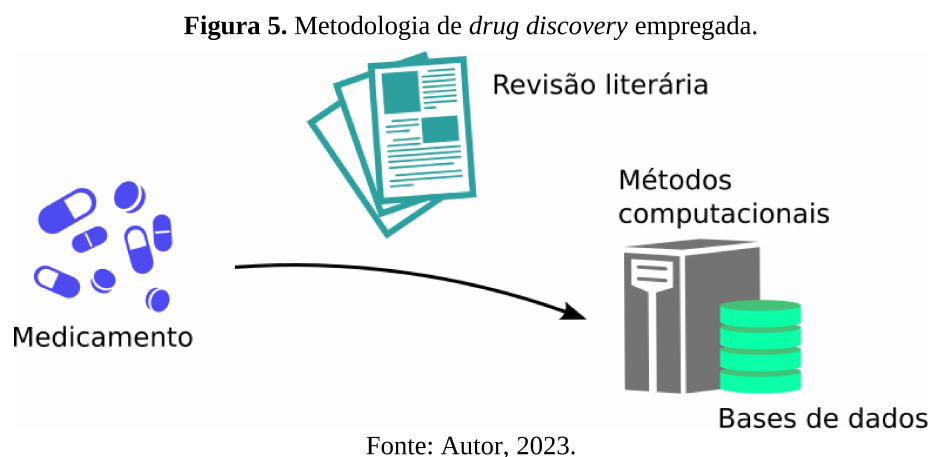
Figura 4. Grupos substituintes empregados.



Fonte: Adaptado de ERTL, 2020; 2022.

Foram adicionados os substituintes de forma a serem produzidos novos candidatos a fármacos por meio da troca de um hidrogênio de ligação C-H dentre todas as possíveis nas moléculas base, no intuito de averiguar o mais simetricamente o possível o espaço amostral de mutações moleculares. Foram gerados ambos os enantiômeros quando possível.

O procedimento sendo empregado, resumido na Figura 5, envolve a modelagem da estrutura de cada um dos candidatos (incluindo os fármacos em sua forma original), seguindo etapas de preparação para que sejam feitas as docagens moleculares de todos eles. As posições advindas das ancoragens, então, são usadas como ponto de partida para simulações de dinâmica molecular e quântica, respectivamente.



4.1 Candidatos a Inibidores

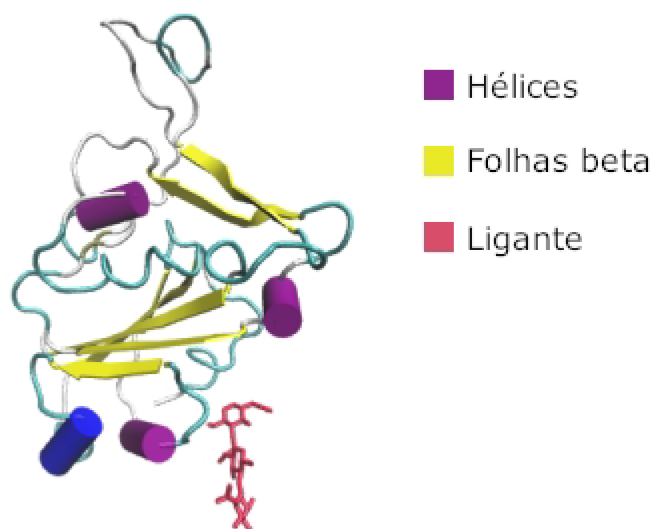
Os candidatos a inibidores foram todas as variantes dos fármacos base que puderam ser geradas pela substituição de átomos de hidrogênio ligados a carbono pelos grupos substituintes.

4.2 Preparação dos Arquivos

4.2.1 Preparação das Proteínas

O arquivo em formato PDB relativo à proteína foi extraído do *website* Protein Data Bank (BERMAN et. al, 2000), sendo ele a glicoproteína 6VW1, cadeia F (SHANG et al., 2020). A visualização 3D da estrutura as proteínas pode ser visualizada na Figura 6, a qual exibe o sistema ligante-receptor.

Figura 6. Proteína *spike* do SARS-CoV-2 e seu ligante cristalográfico NAG.



Fonte: Autor, 2023.

Os arquivos de todas elas tiveram quaisquer moléculas de água presentes removidas no software Visual Molecular Dynamics (VMD) (HUMPHREY; DALKE; SCHULTEN, 1996) antes de serem convertidas no formato PDBQT para o procedimento de ancoragem molecular, por meio do *software* Open Babel (O'BOYLE et al., 2017). A diferença entre ambos os tipos de arquivo é a presença de uma coluna de cargas atômicas no segundo, as quais são empregadas nos cálculos de potencial elétrico da ancoragem molecular.

4.2.2 Preparação dos Ligantes

Todos os fármacos modelados para posteriores estudos computacionais foram preparados na aplicação web MolView (BERGWERF, 2015), cuja interface é exemplificada na Figura 6, a qual permite que o arquivo de estrutura química das moléculas sejam salvos no formato MOL. Entretanto, tal formato não é compatível com a aplicação da metodologia de ancoragem, uma vez que não traz as cargas inerentes às moléculas. A conversão foi feita também por intermédio do Open Babel.

4.3 Ancoragem Molecular

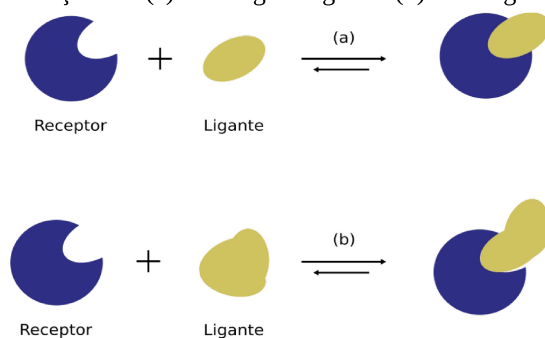
A técnica de ancoragem molecular busca, computacionalmente, encontrar os sítios de interação entre duas moléculas de forma detalhada, mostrando tanto aquelas de caráter ligante quanto não-ligante, e predizendo conformações e afinidades entre ligante e receptor em

ranque. Dessa forma, se permite que alvos biológicos sejam determinados visualmente. Sendo baseada nas estruturas dos componentes envolvidos, a existência de dados de alta resolução provenientes de cristalografias de raios-X e outras técnicas de análise é de fundamental importância permitiram o desenvolvimento do método (DE AZEVEDO, 2019; FAN; FU; ZHANG, 2019).

Se tratando de *drug design*, ou *drug discovery*, é um método de modelagem molecular de candidatos a fármacos. Dentre os métodos de ancoragem molecular, o modelo mais básico do MD, a ancoragem rígida, é uma técnica de caráter estático, e que, portanto, pode ser representada pela ideia dos modelos de chave-fechadura (MEDEIROS FILHO et al; 2020). Ao contrário de biomoléculas estáticas, como esse modelo descreve, as proteínas podem ser melhor descritas como entidades vivas, constantemente tendo as suas estruturas tridimensionais variando em função do tempo e outros fatores (ORELLANA, 2019). Uma vez que tanto o receptor quanto o ligante interagem um com o outro, mudanças estruturais podem ser consequências diretas desse fenômeno (TRIPATHI; BANKAITIS, 2017) e, até a data técnicas de ancoragem que permitem a flexibilidade do receptor (ancoragem flexível), são desafiadoras por demandarem um elevado custo computacional (DE AZEVEDO, 2019).

Levando em consideração as limitações de ambos os modelos mencionados, foi escolhido para este trabalho o modelo denominado ancoragem semiflexível. Ele interpreta o receptor como sendo uma entidade estática, e permite ao ligante graus de liberdade conformacionais que provêm a possibilidade de o mesmo se adequar mais precisamente aos sítios contidos na proteína.

Figura 7. Ilustração das (a) ancoragem rígido e (b) ancoragem semiflexível.



Fonte: Autor, 2023.

Para tanto, a ancoragem molecular foi empregado neste trabalho unicamente como ponto de partida para posteriores simulações de dinâmica molecular, as quais averiguam a estabilidade do ligante na posição encontrada pela ancoragem (GAJJAR; DHAMELIYA; SHAH, 2021), a qual permite que tanto a proteína quanto o ligante se adaptem flexivelmente às influências mutuamente existentes entre eles.

4.3.1 Ancoragem Semiflexível (AS)

O AS acontece partindo do pressuposto de que a proteína já está em sua conformação ideal para a interação com o ligante. Este é o único que possui graus de liberdade móveis. A metodologia em questão emprega uma ferramenta a qual utiliza de um algoritmo de busca do tipo incremental, seguindo o padrão:

$$\theta_{inc} = \prod_{i=1}^N \prod_{j=1}^{n_{inc}} \frac{360^\circ}{\theta_{i,j}} , \quad (3)$$

na qual N é o número de ligações rotacionáveis, n_{inc} o número de incrementos e $\theta_{i,j}$ o incremento angular j a ser feito na i -ésima ligação rotacionável (COUMAR, 2021).

4.3.2 AutoDock Vina

O *software* empregado neste trabalho para a realização das docagens é o AutoDock Vina (Vina) (TROTT; OLSON, 2010; EBERHARDT et al, 2021), o qual utiliza de um algoritmo genérico de busca e um método semi-empírico de avaliação.

4.3.2.1 Algoritmo de Busca

O algoritmo de busca, segundo os autores, está mais próximo do *machine learning* do que de uma orientação à física propriamente dita. Seu funcionamento é descrito abaixo:

$$c = \sum_{i < j} f_{ti tj}(r_{ij}) = c_{inter} + c_{intra} , \quad (4)$$

tal que o somatório (c) é feito sobre todos os pares de átomos que podem se mover de forma relativa um ao outro, separados por três ligações covalentes consecutivas. Cada átomo i tem

atribuído um tipo t_i para o cálculo de um conjunto simétrico de funções de interação f com base na distância interatômica (r_{ij}). O termo c pode ser interpretado como a soma das componentes inter e intramoleculares (c_{inter}, c_{intra}).

A energia de afinidade (energia livre de ligação) (s_i) é calculada a partir da componente intramolecular da conformação a qual apresentou o menor ranque (c_{intra1}):

$$s_i = g(c_i - c_{intra1}) , \quad (5)$$

onde g representa a função de ranqueamento.

Para essa função, átomos de hidrogênio são inexistentes. O raio de van der Waals dos átomos (R) é empregado na determinação da superfície de distância (d_{ij}), um componente necessário para a obtenção da função f . É abordado um valor de raio de corte equivalente a 8Å.

$$d_{ij} = r_{ij} - R_{ti} - R_{tj} , \quad (6)$$

$$f_{t_i t_j}(r_{ij}) \equiv h_{t_i t_j}(d_{ij}) , \quad (7)$$

tal que h representa a soma ponderada das interações estéreas encontradas, com base na Tabela 1 abaixo, cujos termos são descritos na Figura 8. O método traz inspiração do método *X-score* (WANG; LAI; WANG, 2002).

Tabela 1. Pesos e termos da função de ranqueamento.

Termo	Peso
Gauss₁	-0,0356
Gauss₂	-0,00516
Repulsão	0,840
Hidrofóbico	-0,0351
Ligação de hidrogênio	-0,587
N_{rot}	0,0585

Nota. Valores retirados do trabalho de Trott e Olson (2010).

Fonte: Autor, 2023.

O N_{rot} restante representa o número de ligações passíveis de rotação dentre os átomos mais pesados que hidrogênio do ligante. Ele é empregado na equação (6) abaixo, a qual define a função g independente de conformação.

$$g(c_{inter}) = \frac{c_{inter}}{1 + wN_{rot}}, \quad (8)$$

onde w simboliza o peso associado (TROTT; OLSON, 2010).

Figura 8. Explicação dos termos da Tabela 1.

$$\begin{aligned} gauss_1(d) &= e^{-2(d/0,5 \text{ \AA})} \\ gauss_2(d) &= e^{-2((d-3 \text{ \AA})/2 \text{ \AA})} \\ repuls\tilde{a}o(d) &= \begin{cases} d^2, & \text{se } d < 0 \\ 0, & \text{se } d \geq 0 \end{cases} \\ hidrof\tilde{o}bico(d) &= \begin{cases} 1, & \text{se } d < 0,5 \text{ \AA} \\ 0, & \text{se } d \geq 1,5 \text{ \AA} \\ LI, & \text{se } 0,5 \leq d < 1,5 \text{ \AA} \end{cases} \\ liga\tilde{c}\tilde{a}o \text{ de hidrog\tilde{e}nio}(d) &= \begin{cases} 1, & \text{se } d \leq 0,7 \text{ \AA} \\ 0, & \text{se } d \geq 0 \\ LI, & \text{se } 0 < d < 0,7 \end{cases} \end{aligned}$$

Nota. LI = lineamente interpolado.

Fonte: Autor, 2023.

4.3.2.2 Algoritmo de Otimização

O método *quasi-Newton* Broyden-Fletcher-Goldfarb-Shanno (BFGS) foi escolhido para a efetivação da busca local iterável da busca global. Além de levar em conta a função de ranqueamento, ele aborda o seu gradiente; o vetor que carrega as derivadas da função de ranqueamento com relação às posições e orientações do ligante, tais quais os valores das torções das ligações rotacionáveis dele e de quaisquer resíduos flexíveis existentes. Mesmo que a avaliação do gradiente venha a demorar mais do que a análise das conformações em si, o ganho em tempo de otimização faz com que o custo global do processo seja menor. (TROTT; OLSON, 2010).

4.3.2.3 Resultados

A saída padrão do Vina é constituída, basicamente, do resultado da ancoragem molecular, o qual vem em formato de tabela contendo o índice da pose avaliada, sua afinidade em kcal/mol e o valor dos RMSDs dos tipos *lower bound* (lb) e *upper bound* (ub). A pose que apresenta o melhor resultado de afinidade é tomada como referência para o cálculo dos RMSDs das demais, e por isso é sempre atribuída o valor 0 para os seus próprios. Segundo o manual do programa, o primeiro tipo (lb) é calculado se pegando cada átomo da molécula de referência e buscando o átomo de mesmo tipo mais próximo na pose alvo, sendo medido o RMSD, e também o oposto: partindo-se da molécula alvo, é calculado o RMSD com relação ao átomo de mesmo tipo mais próximo da molécula de referência, sendo mantido o maior valor de desvio quadrático encontrado. O segundo tipo (ub), é determinado pela comparação das posições dos mesmos átomos em ambas as poses.

4.4 Dinâmica Molecular

O método dinâmica molecular é uma consistente forma computacional de averiguar interações e movimentos relacionados às moléculas constituintes dos sistemas em escala atomística. Isso permite o reconhecimento de propriedades das propriedades macroscópicas com a finalidade de se comparar os resultados das simulações com aqueles obtidos empiricamente.

A técnica se baseia nas leis de Newton da mecânica clássica para serem obtidos os movimentos atômicos e as propriedades dinâmicas dos componentes do sistema, sendo os seus resultados extremamente acurados graças ao pesado embasamento físico do método. Segundo os autores Zhou e Liu (2022), a dinâmica molecular consegue gerar saídas que costumam estar alinhadas com os aquelas oriundas da mecânica quântica, mas com o diferencial de serem métodos computacionalmente mais baratos quando empregados de forma pura.

O emprego da mecânica é visível pela definição do operador Hamiltoniano energia to-

tal (eq. 9).

$$H(R^N, P^N) = \frac{\sum_i^N \sum_a P_{i,a}^2}{2m_i} + U(R^N) , \quad (9)$$

no qual R^N representa o conjunto das coordenadas espaciais e P^N o conjunto dos momentos de todas as partículas constituintes do sistema. Os somatórios fazem jus às componentes cinética e U à potencial da energia, sendo α aquele quem carrega os valores relativos às coordenadas x , y e z dos átomos e m_i a sua massa.

As forças atuantes sobre o sistema são equivalentes aos vetores gradientes responsáveis por fazerem com que cada um dos átomos do sistema seja movido para a posição que minimize a energia do mesmo, como descreve a eq. 10:

$$F_i(R^N) = \frac{-\partial U(R^N)}{\partial R_i} , \quad (10)$$

Em comunhão com a Segunda Lei de Newton:

$$m_i a_i = F_i(R^N) , \quad (11)$$

A eq. 8, se integrada com relação ao tempo, permite o cálculo das trajetórias dos átomos deterministicamente (ZHOU; LIU, 2022).

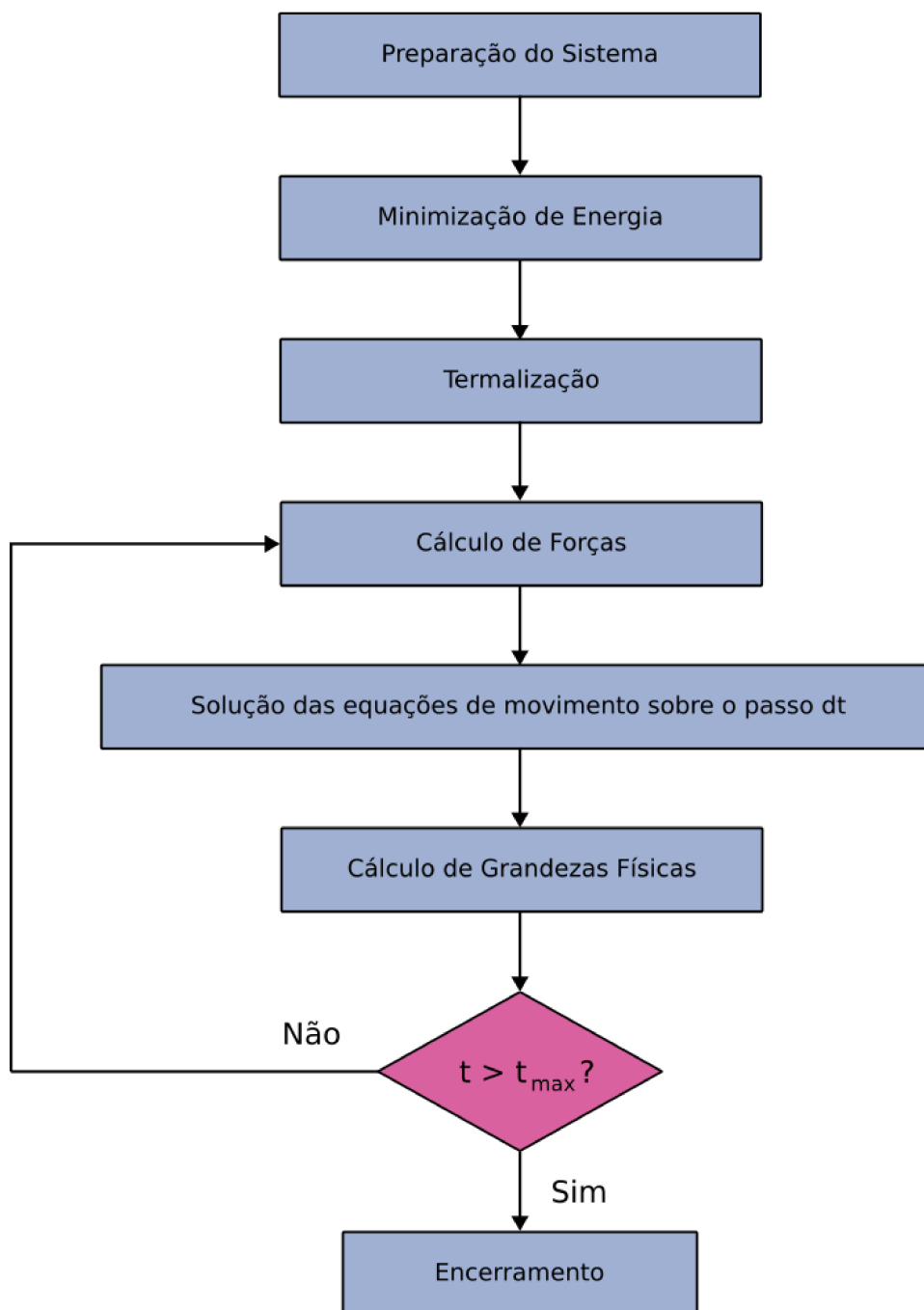
4.4.1 Preparação dos arquivos de entrada

O fluxograma do procedimento de dinâmica molecular sendo empregado neste trabalho é descrito na Figura 9. Ele começou com a preparação do sistema, passo qual consiste da conversão do arquivo PDBQT da saída da AM em PDB, o qual foi hidrogenado de forma programática pela API Python do *software* Pymol, gerando um novo arquivo PDB contendo todos os átomos da molécula. Esse novo PDB foi, então, convertido em MOL2 para que a molécula pudesse ser submetida ao SwissParam (ZOETE et al., 2011), serviço o qual provém

a parametrização completa da molécula. A estrutura originada da ancoragem molecular, então foi otimizada pelo *software* Orca via Teoria da Densidade Funcional (TDF), o qual também computa as cargas na mesma execução, sendo utilizadas cargas de Mulliken. Foram empregados o funcional m062x juntamente ao conjunto de base 6-31g(d, p). Com os valores de cargas provenientes dele, a coluna respectiva do PSF gerado pelo SwissParam foi atualizado com as novas medidas, momento a partir do qual o ligante estava pronto para o passo de combinação dos arquivos do ligante e da proteína. Quanto a esta, o PSF foi gerado via Visual Molecular Dynamics (VMD) (HUMPHREY; DALKE; SCHULTEN, 1996) *scripting*. Com os arquivos referentes ao sistema, foram então adicionadas moléculas do solvente (água TIP3P) a uma caixa cujo tamanho era composto pela extensão da proteína em todas as dimensões irradiando do centro acrescida de 15 Å. Finalmente, o pacote *ionize* foi empregado a fim de neutralizar as cargas do sistema com adição de íons de K⁺ e Cl⁻ para neutralizar as cargas existentes no sistema solvatado.

Com o sistema devidamente solvatado e neutralizado, era conduzida a etapa de minimização das energias. Isso era feito por meio de um ensemble NPT, mantendo-se constantes a temperatura e a pressão por meio de Pistões de Langevin. A temperatura era de 0 K. Os cálculos eram conduzidos a passos de 2 fs com *cutoff* de 12 Å para as interações eletrostáticas. Às moléculas de água eram atribuídas ligações rígidas.

Seguindo a minimização, era conduzida a termalização do sistema, a qual elevava a temperatura a 298 K a incrementos de 0,01 K/passos, seguindo os mesmos parâmetros mencionados anteriormente mais a adição do *Particle Mesh Ewald* (PME) a uma distância entre os *gridpoints* de 1 Å. A subsequente equilibração do sistema termalizado era conduzida com parâmetros idênticos a 100 ns de simulação.

Figura 9. Fluxograma das etapas da DM.

Fonte: Autor, 2023.

4.4.2 Nanoscale Molecular Dynamics (NAMD)

O *software* no qual as simulações de DM foram rodadas foi o NAMD, o qual foi criado com o intuito de ser um programa de alta performance capaz de lidar tanto com arquiteturas de CPU quanto GPU, sendo escrito em C++ (PHILLIPS JAMES et al, 2020). O campo de força empregado foi o CHARMM.

4.4.3 Campo de Força

Neste estudo, foi utilizado o modelo do campo de força Chemistry at Harvard Macromolecular Mechanics (CHARMM) para todos os átomos (VANOMMESLAEGHE et al, 2010; VANOMMESLAEGHE; MACKERELL, 2012; VANOMMESLAEGHE; RAMAN; MACKERELL, 2012; YU et al, 2012; GUTIÉRREZ et al, 2016), com adição dos dados referentes a proteínas (BEST et al, 2012) e lipídios (KLAUDA et al, 2010).

4.3.4 Análise dos resultados

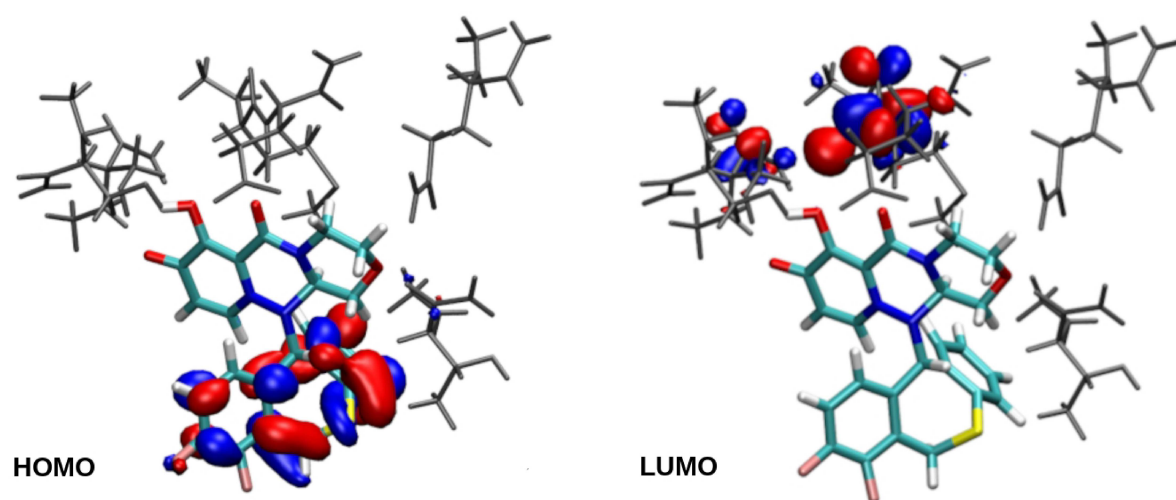
Como análise das trajetórias de dinâmica molecular foram calculados a quantidade de ligações de hidrogênio a fim de se encontrar a conformação mais estável do complexo ligante-receptor. Para tanto, foi utilizado o *plugin* hbonds faz a contagem de todas as ligações de hidrogênio presentes em cada *frame* da trajetória (GUMBART; LUO, 2011).

4.4 Cálculo de Estrutura Eletrônica

Da saída das simulações de dinâmica molecular foram escolhidos um *frame* por trajetória, que possua número de ligações de hidrogênio igual ao número máximo encontrado nela, e então será feita a seleção, no VMD, do ligante e de todos os aminoácidos da proteína que apresentem quaisquer átomo presente num raio de até 5 Å dele. O arquivo PDB salvo no *software* mencionado será convertido em formato XYZ para a posterior produção do arquivo de *input* para o Orca.

A visualização dos orbitais moleculares (Figura 10) dos sistemas que passaram pela Teoria da Densidade Funcional Dependente do Tempo (TDFDT) executada foi feita no *software* VMD, o qual consegue interpretar arquivos no formato CUBE. Tal formato é obtido a partir do arquivo GBW da saída do Orca, por meio do programa *orca_plot*.

Figura 10. Exemplo de visualização de arquivo CUBE no programa VMD.



Fonte: Autor, 2023.

5 Resultados

5.1 Fármacos Análogos

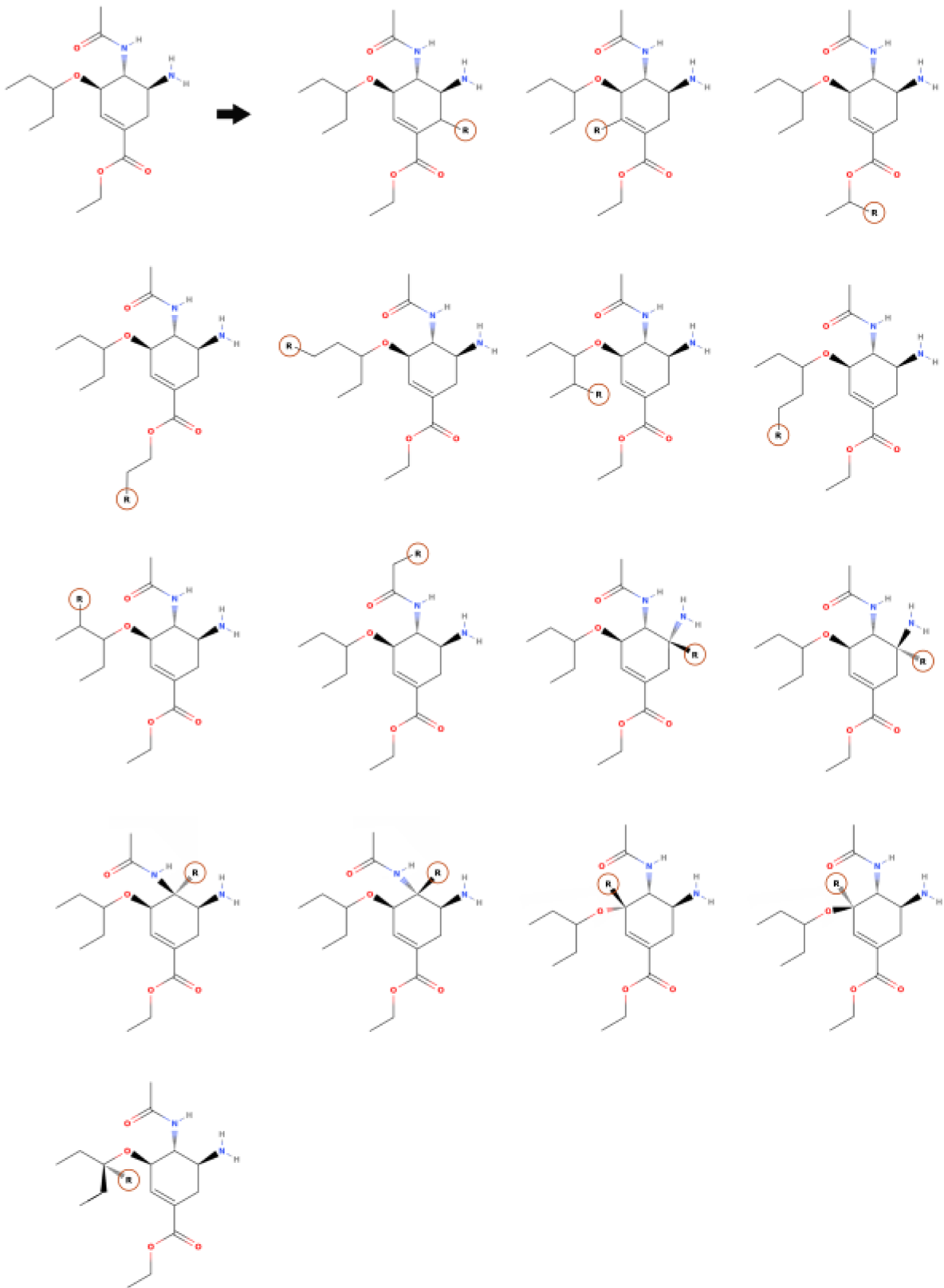
Foram gerados 2053 fármacos análogos segundo as possibilidades descritas nas Figuras 11, 12 e 13, as quais exibem as estruturas dos fármacos base acrescidas dos grupos substituintes onde as antes se encontravam ligações C–H, com enantiômeros referentes a carbonos assimétricos preexistentes tendo sido gerados quando cabíveis. As moléculas foram salvas no formato PDB, e convertidas em PDBQT via OpenBabel para subsequentes docagens moleculares.

Para efeitos de análise, as moléculas mencionadas foram divididas em três grupos: (I) com relação ao fármaco no qual sua estrutura foi baseada, sendo os grupos (B) derivados do ácido baloxavir, (O) derivados do oseltamivir e (Z) derivados do zanamivir; (II) com relação ao grupo substituinte que foi inserido, dos quais derivam os grupos R1 a R46; e (III) com relação ao átomo de carbono o qual teve a substituição feita em um de seus hidrogênios assim como sua isomeria (R ou S) caso a possuísse, indo de M1 a M16 para os grupos B e O e até M13 para o grupo Z. É válido lembrar os índices são individuais para cada um dos grupos B, O e Z, e também que o índice dos referidos átomos de carbono não segue os padrões da IUPAC. A Tabela 2 mostra os SMILES isoméricos das referidas moléculas, obtidos da base de dados do Molview, assim como a Figura 14 mostra, na estrutura de cada fármaco, o índice das referidas mutações segundo a mencionada baseada na posição SMILES de cada átomo.

Tabela 2. Fármacos comerciais e seus respectivos SMILES.

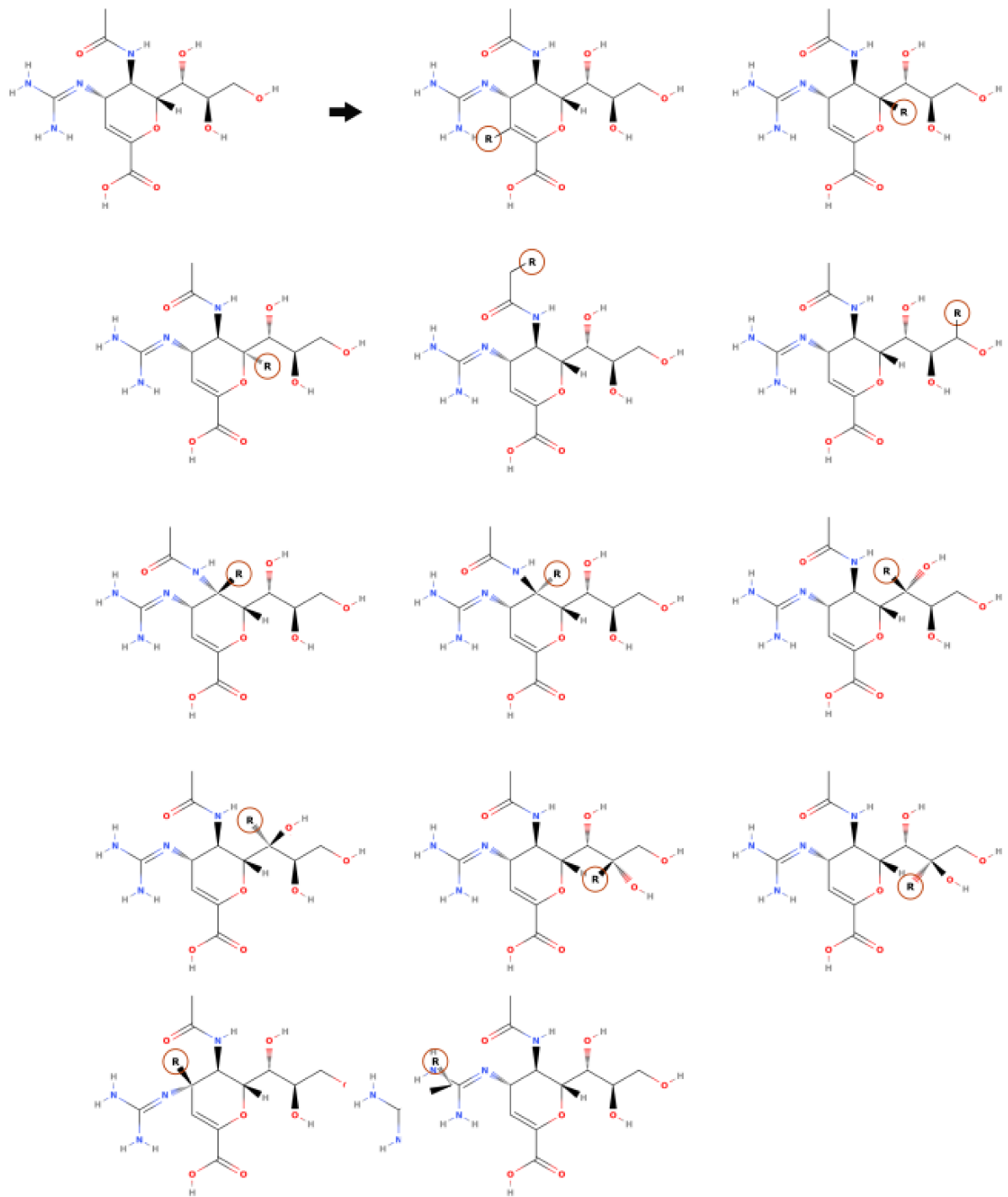
Fármaco	SMILES isoméricos
Ácido baloxavir	<chem>C1COC[C@@H]2N1C(=O)C3=C(C(=O)C=CN3N2[C@H]4C5=C(CSC6=CC=CC=C46)C(=C(C=C5)F)F)O¹</chem>
Oseltamivir	<chem>CCC(CC)O[C@@H]1C=C(C[C@@H])([C@H]1NC(=O)C)N)C(=O)OCC¹</chem>
Zanamivir	<chem>CC(=O)N[C@@H]1[C@H](C=C(O[C@H]1[C@@H])([C@@H](CO)O)O)C(=O)O)N=C(N)N¹</chem>

Notas. Dados retirados dos cartões de informação das respectivas moléculas do website Molview.Org.
Fonte: Autor, 2023.

Figura 11. Modificações estruturais do fármaco zanamivir.

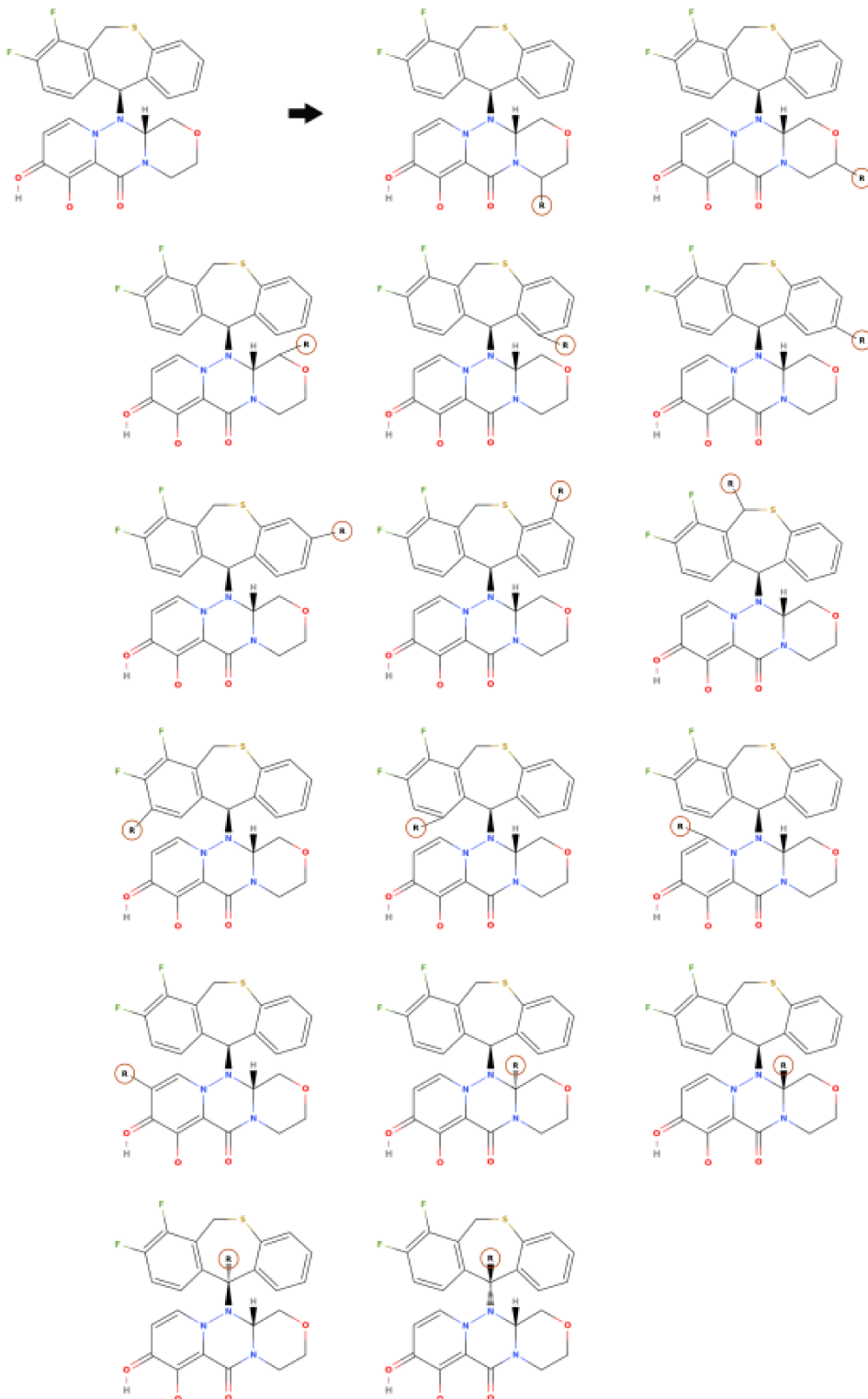
Fonte: Autor, 2023.

Figura 12. Modificações estruturais do fármaco oseltamivir.



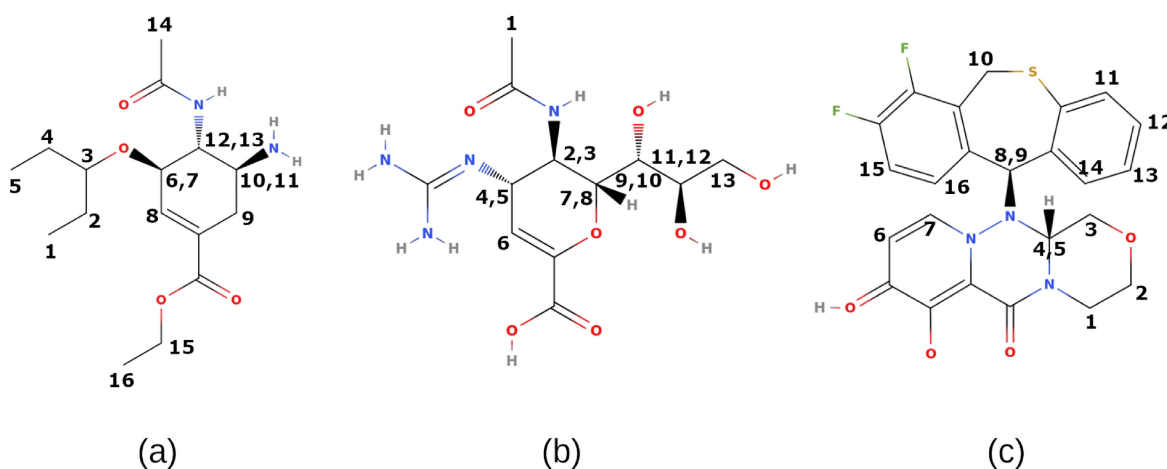
Fonte: Autor, 2023.

Figura 13. Modificações estruturais do fármaco ácido baloxavir.



Fonte: Autor, 2023.

Figura 14. Estruturas dos fármacos com base na indexação dos carbonos substituídos: (a) oseltamivir, (b) zanamivir e (c) ácido baloxavir.



Fonte: Autor, 2023.

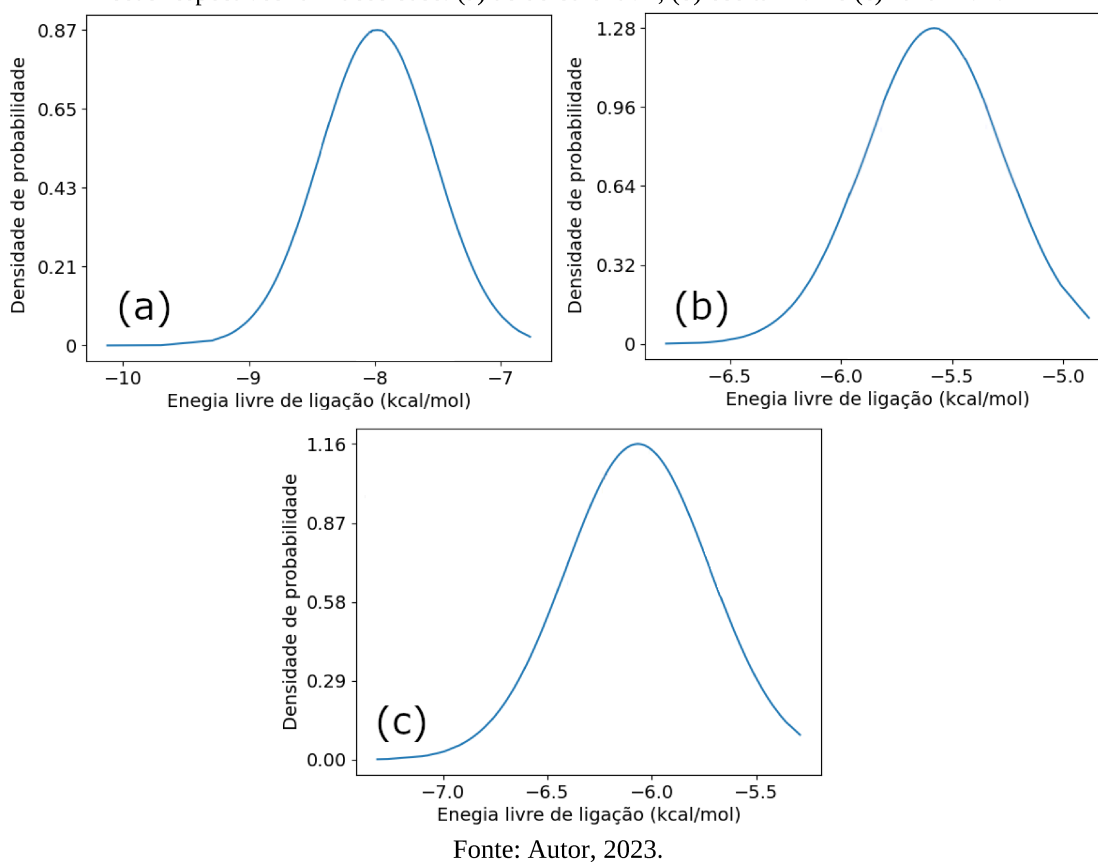
5.2 Ancoragem Molecular

A ancoragem molecular dos fármacos foi conduzida em triplicatas realizadas pelo AutoDock Vina em um computador com 24 núcleos de CPU. O *software* foi incapaz de empregar todos para o valor de exaustividade igual a 8, sendo o fato mencionado em um aviso do próprio programa durante a sua inicialização. Uma vez que apenas 8 núcleos eram alocados automaticamente de cada vez, foram iniciadas baterias de AMs de cada grupo (B: derivados do ácido baloxavir, O: derivados do oseltamivir e Z: derivados do zanamivir) para rodarem paralelamente, aproveitando, assim, todo o recurso disponível na referida máquina. Foi observado um tempo de execução de aproximadamente 16 horas de *wallclock time* para que as ancoragens de todas as 2053 moléculas submetidas fossem concluídas.

A Figura 15 abaixo revela a performance relativa dos análogos se agrupados com relação ao fármaco do qual foi derivado. Em todos os casos, é notável a assimetria das funções densidade de probabilidade, desviadas para a direita em função da frequência de ocorrência dos valores de maior energia do que a média. As medidas de tendência central dos compostos baseados no baloxavir e no oseltamivir estão similarmente distribuídas, com um valor de mediana próximo à média, à sua direita, indicando que existe uma maior frequência de valores localizados na parte de energia livre maior do que média. A única diferença notável quanto à distribuição das medidas de dispersão da curva do zanamivir é a posição da moda. De forma

geral, portanto, os fármacos do grupo B apresentaram uma melhor performance quanto à energia de afinidade com relação aos outros, tanto no que diz respeito à média, de -7,983 contra -5,581 e -6,067 kcal/mol dos grupos O e Z, quanto ao melhor valor à esquerda para a medida energética, sendo as afinidades mais baixas para os grupos B, O e Z de -10,123, -6,787 e -7,314 kcal/mol, respectivamente.

Figura 15. Distribuições de probabilidade representando as energias de afinidade da conformação mais estável dos ligantes com relação ao receptor. As curvas descrevem o comportamento dos análogos agrupados quanto aos seus respectivos fármacos base: (a) ácido baloxavir, (b) oseltamivir e (c) zanamivir.

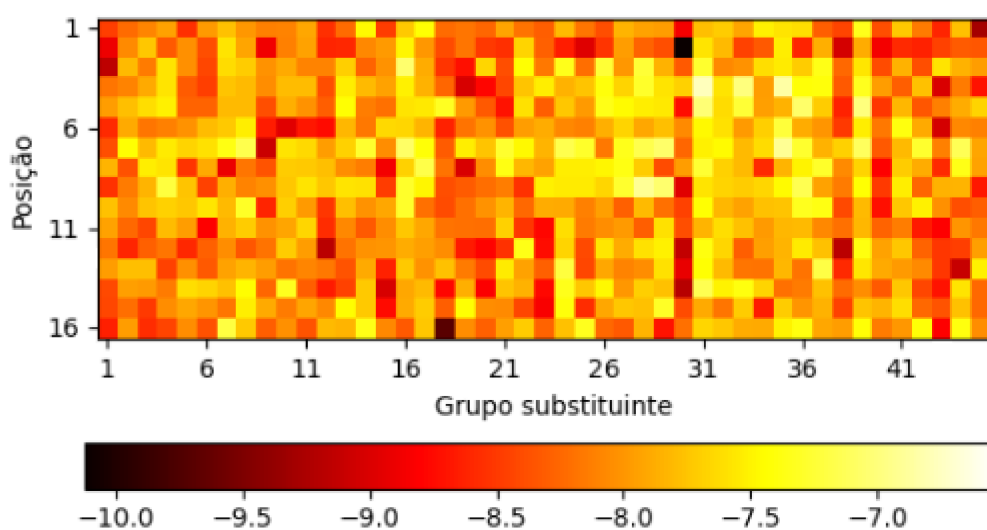


Quanto à modificação estrutural mais efetiva, uma análise visual dos mapas de calor presentes nas Figura 16, 17 e 18 mostram um claro padrão no qual o grupo substituinte 30 causou um abaixamento no valor de energia livre para todas os grupos de fármacos base. Além disso, é possível visualizar trações de melhor performance nos grupos R18, R38 e R40 para a classe B. As espécies B-1-3 e B-44-1 também exibiram um bom resultado. Os análogos A classe O, por sua vez, exibiram boa afinidade quando inseridos os grupos 12, 18, 20, 25, 33, 38 e 43, assim como uma aparição pontual relativa ao análogo O-13-15. Finalmente, o grupo Z também mostrou desempenho acima da média quando acrescido dos substituintes 25, 38 e 43. Traços de melhor performance são observados nas colunas 12, 18, 20 e 33. A similar-

idade dos substituintes presentes nas moléculas ancoradas com menor energia livre de ligação para O e Z pode ser oriunda da maior similaridade estrutural dos respectivos fármacos comerciais entre si do que com o ácido baloxavir. Os substituintes 12, 18, 20, 25, 30, 33 e 38 apresentam anéis aromáticos em suas estruturas, sendo que 12 e 18 não apresentam átomos eletronegativos ligados a eles e desempenharam de forma menos eficaz do que aqueles que os possuíam. No caso da classe B, a presença de dois resíduos contendo oxigênio parecem ter sido os responsáveis pela menor energia de afinidade dos análogos R30 e R38.

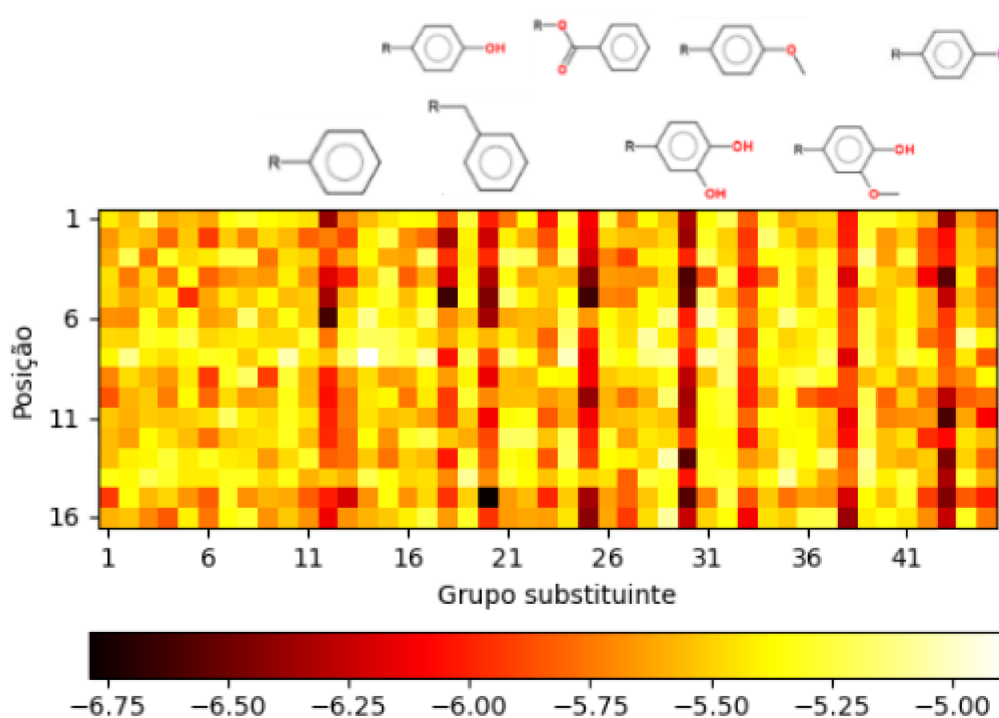
O ligante proveniente do dado cristalográfico, NAG, cadeia V do PDBID 6M17, se encontrava próximo dos resíduos ASN343, SER371 e VAL367 a um raio de corte de 5 Å a partir de qualquer átomo presente em sua estrutura. A maioria dos ligantes mencionados no parágrafo anterior como apresentando bons resultados, apresentou uma distância de aproximadamente 20 Å com relação à posição de NAG, no que diz respeito a centros cartesianos. O análogo B-37-2 se afastou ainda mais, apresentando uma distância de cerca de 47 Å. Assim sendo, é evidente que a energia de afinidade não pode ser o único critério a ser levado em conta quando se estuda candidatos a fármacos, uma vez que regiões de menor interesse farmacológico podem conduzir a uma maior probabilidade de ligação com o medicamento avaliado do que o alvo do estudo.

Figura 16. Mapa de calor das substituições feitas com a energia de ligação da conformação de afinidade mais estável da ancoragem molecular do grupo B.



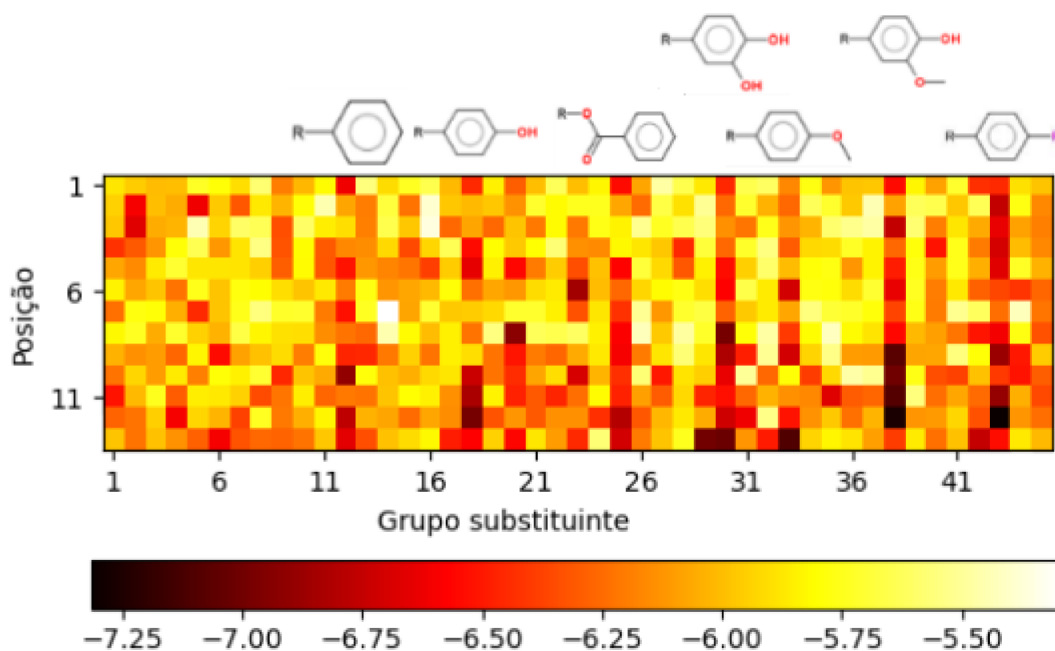
Fonte: Autor, 2023.

Figura 17. Mapa de calor das substituições feitas com a energia de ligação da conformação de afinidade mais estável da ancoragem molecular do grupo O.



Fonte: Autor, 2023.

Figura 18. Mapa de calor das substituições feitas com a energia de ligação da conformação de afinidade mais estável da ancoragem molecular do grupo Z.

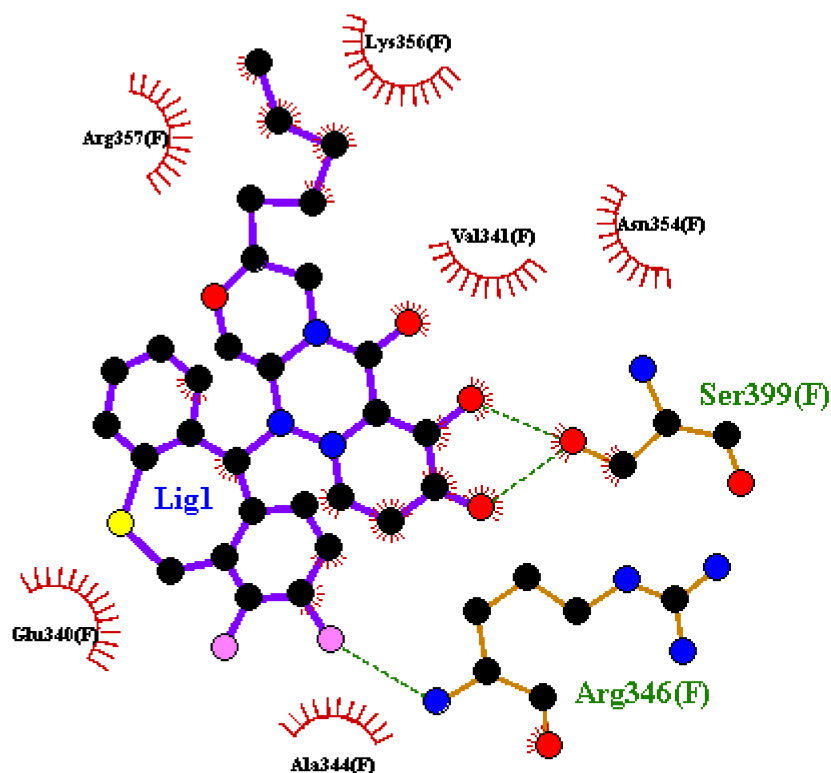


Fonte: Autor, 2023.

As espécies, à exceção do análogo B-37-2, apresentaram uma forte tendência a se aproximarem dos resíduos ALA348, ASN354, SER399 e THR345, com PHE347, LYS356 e ARG346 aparecendo em menor frequência. O-12-1, O-12-4, O-12-6, O-31-1, O-31-4, O-44-1, O-44-15 conseguiram ser ancorados a uma distância de, também aproximadamente 20 Å de NAG, porém foram capazes de se aproximarem do resíduo ASN343, apenas, dentre os mencionados. Resíduos comuns dentre eles envolvem ALA344, ALA348, ARG346, ARG457, ARG509, GLU340, GLY339, LYS356, PHE342, PHE347, VAL341 e THR345. O restante dos análogos foi conduzido a um sítio diferente durante o processo, distando algo entre 44 e 46 Å, onde foram ancorados perto dos resíduos ARG457, THR470, PRO491, ARG454, ASP467, ILE472, SER469, GLN474, TYR473, LYS458, SER459, GLU471, PHE456, comumente. No caso do grupo Z houve a mesma divisão em termos de distância de NAG, dentre os quais aqueles que mais se aproximaram deste foram as espécies Z-12-1, Z-12-10, Z-12-13, Z-18-5, Z-18-12, Z-18-13, Z-23-6, Z-23-10, Z-30-8, Z-33-6, Z-33-13, Z-38-10, Z-38-11, Z-38-12 e Z-43-12, os quais apresentaram semelhante afinidade pela região dos resíduos PHE342, ALA348, ARG346, ALA344, THR345, VAL511, ARG509, ASN354, PHE347, SER399, VAL341, GLU340, LYS356. Dos análogos mencionados, apenas Z-12-10 apresentou aproxi-

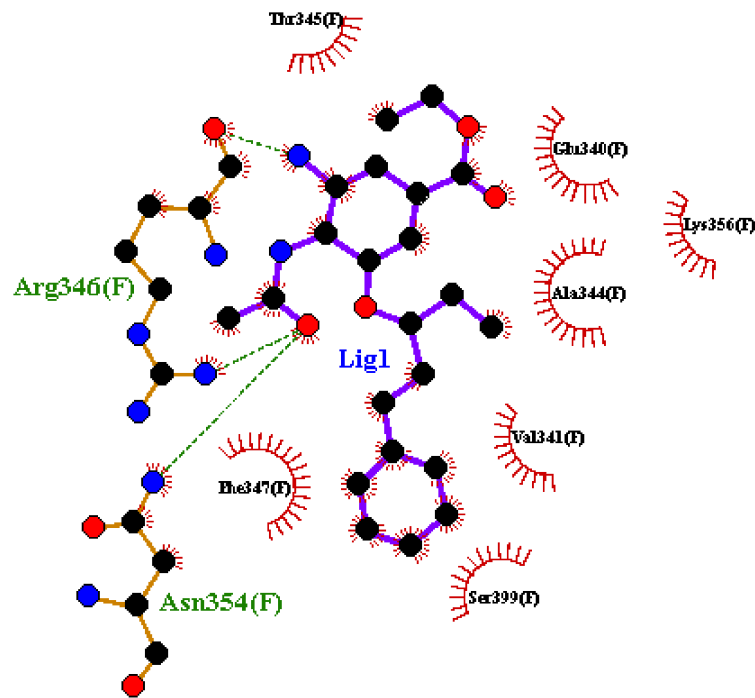
mação a ASN343, enquanto nenhum dos outros resíduos capturados no raio de corte de NAG foi exibido. As interações mais características para cada grupo mencionadas podem ser visualizadas por meio das figuras 19, 20 e 21, geradas pelo *software* LigPlot⁺ (LASKOWSKI; SWINDELLS, 2011).

Figura 19. Interações entre o ligante B-31-2 e o receptor.



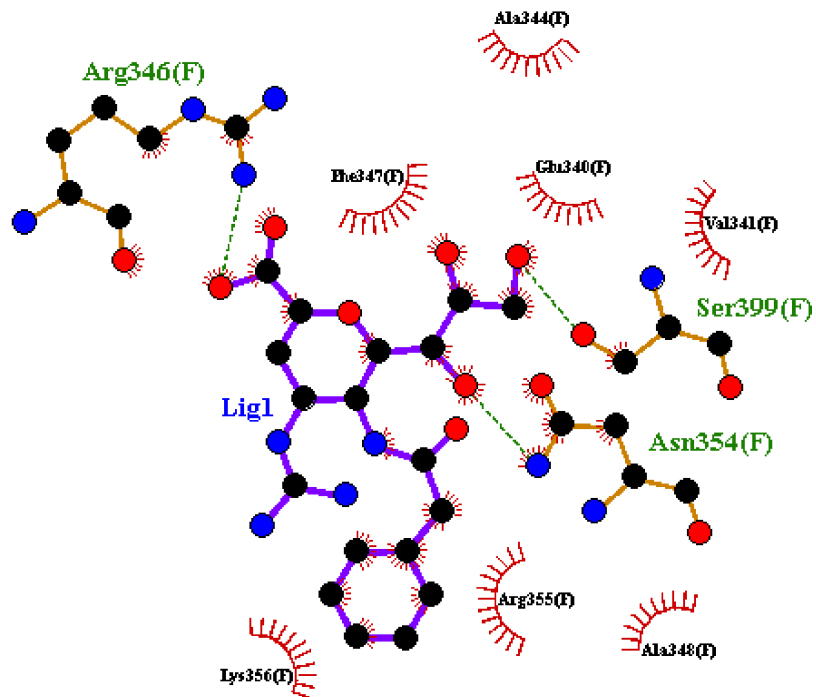
Fonte: Autor, 2023.

Figura 20. Interações entre o ligante O-12-1 e o receptor.



Fonte: Autor, 2023.

Figura 21. Interações entre o ligante O-12-1 e o receptor.

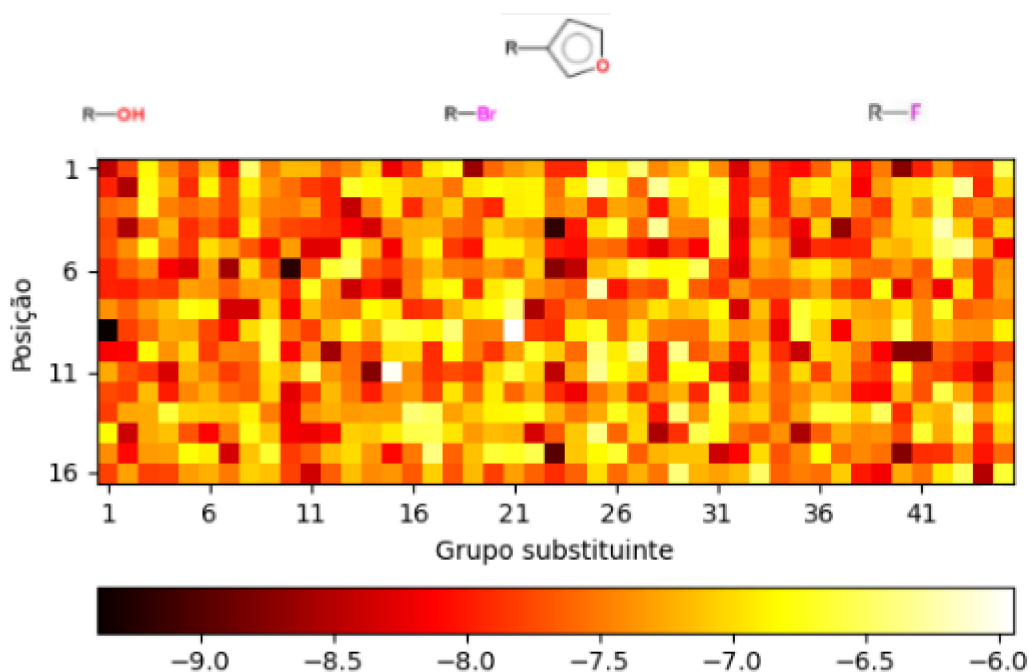


Fonte: Autor, 2023.

Por meio das Figuras 22, 23 e 24 no entanto, é possível visualizar os mapas de calor

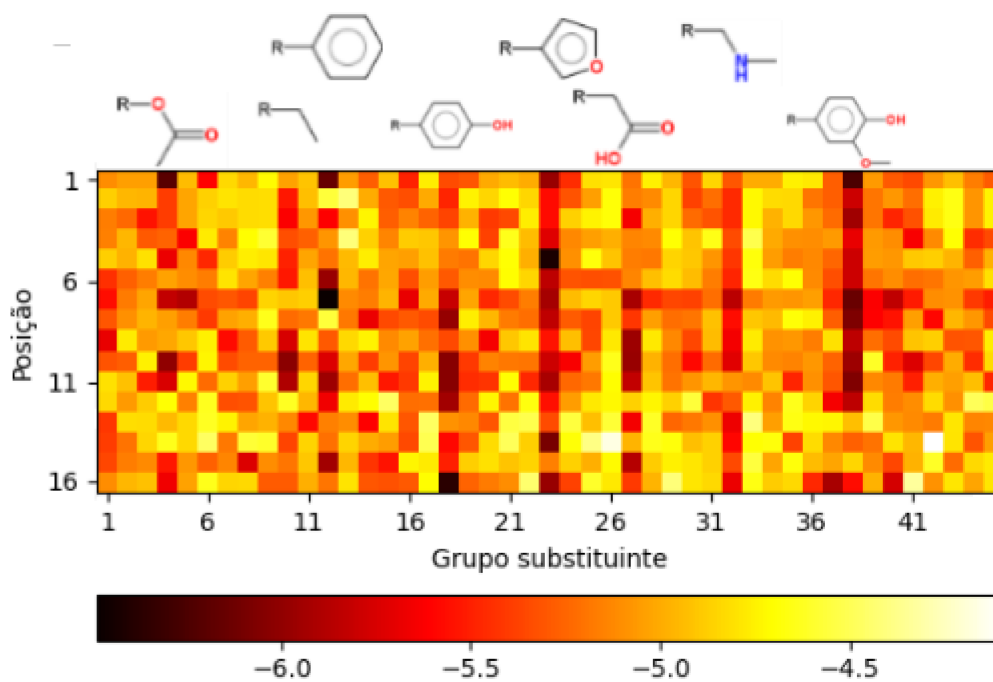
relativos à energia de afinidade das conformações de cada análogo que mais se aproximou da posição de NAG, sendo que, da triplicata, foi escolhida a repetição a qual demonstrou o melhor (menor) valor de distância euclidiana. Os menores valores de energia livre tendem a aparecer pontualmente para o grupo B para as classes C1, C4, C6, C9, C14 e C15, sem sequências definidas como foi visto na figura anterior. Analisando pelos agrupamentos R, também, não são observados padrões dotados de continuidade. Destes, B-18-1 foi quem mais se aproximou de NAG, a uma distância de 3 Å, e energia de afinidade de -8.630 kcal/mol, sendo localizados próximos a ele os resíduos ASN343, VAL36 e SER371. Os resultados dos outros grupos não exibem resultados tão satisfatórios em termos de distância e proximidade dos resíduos. As interações da molécula com a proteína são visíveis na Figura 25.

Figura 22. Mapa de calor dos análogos do grupo B com a energia de ligação da conformação cuja ancoragem molecular mais se aproximou de NAG.



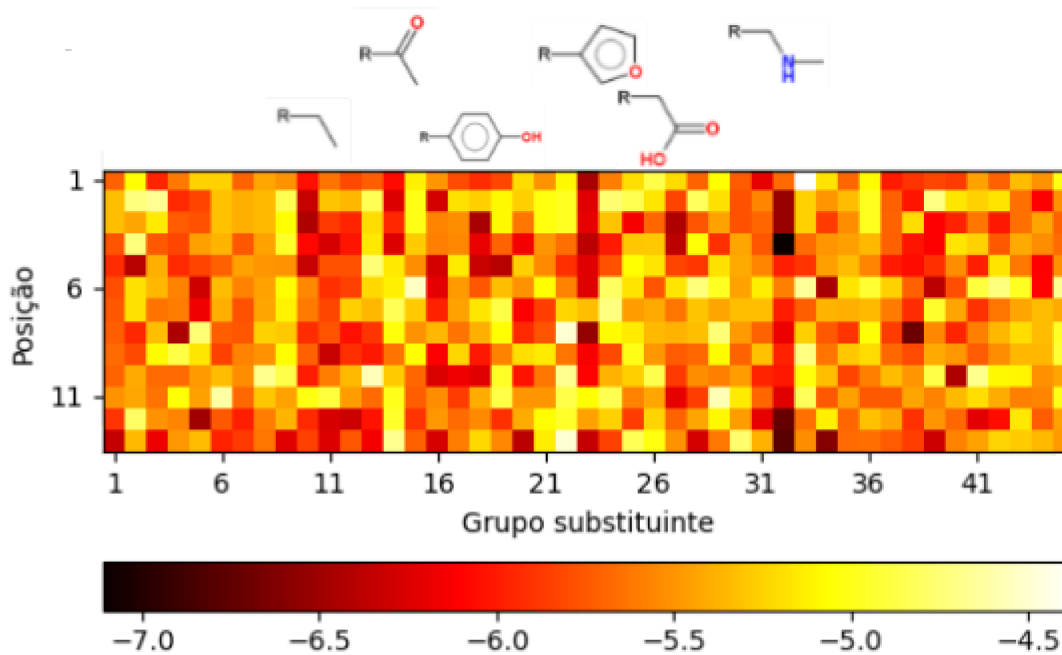
Fonte: Autor, 2023.

Figura 23. Mapa de calor dos análogos do grupo O com a energia de ligação da conformação cuja ancoragem molecular mais se aproximou de NAG.



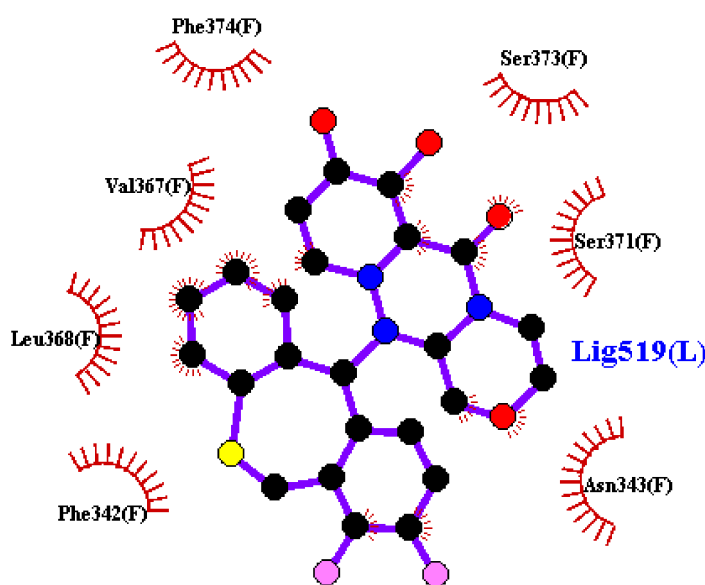
Fonte: Autor, 2023.

Figura 24. Mapa de calor dos análogos do grupo Z com a energia de ligação da conformação cuja ancoragem molecular mais se aproximou de NAG.



Fonte: Autor, 2023.

Figura 25. Interações entre o ligante B-18-1 e o receptor.



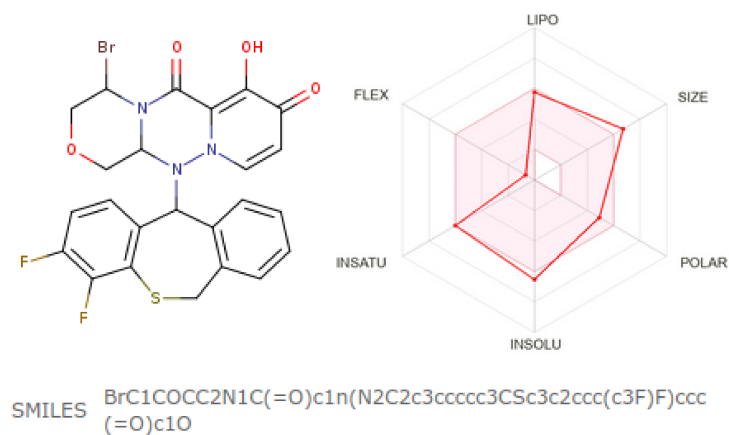
Fonte: Autor, 2023.

Como mencionado no fluxograma, seguindo o protocolo proposto, conduziu-se a dinâmica molecular do fármaco mencionado, o qual foi equilibrado dentro de 100 ns de simulação. Da referida trajetória, selecionou-se o quadro no qual foi observada uma dentre os maiores números de ligações de hidrogênio, sendo que não foram encontradas mais de 2 em toda a trajetória do análogo.

5.3 Filtros de viabilidade oral

A molécula foi avaliada segundo as regras de Lipinsky, Ghose, Vaber, Egan e Muegge. Nem todos os parâmetros listados para a RM foram empregados, no entanto. Uma vez que o fármaco base já satisfaziam os critérios: número de carbonos > 4 e número de heteroátomos > 1 , assumiu-se que todos os análogos deles derivados também passariam, sendo aprovada em todos eles. Algumas das propriedades físico-químicas mensuradas para a molécula via RDKit são descritas na Figura 26, contendo um gráfico de teia gerado no serviço web SwissADME (DAINA; MICHIELIN; ZOETE, 2017) usando o SMILES ilustrado abaixo.

Figura 26. Ilustração da estrutura do fármaco selecionado e dos resultados obtidos para alguns dos parâmetros calculados.

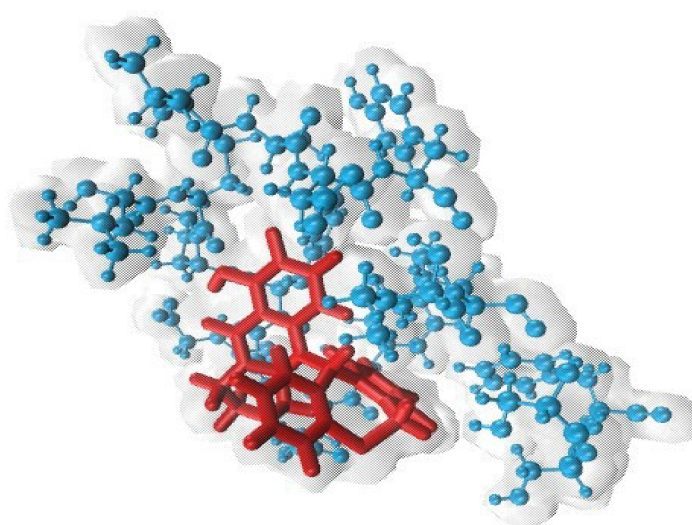


Fonte: DAINA; MICHIELIN; ZOETE, 2017.

4.3 Cálculo de Estrutura Eletrônica

Os cálculos envolvidos na TDFDT conduzida resultam na obtenção das energias associadas a cada orbital atômico, a exemplo da superfície vista na Figura 27, assim como nas energias de excitação necessárias para promover as transições eletrônicas deduzidas. Foram calculadas as energias de transição de dez eventos, como demonstra a Tabela 3.

Figura 27. Superfície eletrônica calculada para o recorte.



Fonte: Autor, 2023.

Tabela 3. Energias de transição para o sistema equilibrado a um raio de corte de 5 Å do ligante.

Estado	Gap (eV)	f_{osc}	Número de onda (cm-1)	Transições
1	0,330	$5,4427 \times 10^{-5}$	2659,7	H ₅ → L ₂ H ₁ → L ₂ H ₁ → L ₅
2	0,705	$9,1635 \times 10^{-5}$	5683,2	H ₅ → L ₂ H ₅ → L ₅ H ₁ → L ₂
3	0,745	$6,53 \times 10^{-7}$	6007,2	H ₁ → L ₁
4	0,904	$3,1185 \times 10^{-5}$	7288,9	H ₂ → L ₂ H ₂ → L ₅
5	0,935	$5,41 \times 10^{-7}$	7543,1	H ₂ → L ₁
6	0,983	$5,0446 \times 10^{-5}$	7926,2	H ₁ → L ₂ H ₁ → L ₅
7	1,114	$2,526 \times 10^{-6}$	8987,6	H ₁ → L ₄
8	1,149	$1,3186 \times 10^{-4}$	9297,3	H ₂ → L ₂ H ₂ → L ₅
9	1,164	$2,35 \times 10^{-7}$	9389,8	H ₄ → L ₂
10	1,191	3×10^{-9}	9605,0	H ₁ → L ₃

Fonte: Autor, 2023.

As transições observadas indicam uma tendência da transição 8 em ser a mais favorecida, a analisar pelo valor da força do oscilador f_{osc} . Porém, todas as transições são tão fracas, segundo o mesmo fator, que é improvável que qualquer uma delas ocorra, de fato, num experimento realista, uma vez que, como Lestrage, Egidi, e Li (2015) afirmam, o valor da força de oscilação é equivalente à probabilidade de ocorrência da transição, logo, pertencente aos extremos do intervalo [0; 1].

6 Conclusões

Os resultados demonstram uma maior tendência de grupos substituintes de natureza aromática a produzirem análogos farmacológicos com maior afinidade com a proteína *spike* do SARS-CoV-2, entretanto com um sítio diferente do alvo biológico. O ligante que mais se aproximou do local desejado foi obtido por meio da adição de um átomo eletronegativo volumoso Br ao ácido baloxavir, caracterizando uma energia de afinidade de -8,640 kcal/mol. De maneira geral, variantes do fármaco baloxavir apresentaram maior afinidade com a proteína do que os derivados das outras espécies tanto quanto para a menor afinidade global da AM, média de -7,930 contra -5,581 e -6,067 kcal/mol, quanto em casos de maior afinidade pelo sítio especificado, aproximadamente 0,5 kcal/mol menor em todos os casos. Acredita-se que adição de grupos com ligações rotacionáveis com propensão a formar tais ligações pode ser favorável à produção de variantes com maior chance de sucesso do que as avaliadas.

Uma vez que o melhor resultado da ancoragem molecular, segundo os critérios estabelecidos para seleção, conduziram a uma estrutura que exibiu baixo nível de ligações de hidrogênio ao longo de 100 ns de trajetória, Os resultados da TDFDT indicam um baixo nível de probabilidade de transferência de densidade de carga levando em conta os valores de força do oscilador calculados, cuja maior ordem foi de 10^{-4} , enfatizando a baixa eficiência da espécie B-18-1 como antiviral contra a proteína *spike* do SARS-CoV-2.

REFERÊNCIAS BIBLIOGRÁFICAS

ABU-SALEH, A. A. A.; YADAV, A.; POIRIER, R. A. **Accelerating the discovery of the beyond rule of five compounds that have high affinities toward SARS-CoV-2 spike RBD.** Journal of Biomolecular Structure and Dynamics, v. 41, n. 6, p. 2518-2527, 2023. DOI: 10.1080/07391102.2022.2036640. Disponível em: https://www.tandfonline.com/doi/full/10.1080/07391102.2022.2036640?casa_token=hxVJIJobXaEAAAAA%3A9vdqfkWRdHFufu1-6-rL_4lvDQPwTE7-vIqVLxkS8TAVTyikFfNul7tw1cjWNu4-RMJ_eJOEZwLSEAUrh-w.

AKSHAYA, J. et al. **A comparison between steepest descent and non-linear conjugate gradient algorithms for binding energy minimization of organic molecules.** Journal of Physics: Conference Series. IOP Publishing, 2023. p. 012004. DOI: 10.1088/1742-6596/2484/1/012004. Disponível em: <https://iopscience.iop.org/article/10.1088/1742-6596/2484/1/012004/meta>.

ALVES, V. M. et al. **Quimioinformática: uma introdução.** Química Nova, v. 41, p. 202-212, 2018. DOI: 10.21577/0100-4042.20170145. Disponível em: <https://www.scielo.br/j/qn/a/K6fzXSJWRGLmqzkbfbBsqrj/?lang=pt>.

ASHOUR, H. M. et al. **Insights into the recent 2019 novel coronavirus (SARS-CoV-2) in light of past human coronavirus outbreaks.** Pathogens, v. 9, n. 3, p. 186, 2020. DOI: 10.3390/pathogens9030186. Disponível em: <https://www.mdpi.com/2076-0817/9/3/186>.

BAERENDS, E. J. et al. **A DFT/TDFDT interpretation of the ground and excited states of porphyrin and porphyrazine complexes.** Coordination chemistry reviews, v. 230, n. 1-2, p. 5-27, 2002. DOI: 10.1016/S0010854502000930. Disponível em: https://www.sciencedirect.com/science/article/pii/S0010854502000930?casa_token=SPqxPLSGo54AAAAA:zSu1_8F2-XH4mqsBj_4rn9TeXlAkjaRp8Yb2e2aTHz7S1VqT7TvO2d46BWEwxwBfyRVepEmO0mC-BP.

BARDSLEY-ELLIOT, A.; NOBLE, S. **Oseltamivir.** Drugs, v. 58, p. 851-860, 1999. DOI: 10.2165/00003495-199958050-00007. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/1059-5865/>.

BEEKE, A. D. **Density-functional thermochemistry III: The role of exact exchange.** J. The Journal of Chemical Physics, v. 98, n. 7, p. 5648-6, 1993. DOI: 10.1063/1.464913. Disponível em: <https://pubs.aip.org/aip/jcp/article-abstract/98/7/5648/842114/Density-functional-thermochemistry-III-The-role-of-exact-exchange?redirectedFrom=fulltext>.

BELL, E. W.; ZHANG, Y. **DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism.** Journal of Cheminformatics, v. 11, n. 1, p. 1-9, 2019. DOI: 10.1186/s13321-019-0362-7. Disponível em: <https://jcheminf.biomedcentral.com/articles/10.1186/s13321-019-0362-7?ref=https://githubhelp.com>.

BERGWERF, H. **MolView.** 2015. Disponível em: <https://molview.org/>.

BERMAN, H. M. et al. **The protein data bank.** Nucleic acids research, v. 28, n. 1, p. 235-

242, 2000. DOI: 10.1093/nar/28.1.235. Disponível em: <https://academic.oup.com/nar/article/28/1/235/2384399?login=false>.

BEST, R. B. et al. **Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles.** Journal of chemical theory and computation, v. 8, n. 9, p. 3257-3273, 2012. DOI: 10.1021/ct300400x. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/ct300400x>.

BOHETS, H. et al. **Strategies for absorption screening in drug discovery and development.** Current topics in medicinal chemistry, v. 1, n. 5, p. 367-383, 2001. DOI: 10.2174/1568026013394886. Disponível em: <https://www.ingentaconnect.com/content/ben/c-tmc/2001/00000001/00000005/art00003>.

BUEHLER, M. J. **Multiscale modeling at the interface of molecular mechanics and natural language through attention neural networks.** Accounts of Chemical Research, v. 55, n. 23, p. 3387-3403, 2022. DOI: 10.1021/acs.accounts.2c00330. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.accounts.2c00330>.

CARABELLI, A. M. et al. **SARS-CoV-2 variant biology: immune escape, transmission and fitness.** Nature Reviews Microbiology, p. 1-16, 2023. DOI: 10.1038/s41579-022-00841-7. Disponível em: <https://www.nature.com/articles/s41579-022-00841-7>.

CARVALHO, T.; KRAMMER, F.; IWASAKI, A.. **The first 12 months of COVID-19: a timeline of immunological insights.** Nature Reviews Immunology, v. 21, n. 4, p. 245-256, 2021. DOI: 10.1038/s41577-021-00522-1. Disponível em: <https://www.nature.com/articles/s41577-021-00522-1>.

CASIDA, M. E.; HUIX-ROTLANT, M. **Progress in time-dependent density-functional theory.** Annual review of physical chemistry, v. 63, p. 287-323, 2012. DOI: 10.1146/annurev-physchem-032511-143803. Disponível em: https://www.annualreviews.org/doi/abs/10.1146/annurev-physchem-032511-143803?casa_token=YCCUAhHveqwAAAAA%3ANXudGeVN-y20bpV4u1rUCjONEUhhZmDq4q-Dtl9Wk8UQ6-ct-Tfo6ssBdftHWotriP8uDG_RU1N9ci7-YalA.

CHAKRABORTY, D.; CHATTARAJ, P. K. **Conceptual density functional theory based electronic structure principles.** Chemical Science, v. 12, n. 18, p. 6264-6279, 2021. DOI: <https://doi.org/10.1039/D0SC07017C>. Disponível em: <https://pubs.rsc.org/en/content/articlehtml/2021/sc/d0sc07017c>.

CHAMARY, J. V. **Coronavirus Uses Spikes To Break Into Cells - Here's How To Stop It.** Forbes, 17 abr. 2020. Disponível em: <https://www.forbes.com/sites/jychamary/2020/04/17/coronavirus-spike-protein/?sh=4d02d1992b27>. Acesso em: 17 fev. 2023.

CHENG, Tiejun et al. Computation of octanol– water partition coefficients by guiding an additive model with knowledge. **Journal of chemical information and modeling**, v. 47, n. 6, p. 2140-2148, 2007. DOI: 10.1021/ci700257y. Disponível: <https://pubs.acs.org/doi/abs/10.1021/ci700257y>.

CHOUDHURY, C.; MURUGAN, N. A.; PRIYAKUMAR, U. D. **Structure-based drug repurposing: Traditional and advanced AI/ML-aided methods**. *Drug Discovery Today*, 2022. DOI: 10.1016/j.drudis.2022.03.006. Disponível em: https://www.sciencedirect.com/science/article/pii/S135964462200112X?casa_token=c3bRTyyBDnUAAAAA:XkdeAeI_3xusKQMILNZtlh4D-AUXDM7qfyLXyRnFT7SeIji0PITy3jezpQSaSyBtLSEPAOZRUUgC.

COSTANZI, E. et al. **Structural and biochemical analysis of the dual inhibition of MG-132 against SARS-CoV-2 main protease (Mpro/3CLpro) and human cathepsin-L**. *International journal of molecular sciences*, v. 22, n. 21, p. 11779, 2021. DOI: 10.3390/ijms2-22111779. Disponível em: <https://www.mdpi.com/1422-0067/22/21/11779>.

COUMAR, M. S. **Molecular Docking for Computer-Aided Drug Design: Fundamentals, Techniques, Resources and Applications**. 1. ed. Reino Unido: Elsevier, 2021.

COVID, C. D. C. et al. **Severe outcomes among patients with coronavirus disease 2019 (COVID-19)—United States, February 12–March 16, 2020**. *Morbidity and mortality weekly report*, v. 69, n. 12, p. 343, 2020. Disponível em: https://www.cdc.gov/mmwr/volumes/69/wr/mm6912e2.htm?s_cid=mm6912e2_w.

DAINA, A.; MICHELIN, O.; ZOETE, V. **SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules**. *Scientific reports*, v. 7, n. 1, p. 42717, 2017. DOI: 10.1038/srep42717. Disponível em: <https://www.nature.com/articles/srep42717>.

DE AZEVEDO, W. F. **Docking screens for drug discovery**. New York, NY: Humana Press, 2019.

DE PAULA JUNIOR, V. F. et al. **Quantum Biochemistry and MM-PBSA Description of the ZIKV NS2B-NS3 Protease: Insights into the Binding Interactions beyond the Catalytic Triad Pocket**. *International Journal of Molecular Sciences*, v. 23, n. 17, p. 10088, 2022. DOI: 10.3390/ijms231710088. Disponível em: <https://www.mdpi.com/1422-0067/23-17/10088>.

DENSITY Functional (DFT) Methods. Gaussian, 2022. Disponível em: <<https://gaussian.com/dft/>>. Acesso em: 10 fev. 2023.

DUFRASNE, F. **Baloxavir marboxil: an original new drug against influenza**. *Pharmaceuticals*, v. 15, n. 1, p. 28, 2022. DOI: 10.3390/ph15010028. Disponível em: <https://www.mdpi.com/1424-8247/15/1/28>.

EBERHARDT, J. et al. **AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings**. *Journal of chemical information and modeling*, v. 61, n. 8, p. 3891-3898, 2021. DOI: 10.1021/acs.jcim.1c00203. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.jcim.1c00203>.

EGAN, W. J.; MERZ, K. M.; BALDWIN, J. J. **Prediction of drug absorption using multivariate statistics**. *Journal of medicinal chemistry*, v. 43, n. 21, p. 3867-3877, 2000. DOI: 10.1021/jm000292e. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/jm000292e>.

EGBERT, M. et al. **Why some targets benefit from beyond rule of five drugs.** Journal of medicinal chemistry, v. 62, n. 22, p. 10005-10025, 2019. DOI: 10.1021/acs.jmedchem.8b017-32. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.jmedchem.8b01732>.

ERTL, P. **Identification of bioisosteric substituents by a deep neural network.** Journal of Chemical Information and Modeling, v. 60, n. 7, p. 3369-3375, 2020. DOI: 10.1021/acs.jcim.-0c00290. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.jcim.0c00290>.

ERTL, P. **Substituents of life: The most common substituent patterns present in natural products.** Bioorganic & medicinal chemistry, v. 54, p. 116562, 2022. DOI: 10.1016/j.bmc.2021.116562. Disponível em: https://www.sciencedirect.com/science/article/pii/S0968089621005708?casa_token=dl3qUBREjmUAAAAA:g7TEHpCm927Ry6gLg6htO-P9imU3lH5bW-3TeFCGXV3S6UFQhjAYtGVolSE3yZnnodS4MHSWM9b1j.

FAN, J.; FU, A.; ZHANG, L. **Progress in molecular docking.** Quantitative Biology, v. 7, p. 83-89, 2019. DOI: 10.1007/s40484-019-0172-y. Disponível em: <https://link.springer.com/article/10.1007/s40484-019-0172-y>.

GAJJAR, N. D.; DHAMELIYA, T. M.; SHAH, G. B. **In search of RdRp and Mpro inhibitors against SARS CoV-2: Molecular docking, molecular dynamic simulations and ADMET analysis.** Journal of molecular structure, v. 1239, p. 130488, 2021. DOI: 10.1016/j.molstruc.2021.130488. Disponível em: https://www.sciencedirect.com/science/article/pii/S0022286021006219?casa_token=B43STW1U55gAAAAA:PB7iIu7qt3hAf6-FF2lAb-Yd-dHRfC09GcdBgD4fmjCuLU2fuyOeihUTSNzserPL8lGpVQzjdlZAY.

GHOSH, K. et al. **Chemical-informatics approach to COVID-19 drug discovery: Exploration of important fragments and data mining based prediction of some hits from natural origins as main protease (Mpro) inhibitors.** Journal of molecular structure, v. 1224, p. 129026, 2021. DOI: 10.1016/j.molstruc.2020.129026. Disponível em: https://www.sciencedirect.com/science/article/pii/S0022286020313491?casa_token=KoPCt_sgHV0AAAAA:jFd1QYT7XXBkjYjJRRbV-_aBdrk_bAZ0B-pPvKuRQgQuHCvcfSBlC7O_rEQPuH6E-e6RfAINYMT0r.

GUMBART, J. C.; LUO, D. **HBonds Plugin, Version 1.2. 2011.** Disponível em: <<https://www.ks.uiuc.edu/Research/vmd/plugins/hbonds/>>. Acesso em: 13 fev. 2023.

GUTIÉRREZ, I. S. et al. **Parametrization of halogen bonds in the CHARMM general force field: Improved treatment of ligand–protein interactions.** Bioorganic & medicinal chemistry, v. 24, n. 20, p. 4812-4825, 2016. DOI: 10.1016/j.bmc.2016.06.034. Disponível em: https://www.sciencedirect.com/science/article/pii/S0968089616304576?casa_token=1zsFFQ-o12BQAAAAA:sbG55KhaBuny-W2H0E2R66YaN_N3TgTRQ4bUkiaBJ5cY5JRImiormq-P-x8w1fAsJTs8lvFcqjlx.

HANSSON, T.; OOSTENBRINK, C.; VAN GUNSTEREN, W. **Molecular dynamics simulations.** Current opinion in structural biology, v. 12, n. 2, p. 190-196, 2002. DOI: 10.1016/S0959440X(02)00308-1. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0959440X02003081>.

HE, F.; DENG, Y.; LI, W. **Coronavirus disease 2019: What we know?** Journal of medical

virology, v. 92, n. 7, p. 719-725, 2020. DOI: 10.1002/jmv.25766. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.25766>.

HEHRE, W. J. **A guide to molecular mechanics and quantum chemical calculations**. Irvine, CA: Wavefunction, 2003.

HU, B. et al. **Characteristics of SARS-CoV-2 and COVID-19**. Nature Reviews Microbiology, v. 19, n. 3, p. 141-154, 2021. DOI: 10.1038/s41579-020-00459-7. Disponível em: <https://www.nature.com/articles/s41579-020-00459-7>.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. **VMD: visual molecular dynamics**. Journal of molecular graphics, v. 14, n. 1, p. 33-38, 1996. DOI: 10.1016/0263-7855(96)0001-8-5. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/-02637855960001-85>.

JABEEN, A.; RANGANATHAN, S. **Applications of machine learning in GPCR bioactive ligand discovery**. Current opinion in structural biology, v. 55, p. 66-76, 2019. DOI: 10.1016/j.sbi.2019.03.022. Disponível em: https://www.sciencedirect.com/science/article/pii/S0959440X19300181?casa_token=6bAGcdkSYOQAAAAA:gyElt0S20GaN36SGdgNdwTE-ym35zZjMQcWBI1nRK4PnlXVDsVRElwOiuGC1YB_XTxKZdZvxqeTUa.

JAYAWEERA, M. et al. **Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy**. Environmental research, v. 188, p. 109819, 2020. DOI: 10.1016/j.envres.2020.109819. Disponível em: https://www.sciencedirect.com/science/article/pii/S0013935120307143?casa_token=bn711a_Xmb4AAAAA:MIBY7acKP9I-EID6rGQq1k71nTYnAVd5z3KNu9A5yVCKVCim9MJevjsd8xMLf-ukAQWlVyoN8waR.

KALATHIYA, U. et al. **Highly conserved homotrimer cavity formed by the SARS-CoV-2 spike glycoprotein: a novel binding site**. Journal of clinical medicine, v. 9, n. 5, p. 1473, 2020. DOI: 10.3390/jcm9051473. Disponível em: <https://www.mdpi.com/2077-0383/9/5/1473>.

KINGHORN, D. **Molecular Dynamics Benchmarks GPU Roundup GROMACS NAMD2 NAMD 3alpha on 12 GPUs**. 2022. Disponível em: <https://www.pugetsystems.com/labs/hpc/Molecular-Dynamics-Benchmarks-GPU-Roundup-GROMACS-NAMD2-3alpha-on-12-GPUs-2330>. Acesso em: 22 jul. 2023.

KLAUDA, J. B. et al. **Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types**. The journal of physical chemistry B, v. 114, n. 23, p. 7830-7843, 2010. DOI: 10.1021/jp101759q. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/jp101759q>.

KOVALEVSKY, A. et al. **Unmasking the conformational stability and inhibitor binding to SARS-CoV-2 main protease active site mutants and miniprecursor**. Journal of Molecular Biology, v. 434, n. 24, p. 167876, 2022. DOI: 10.1016/j.jmb.2022.167876. Disponível em: https://www.sciencedirect.com/science/article/pii/S002228362200496X?casa_token=FL95W80dmhwAAAAA:zaEW_vELuZHhi8MM1kqj10DcHYbDrC8Q6ipEvVlpl_Ah06ae0IOOImgAHlxoeo-F42Q-XhUP_4C.

KRÄMER, S. D. et al. **When barriers ignore the “rule-of-five”**. Advanced drug delivery re-

views, v. 101, p. 62-74, 2016. DOI: 10.1016/j.addr.2016.02.001. Disponível em: https://www.sciencedirect.com/science/article/pii/S0169409X16300515?casa_token=Pz7l8s4-jpMA-AAAA:kqDF_-JyHtbxEMstyFhY5z5jgEjlymC-3_fm4AmzIZDgfPLJN829wQqDSuZBakT-OIvB8PbA9MSB7.

KÜHNE, T. D. et al. **CP2K: An electronic structure and molecular dynamics software package-Quickstep: Efficient and accurate electronic structure calculations.** The Journal of Chemical Physics, v. 152, n. 19, p. 194103, 2020. DOI: 10.1063/5.0007045. Disponível em: <https://pubs.aip.org/aip/jcp/article/152/19/194103/199081>.

LASKOWSKI, R. A.; SWINDELLS, M. B. **LigPlot+: multiple ligand-protein interaction diagrams for drug discovery.** Journal of Chemical information and Modeling, v. 51, n. 10, p. 2778-2786, 2011. DOI: 10.1021/ci200227u. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/CI200227U>.

L'EXACT, M. et al. **Beyond rule-of-five: Permeability assessment of semi-PEPTIDIC macrocycles.** Biochimica et Biophysica Acta (BBA)-Biomembranes, p. 184196, 2023. DOI: 10.1016/j.bbamem.2023.184196. Disponível em: https://www.sciencedirect.com/science/article/pii/S0005273623000780?casa_token=0VmkriA9ZDEAAAAA:cAmApR2FkGM7mdqOo-OgQtjluaiiNrTSvJmLq_QomruTFObiKxcQhl9FpiKlbZftJX9LnaxOO36Vd.

LESTRANGE, P. J.; EGIDI, F.; LI, X. **The consequences of improperly describing oscillator strengths beyond the electric dipole approximation.** The Journal of Chemical Physics, v. 143, n. 23, 2015. DOI: 10.1063/1.4937410. Disponível em: <https://pubs.aip.org/aip/jcp/article/143/23/234103/194033>.

LIPINSKI, C. A. et al. **Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.** vol. 46, issue 1-3. Adv Drug Deliv Rev, p. 3-26, 2001. DOI: S0169-409X(96)00423-1. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0169409X96004231>.

MAHESWARI, C. U.; MUTHU, S.; SUNDIUS, T. **Combined quantum-mechanics/molecular-mechanics, molecular docking studies on hemorrhoid drug.** Results in Chemistry, v. 5, p. 100962, 2023. DOI: 10.1016/j.rechem.2023.100962. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2211715623002011>.

MALLA, T. R. et al. **Penicillin derivatives inhibit the SARS-CoV-2 main protease by reaction with its nucleophilic cysteine.** Journal of Medicinal Chemistry, v. 65, n. 11, p. 7682-7696, 2022. DOI: 10.1021/acs.jmedchem.1c02214. Disponível em: <https://pubs.acs.org/doi/full/10.1021/acs.jmedchem.1c02214>.

MANDAL, I.; MANNA, S.; VENKATRAMANI, R. **UV-Visible Lysine-Glutamate dimer excitations in protein charge transfer spectra: TDFDT descriptions using an optimally tuned CAM-B3LYP functional.** The Journal of Physical Chemistry B, v. 123, n. 51, p. 10967-10979, 2019. DOI: 10.1021/acs.jpcc.9b07827. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.jpcc.9b07827>.

MARTÍNEZ, L. **Automatic identification of mobile and rigid substructures in molecular dynamics simulations and fractional structural fluctuation analysis.** PloS one, v. 10, n. 3,

p. e0119264, 2015. DOI: 10.1371/journal.pone.0119264. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0119264>.

MEDEIROS FILHO, F. C. et al. **Estudo Da Inibição Da Acetilcolinesterase Por Docking Molecular: Aplicação No Tratamento Da Doença Do Alzheimer**. Educação, Ciência E Saúde, v. 7, n. 2, p. 18, 2020. DOI: 10.20438/ccs.v7i2.297. Disponível em: <http://periodicos.ces.ufcg.edu.br/periodicos/index.php/99cienciaeducacaosaude25/article/view/297>.

MELI, R.; BIGGIN, P. C. **Spyrmsd: symmetry-corrected RMSD calculations in Python**. Journal of Cheminformatics, v. 12, n. 1, p. 1-7, 2020. DOI: 10.1186/s13321-020-00455-2. Disponível em: <https://jcheminf.biomedcentral.com/articles/10.1186/s13321-020-00455-2>.

MISRA, M.; YADAV, A. K. **Assessment of Available AMBER Force Fields to Model DNA-Ligand Interactions**. Biointerface Research on Applied Chemistry, v. 13, issue 2, 2022. DOI: 10.33263/BRIAC132.156. Disponível em: https://www.researchgate.net/profile/Manas-Misra/publication/359786912_Assessment_of_Available_AMBER_Force_Fields_to_Model_DNA-Ligand_Interactions/links/624e5c0def013420665de0e7/Assessment-of-Availab-le-AMBER-Force-Fields-to-Model-DNA-Ligand-Interactions.pdf.

MORIGUCHI, I. et al. **Simple method of calculating octanol/water partition coefficient**. Chemical and pharmaceutical bulletin, v. 40, n. 1, p. 127-130, 1992. DOI: 10.1248/cpb.40.127. Disponível em: https://www.jstage.jst.go.jp/article/cpb1958/40/1/40_1_127/_article/-char/ja/.

MORIGUCHI, I. et al. **Comparison of reliability of log P values for drugs calculated by several methods**. Chemical and pharmaceutical bulletin, v. 42, n. 4, p. 976-978, 1994. DOI: 10.1248/cpb.42.976. Disponível em: https://www.jstage.jst.go.jp/article/cpb1958/42/4/42_4_976/_article/-char/ja/.

MORRIS, G. M. et al. **AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility**. Journal of computational chemistry, v. 30, n. 16, p. 2785-2791, 2009. DOI: 10.1002/jcc.21256. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21256>.

MUEGGE, I.; HEALD, S. L.; BRITTELLI, D. **Simple selection criteria for drug-like chemical matter**. Journal of medicinal chemistry, v. 44, n. 12, p. 1841-1846, 2001. DOI: 10.1021/jm015507e. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/jm015507e>.

NARAYANAN, A. et al. **Identification of SARS-CoV-2 inhibitors targeting Mpro and PLpro using in-cell-protease assay**. Communications Biology, v. 5, n. 1, p. 169, 2022. DOI: 10.1038/s42003-022-03090-9. Disponível em: <https://www.nature.com/articles/s42003-022-03090-9>.

NATARELLI, L. et al. **MicroRNAs and long non-coding RNAs as potential candidates to target specific motifs of SARS-CoV-2**. Non-coding RNA, v. 7, n. 1, p. 14, 2021. DOI: 10.3390/ncrna7010014. Disponível em: <https://www.mdpi.com/2311-553X/7/1/14>.

NATIONAL SCIENCE FOUNDATION – NSF; NATIONAL INSTITUTE OF HEALTH –

NIH. **AutoDock Vina Manual**. 2018. Disponível em: <<https://vina.scripps.edu/manual/>>. Acesso em: 13 fev. 2023.

NERENBERG, P. S.; HEAD-GORDON, T. **New developments in force fields for biomolecular simulations**. *Current opinion in structural biology*, v. 49, p. 129-138, 2018. DOI: 10.1016/j.sbi.2018.02.002. Disponível em: https://www.sciencedirect.com/science/article/pii/S0959440X17301604?casa_token=WU-5q5IFu7YAAAAA:o-MwmuBEzX1IAmWSI-FWk2XGWb1bq0iUw8esJDrD3xN99WoH7ut2uupzAg8ZOVCL2mOwxFbgiI3RG.

NEESE, F. et al. **The ORCA quantum chemistry program package**. *The Journal of chemical physics*, v. 152, n. 22, p. 224108, 2020. DOI: 10.1063/5.0004608. Disponível em: <https://pubs.aip.org/aip/jcp/article/152/22/224108/1061982>.

O'BOYLE, N. M. et al. **Open Babel: An open chemical toolbox**. *Journal of cheminformatics*, v. 3, n. 1, p. 1-14, 2011. DOI: 10.1186/1758-2946-3-33. Disponível em: <https://jcheminf.biomedcentral.com/articles/10.1186/1758-2946-3-33?report=reader>.

O'BOYLE, N. M. et al. **The Open Babel Package**, versão 2.3.2 [S.l.], 27 jul. 2017. Disponível em: <<http://openbabel.org>>. Acesso em: 16 nov. 2022.

ORELLANA, L. **Large-scale conformational changes and protein function: breaking the in silico barrier**. *Frontiers in molecular biosciences*, v. 6, p. 117, 2019. DOI: 110.3389/fmolb.2019.00117. Disponível em: <https://www.frontiersin.org/articles/10.3389/fmolb.2019.00117/full>.

ORGANIZAÇÃO MUNDIAL DE SAÚDE – OMS. **WHO Coronavirus (COVID-19) Dashboard**. 2023. Disponível em: <<https://covid19.who.int/>>.

PARVATHANENI, V. et al. **Drug repurposing: a promising tool to accelerate the drug discovery process**. *Drug discovery today*, v. 24, n. 10, p. 2076-2085, 2019. DOI: 10.1016/j.drudis.2019.06.014. Disponível em: https://www.sciencedirect.com/science/article/pii/S1359644618302599?casa_token=Zb6ksY50bIIAAAAA:h0VxhhZdnBwZdVqiYnSKgzi-m9c-Djrt3wY1c6i6YQNNQgZf9elFj8VQfMwc_pozI0gKxMCi5HMttw.

PHILLIPS JAMES, C. et al. 2020. **Scalable molecular dynamics on CPU and GPU architectures with NAMD**. *J. Chem. Phys.*, v. 153, n. 044130, p. 10.1063, 2020. DOI: 10.1063/5.0014475. Disponível em: <https://pubs.aip.org/aip/jcp/article/153/4/044130/106495-3>.

POLLASTRI, M. P. **Overview on the Rule of Five**. *Current protocols in pharmacology*, v. 49, n. 1, p. 9.12. 1-9.12. 8, 2010. DOI: 10.1002/0471141755.ph0912s49. Disponível em: <https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/0471141755.ph0912s49>.

RDKit: Open-source cheminformatics. 2023. Disponível em: <<https://www.rdkit.org/>>. Acesso em: 04 de jul de 2023.

RENNER, J.; SCHMIDT, M; DE LAS HERAS, D. **Custom flow in molecular dynamics**. *Physical Review Research*, v. 3, n. 1, p. 013281, 2021. DOI: 10.1103/PhysRevResearch.3.013281. Disponível em: <https://journals.aps.org/prresearch/abstract/10.1103/PhysRevResearch.3.013281>.

search.3.013281.

SARGSYAN, K.; GRAUFFEL, C.; LIM, C. **How molecular size impacts RMSD applications in molecular dynamics simulations.** Journal of chemical theory and computation, v. 13, n. 4, p. 1518-1524, 2017. DOI: 10.1021/acs.jctc.7b00028. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.jctc.7b00028>.

SCHÜTTE, C.; KLUS, S.; HARTMANN, C. **Overcoming the timescale barrier in molecular dynamics: Transfer operators, variational principles and machine learning.** Acta Numerica, v. 32, p. 517-673, 2023. DOI: 10.1017/S0962492923000016. Disponível em: <https://www.cambridge.org/core/journals/acta-numerica/article/overcoming-the-timescale-barrier-in-molecular-dynamics-transfer-operators-variational-principles-and-machine-learning/AFE1143DC57B9D2B33774997A620E277>.

SHANG, J. et al. **Structural basis of receptor recognition by SARS-CoV-2.** Nature, v. 581, n. 7807, p. 221-224, 2020. DOI: 10.1038/s41586-020-2179-y. Disponível em: <https://www.nature.com/articles/s41586%2020%202179-y>.

SEWELL, T. D. et al. **A molecular dynamics simulation study of elastic properties of HMX.** The Journal of chemical physics, v. 119, n. 14, p. 7417-7426, 2003. DOI: 10.1063/1.1599273. Disponível: <https://pubs.aip.org/aip/jcp/article-abstract/119/14/7417/534469/A-molecular-dynamics-simulation-study-of-elastic>.

SINGH, T. U. et al. **Drug repurposing approach to fight COVID-19.** Pharmacological Reports, v. 72, p. 1479-1508, 2020. DOI: 10.1007/s43440-020-00155-6. Disponível em: <https://link.springer.com/article/10.1007/s43440-020-00155-6>.

STANZIONE, F.; GIANGRECO, I.; COLE, J. C. **Use of molecular docking computational tools in drug discovery.** Progress in Medicinal Chemistry, v. 60, p. 273-343, 2021. DOI: 10.1016/bs.pmch.2021.01.004. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0079646821000047>.

TRIPATHI, A.; BANKAITIS, V. A. **Molecular docking: from lock and key to combination lock.** Journal of molecular medicine and clinical applications, v. 2, n. 1, 2017. DOI: 10.16966/2575-0305.106. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5764188/>.

TROTT, O.; OLSON, A. J. **AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.** Journal of computational chemistry, v. 31, n. 2, p. 455-461, 2010. DOI: 10.1002/jcc.21334. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21334?casa_token=L_ZHHIDRo3Y-AAAAA%3AwHI3qvbSa52digWTXcGqjUi3-FRPvD2cDNJAmqlyp_Dj3SdwCEVXZzcA6-cZaIBLLeyuHzNpFafJqJHHeg.

TÜZÜN, B.; SARIPINAR, E. **Molecular docking and 4D-QSAR model of methanone derivatives by electron conformational-genetic algorithm method.** Journal of the Iranian Chemical Society, v. 17, p. 985-1000, 2020. DOI: 10.1007/s13738-019-01835-8. Disponível em: <https://link.springer.com/article/10.1007/s13738-019-01835-8>.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN. **NAMD Energy Plugin**, Version 1.4. 2017. Disponível em: <https://www.ks.uiuc.edu/Research/vmd/plugins/namd-energy/>. Acesso em: 13 fev. 2023.

U.S. FOOD AND DRUG ADMINISTRATION (FDA). **The Drug Development Process**. 2018. Disponível em: <https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process>. Acesso em: 16 mai, 2023.

VANOMMESLAEGHE, K. et al. **CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields**. *Journal of computational chemistry*, v. 31, n. 4, p. 671-690, 2010. DOI: 10.1002/jcc.21367. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21367?casa_token=UVm-ARGyj7d4AAAAA%3Ah6Blmi4YBUKhmjMqi76LXrMBFR9dbNNXU3yn-L76a-rk3_IfetS-a6xAXxtjYpYMcOl_ZNPR6jS5tauWDA.

VANOMMESLAEGHE, K.; MACKERELL JR, A. D. **Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing**. *Journal of chemical information and modeling*, v. 52, n. 12, p. 3144-3154, 2012. DOI: 10.1021/ci300363c. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/ci300363c>.

VANOMMESLAEGHE, K.; RAMAN, E. P.; MACKERELL JR, A. D. **Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges**. *Journal of chemical information and modeling*, v. 52, n. 12, p. 3155-3168, 2012. DOI: 10.1021/ci3003649. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/ci3003649>.

VISSCHER, K. M.; GEERKE, D. P. **Deriving force-field parameters from first principles using a polarizable and higher order dispersion model**. *Journal of chemical theory and computation*, v. 15, n. 3, p. 1875-1883, 2019. DOI: 10.1021/acs.jctc.8b01105. Disponível em: <https://pubs.acs.org/doi/full/10.1021/acs.jctc.8b01105>.

WAGHORN, S. L.; GOA, K. L. **Zanamivir**. *Drugs*, v. 55, n. 5, 1998. DOI: 10.2165/00003495-199855050-00016. Disponível em: <https://link.springer.com/article/10.2165/00003495-199855050-00016>.

WANG, R.; LAI, L.; WANG, S. **Further development and validation of empirical scoring functions for structure-based binding affinity prediction**. *Journal of computer-aided molecular design*, v. 16, p. 11-26, 2002. DOI: 10.1023/A:1016357811882. Disponível em: <https://link.springer.com/article/10.1023/A:1016357811882>.

WANG, Y. et al. **RNA-dependent RNA polymerase of SARS-CoV-2 as a therapeutic target**. *Journal of medical virology*, v. 93, n. 1, p. 300-310, 2021. DOI: 10.1002/jmv.26264. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.26264?casa_token=9qD-R_GoPCF0AAAAA%3ALIM50V4VEFTkiA5fMxHHO9ITKExX2I2pwvvhzIQeUKXz2xuQ-9_naQ2CrEKGm32B7KYKCeuo_k2afogP-uUw.

WANG, Y. et al. **End-to-end differentiable construction of molecular mechanics force fields**. *Chemical Science*, v. 13, n. 41, p. 12016-12033, 2022. DOI: 10.1039/D2SC02739A. Disponível em: <https://pubs.rsc.org/en/content/articlehtml/2022/sc/d2sc02739a>.

WILDMAN, S. A.; CRIPPEN, G. M. **Prediction of physicochemical parameters by atomic contributions**. *Journal of chemical information and computer sciences*, v. 39, n. 5, p. 868-873, 1999. DOI: 10.1021/ci990307l. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/ci990307l>.

YAN, R. et al. **Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2**. *Science*, v. 367, n. 6485, p. 1444-1448, 2020. DOI: 10.1126/science.abb2762. Disponível em: <https://www.science.org/doi/full/10.1126/science.abb2762>.

YANG, H. et al. **Identification of the probable structure of the sAPP α -GABA B R1a complex and theoretical solutions for such cases**. *Physical Chemistry Chemical Physics*, v. 24, n. 20, p. 12267-12280, 2022. DOI: 10.1039/D2CP00569G. Disponível em: <https://pubs.rsc.org/en/content/articlelanding/2022/cp/d2cp00569g/unauth>.

YU, W. et al. **Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations**. *Journal of computational chemistry*, v. 33, n. 31, p. 2451-2468, 2012. DOI: 10.1002/jcc.23067. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.23067?casa_token=Tjyq1FrUy-kAAAAA%3AKaamo-Q4X2CkLL75N0a1GrG5gNSXj94gN84Wag9nmOuqEaLneXXgVyzm_BL_a5uu19hHB-BzrvhdXs_ylYOA,

ZHANG, M.; WILKINSON, B. **Drug discovery beyond the 'rule-of-five'**. *Current opinion in biotechnology*, v. 18, n. 6, p. 478-488, 2007. DOI: 10.1016/j.copbio.2007.10.005. Disponível em: https://www.sciencedirect.com/science/article/pii/S0958166907001279?casa_token=-ialab4wIyMAAAAA:tcwr_iMMIPzmleYNSjCdabTpAQNVt9dnjxRSDDemcRZrxG6e0vzxOagiJrHtf7AKMIDXL21jJr.

ZHANG, Y.; OTANI, A.; MAGINN, E. J. **Reliable viscosity calculation from equilibrium molecular dynamics simulations: A time decomposition method**. *Journal of chemical theory and computation*, v. 11, n. 8, p. 3537-3546, 2015. DOI: 10.1021/acs.jctc.5b00351. Disponível em: <https://pubs.acs.org/doi/abs/10.1021/acs.jctc.5b00351>.

ZHAO, J.; CAO, Y.; ZHANG, L. **Exploring the computational methods for protein-ligand binding site prediction**. *Computational and structural biotechnology journal*, v. 18, p. 417-426, 2020. DOI: 10.1016/j.csbj.2020.02.008. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2001037019304465>.

ZHAO, Y. et al. **Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2**. *Proceedings of the National Academy of Sciences*, v. 119, n. 16, p. e2117142119, 2022. DOI: 10.1073/pnas.2117142119. Disponível em: <https://www.pnas.org/doi/abs/10.1073/pnas.2117142119>.

ZOETE, V. et al. **SwissParam: a fast force field generation tool for small organic molecules**. *Journal of computational chemistry*, v. 32, n. 11, p. 2359-23, 2011. DOI: 10.1002/jcc.21816. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21816?casa_token=DNRk5Bc-GxEAAAAA%3AbD0GXAYAQOMbX6kaMTL8HgyGQFiITknTqEnyGLNcWJQc0lDjBIFQrHSIhHwwMsygHNYWVb8ojgTWKvrLvW.

ZHOU, K.; LIU, B. **Molecular Dynamics Simulation: Fundamentals and Applications.** Amsterdam: Elsevier, 2022.