



**Universidade Federal da Paraíba**  
**Centro de Ciências da Saúde**  
**Programa de Pós-Graduação em Produtos**  
**Naturais e Sintéticos Bioativos**



**Herbert Igor Rodrigues de Medeiros**

**Planejamento e Desenvolvimento Racional de Candidatos a**  
**Fármacos Inibidores de Influenza A**

**João Pessoa – PB**

**2021**

**Herbert Igor Rodrigues de Medeiros**

**Planejamento e Desenvolvimento Racional de Candidatos a Fármacos  
Inibidores de Influenza A**

**Dissertação de Mestrado do Programa de Pós-graduação em Produtos Naturais e Sintéticos Bioativos da UFPB, como requisito para obtenção do título de Mestre em Produtos Naturais e Sintéticos Bioativos, na área de concentração: Farmacoquímica**

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Luciana Scotti

**Área de concentração: Farmacoquímica**

**João Pessoa – PB**

2021

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

M488p Medeiros, Herbert Igor Rodrigues de.

Planejamento e desenvolvimento racional de  
candidatos a fármacos inibidores de Influenza A /  
Herbert Igor Rodrigues de Medeiros. - João Pessoa,  
2021.

105 f. : il.

Orientação: Luciana Scotti.

Dissertação (Mestrado) - UFPB/CCS.

1. Influenza A. 2. Multitarget. 3. Química  
medicinal. 4. Docagem molecular. I. Scotti, Luciana.  
II. Título.

UFPB/BC

CDU 616.921.5(043)

Elaborado por Gracilene Barbosa Figueiredo - CRB-15/794

**Herbert Igor Rodrigues de  
Medeiros**

**Planejamento e Desenvolvimento Racional de Candidatos a Fármacos  
Inibidores de Infoduenza A**

Aprovada em 06/12/2021

**Banca Examinadora**



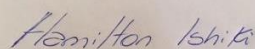
---

Profa. Dra. Luciana Scotti  
(Pós-graduação em Produtos Naturais e Sintéticos Bioativos,  
UFPB)  
Presidente



---

Prof. Dr. Francisco Jaime B. Mendonça Jr.  
(Pós-graduação em Produtos Naturais e Sintéticos Bioativos,  
UFPB)  
Examinador interno



---

Prof. Dr. Hamilton Mitsugu  
Ishiki  
(Pós-graduação em Ciências da Saúde,  
Unoeste)  
Examinador Externo

## DEDICATÓRIA

*Aos meus pais, Francisca Costa (in memoriam) e Paulo Rodrigues, minha companheira Iany Louise e toda a minha família, eu dedico este trabalho, pelos exemplos de honestidade, simplicidade e por todo esforço em me proporcionar a oportunidade de seguir e realizar meus sonhos.*

## AGRADECIMENTOS

Obrigado Meu Senhor Jesus Cristo, sem ti não estaria aqui e agora. Obrigado por ter incluído anjos durante a minha jornada, por estar sempre ao meu lado, por me guiar nesta missão que só o senhor sabe. Mesmo sem ter conhecimentos, sem entender certas passagens, agradeço pelo senhor nunca ter me abandonado. O senhor é só o senhor é dono disto tudo.

A minha grandiosa, santíssima e excelentíssima Mãe, Francisca Costa (*in memoriam*), ao qual sou fruto de uma mãe de caráter e coração imensurável. Não tenho palavras para te agradecer, pois nada chegará perto do que fizestes por mim. Obrigado por não ter medido esforços para me ajudar e por todos os momentos de apoio em família. Sou eternamente grato ao nosso Senhor Jesus Cristo por ter me dado a honra de ter passado 23 anos da minha vida na sua presença, por ter sido seu filho. “Eu sempre vou te amar...”

Ao meu grande pai, Paulo Rodrigues, que sempre fez de tudo e se sacrificava diariamente para que nunca me falte nada. Por ser este pai exemplar, amoroso, que bota os filhos acima de tudo. Sou grato ao nosso Senhor por ter me dado pais tão grandiosos e que em plena consciência lógica, não vi em nenhuma outra família. Obrigado meu pai, meu herói. Agradeço aos meus pais, que confiaram em mim e batalharam desde o meu nascimento para que eu pudesse chegar até aqui, me ensinando sobre a vida e guiando a construção da minha moral nos melhores alicerces.

À minha amada Iany Louise, que sempre esteve no meu lado, sempre me apoiou e não me deixou faltar nada em todos os âmbitos. Sou grato por ter te conhecido, e ainda mais, por estar com você. Te amo com todas as forças, obrigado por ser essa pessoa tão meiga e amorosa, és tudo e um pouco mais que eu sempre sonhei pra mim.

A minha tia, “Tica Rodrigues”, que me impulsionou e me ajudou nos primeiros momentos da minha permanência no curso, foi quem garantiu minha ida para estudar fora, sem ela nunca nem teria começado.

Aos meus tios, Nilzon Cesino, Jedson Cesino, Júnior Cesino, Josinaldo Cesino, Enilson Cesino e Nailson Cesino, que sempre me apoiaram e me ajudaram nesta jornada, sempre estando presentes e serem esses exemplos de tios e pessoas. À minha família como um todo: meus avós, tios, tias e primos, que me mostraram o significado de uma família unida e amorosa. Todos são muito importantes para mim, não poderia ter uma família melhor.

Ao meu querido Prof. Juliano Freitas, que sempre me ajudou, me aconselhou e me apoiou de todas as formas. Sou eternamente grato, por ter o senhor na minha vida, não só na acadêmica, como na social e evidentemente, na espiritual. O senhor foi e continua sendo como um anjo que me acolheu e que me fez crescer de todas as formas. Agradeço por ser muito mais que um professor, por ter acreditado em mim. Não somos da mesma família, mas saiba que o senhor constitui um membro da minha família.

A grande Prof. Júlia Beatriz, que sempre me ajudou desde quando comecei esta jornada, não tenho palavras para expressar todo o meu agradecimento. A senhora é muito mais que uma professora, foi um anjo que apareceu na minha vida. Sem seu apoio, talvez não estivesse aqui e agora, és um dos meus alicerces. Agradeço a sua família, em especial, a dona Magnólia, pessoa que me acolheu e sempre buscou me ajudar de todas as formas, tenho um carinho enorme.

Agradeço a minha orientadora e grandiosa, Profa. Dra. Luciana Scotti, pelo aprendizado, confiança, suporte, orientação e amizade. Tem sido uma honra trabalhar em conjunto com a senhora, meu eterno agradecimento pela cientista e inspiração que és em minha vida.

A todos do Laboratório de Quimioinformática, pelas discussões científicas, apoio e fortalecimento na pesquisa. Agradeço, especialmente, aos pesquisadores Natália, Mayara, Simone, Jéssica, Érika, Natan, além do pesquisador Alex France, pela grande ajuda e todo aprendizado repassado.

Ao professor Marcus Tullius Scotti, pelo esclarecimento de dúvidas e por abrir as portas de seu laboratório para execução desta pesquisa.

À CnPq e à Universidade Federal da Paraíba pelo apoio científico e financeiro.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

## EPÍGRAFE

*“Obrigado minha Mãe, por tudo que eu sou.”*

*Francisca Costa*

## RESUMO

Os vírus de RNA tem sido uma grande problemática para todo o mundo, tendo em vista sua fácil replicação, mutação e transmissão. Especificamente, o influenza A é um vírus de rna responsável por um alto número de óbitos, seja atuando isoladamente ou agravando diversos quadros patológicos já existentes. Diante do quadro que tal vírus trás para a população e levando em consideração os poucos fármacos existentes para o tratamento do influenza A, é de suma importância a busca por candidatos a fármacos que não apresentem tanto efeitos colaterais e que sejam eficazes para o mesmo, em alternativa aos atuais tratamentos, salientando a importância de candidatos que atuem em múltiplos alvos (multitarget), buscando através destes, fármacos mais eficazes e promissores, uma vez que, dentre as principais vantagens, tem a capacidade de bloquear mais de um alvo, obtendo melhor eficácia e perfis de resistência. Deste modo, a química medicinal, através da utilização de ferramentas computacionais, tem sido essencial no processo de planejamento de fármacos, uma vez que possibilita a otimização do tempo e de custos operacionais. Associado a isto, o crescente número de casos decorrentes de doenças endêmicas, como o influenza A, vem estimulando os órgãos de fomento a investir em pesquisas para o desenvolvimento de moléculas que possam ser utilizadas no tratamento de tais. Com isso, o objetivo do presente trabalho foi propor moléculas naturais com potencial atividade anti-influenza A, predita por uma análise de consenso de três modelos de predição de atividade biológica e analisadas por uma triagem virtual hibrida. Frente a isto, o banco de dados de 986 moléculas naturais, foram triadas, através primeiramente, da análise por consenso oriundas de 3 modelos de predição de atividade biológica, o que resultou em apenas 36 moléculas preditas como ativas para o influenza A e com ótimo percentual de confiabilidade, resultante da construção dos 3 modelos preditivos. Posteriormente, tais moléculas foram submetidas a análises do perfil de absorção e biodisponibilidade oral, riscos de toxicidade e metabolismo frente ao CYP 450. Neste sentido, as moléculas resultantes de tais etapas totalizaram em 7 e diante de seus promissores resultados, estas moléculas foram submetidas a docagem molecular, frente as 4 principais proteínas envolvidas no ciclo de replicação do influenza A, a citar: canal M2, hemaglutinina, neuraminidase e RNA polimerase. Com tais resultados, no geral todas as moléculas apresentaram ótimas interações frente as 4 principais proteínas do influenza A, com destaque, para a molécula HER03, que demonstrou ser um potencial candidato a fármaco atuante em múltiplos alvos, tendo em vista resultantes que apresentou scores de energias de interação superior a todos os fármacos controle e para todas as 4 proteínas envolvidas no ciclo de replicação do influenza A. Diante de tais resultados, foi demonstrado que das 986 moléculas naturais, 7 apresentaram potenciais para serem candidatos a fármacos frente ao influenza A, com destaque para a molécula HER03, que apresentou promissores e potenciais resultados para ser uma candidata a fármaco anti influenza A atuante em múltiplos alvos.

**PALAVRAS-CHAVE:** Influenza A, Multitarget, Química medicinal, Docagem molecular.

## ABSTRACT

RNA viruses have been a major problem for the entire world, given their easy replication, mutation and transmission. Specifically, influenza A is an RNA virus responsible for a high number of deaths, whether acting alone or aggravating several existing pathological conditions. Given the situation that this virus brings to the population and taking into account the few drugs used for the treatment of influenza A, it is extremely important to search for drug candidates that do not have as many side effects and that are effective for it, in alternative to current treatments, stressing the importance of candidates that act on multiple targets (multitarget), seeking through these, more effective and promising drugs, since, among the main advantages, it has the ability to block more than one target, obtaining better effectiveness and strength profiles. Thus, medicinal chemistry, through the use of computational tools, has been essential in the drug planning process, as it allows the optimization of time and operating costs. Associated with this, the growing number of cases resulting from endemic diseases, such as influenza A, has encouraged funding agencies to invest in research for the development of molecules that could be used in the treatment of such diseases. Thus, the aim of the present work was to propose natural molecules with potential anti-influenza A activity, predicted by a consensus analysis of three biological activity prediction models and analyzed by a hybrid virtual screening. In view of this, the database of 986 natural molecules were screened, first, through consensus analysis from 3 predictive models of biological activity, which resulted in only 36 molecules predicted to be active for influenza A and with excellent percentage of reliability, resulting from the construction of the 3 predictive models. Subsequently, these molecules were subjected to analyzes of the profile of absorption and oral bioavailability, risks of toxicity and metabolism against CYP 450. Molecular docking, against the 4 main proteins involved in the replication cycle of influenza A, namely: M2 channel, hemagglutinin, neuraminidase and RNA polymerase. With such results, in general all molecules showed excellent interactions against the 4 main proteins of influenza A, with emphasis on the HER03 molecule, which proved to be a potential drug candidate acting on multiple targets, considering the results that presented scores of interaction energies higher than for all control drugs and for all 4 proteins involved in the influenza A replication cycle. In view of these results, it was demonstrated that of the 986 natural molecules, 7 had potential to be drug candidates against influenza A, highlighting the HER03 molecule, which showed promising and potential results to be a candidate for an anti-influenza A drug acting on multiple targets.

**KEYWORDS:** Influenza A, Multitarget, Medicinal chemistry, Molecular docking.

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	21
<b>2 FUNDAMENTAÇÃO TEÓRICA</b> .....	23
<b>2.1 Vírus de RNA</b> .....	23
<b>2.2 Influenza – Aspectos gerais</b> .....	25
<b>2.3 Estrutura e organização viral do influenza</b> .....	28
<b>2.4 Ciclo viral</b> .....	29
<b>2.5 Química medicinal</b> .....	31
<b>2.6 Químioinformática</b> .....	32
<b>2.7 Plantas medicinais</b> .....	35
<b>2.7 Metabólitos secundários</b> .....	36
<b>3 OBJETIVOS</b> .....	41
<b>3.1 Objetivo geral</b> .....	41
<b>3.2 Objetivos específicos</b> .....	41
<b>4 METODOLOGIA</b> .....	42
<b>4.1 Conjunto e tratamento de dados das moléculas ativas</b> .....	42
<b>4.2 Conjunto e tratamento de dados das moléculas candidatas</b> .....	43
<b>4.3 Otimização da arquitetura molecular</b> .....	43
<b>4.4 Descritores moleculares</b> .....	43
<b>4.5 Modelo de predição de atividade biológica – mPRECT</b> .....	45
<b>4.5.1 KNIME Analytics Platforms 3.6</b> .....	45
<b>4.5.2 Random Forest</b> .....	46
<b>4.5.3 Wecka Predictor 3.7</b> .....	47
<b>4.5.4 Curva ROC</b> .....	47

<b>4.6 Regra dos Cinco de Lipinski e %ABS.....</b>	<b>48</b>
<b>4.7 Predição dos riscos de toxicidade.....</b>	<b>48</b>
<b>4.8 Metabolismo.....</b>	<b>49</b>
<b>4.9 Docking Molecular.....</b>	<b>50</b>
<b>5 RESULTADOS E DISCUSSÃO.....</b>	<b>51</b>
<b>5.1 Modelo de predição de atividade biológica.....</b>	<b>51</b>
<b>5.2 Análise da regra de Lipinski, TPSA e Taxa de absorção (%ABS).....</b>	<b>57</b>
<b>5.3 Toxicidade.....</b>	<b>64</b>
<b>5.4 Metabolização no fígado.....</b>	<b>71</b>
<b>5.5 Docking molecular.....</b>	<b>78</b>
<i>5.5.1 Docagem molecular do canal M2.....</i>	<i>80</i>
<i>5.5.2 Docagem molecular da hemaglutinina.....</i>	<i>83</i>
<i>5.5.3 Docagem molecular da neuraminidase.....</i>	<i>85</i>
<i>5.5.4 Docagem molecular da polimerase.....</i>	<i>87</i>
<b>6 CONCLUSÃO.....</b>	<b>90</b>

## **REFERÊNCIAS**

## LISTA DE ABREVIATURAS, SIGLAS E FÓRMULAS

SARS-CoV-2	Síndrome Respiratória Aguda Grave do Corona vírus 2
CADD	Computational Aided-Drug Design
OMS	Organização Mundial da Saúde
CDC	Centers for Disease Control and Prevention
dsRNA	Vírus RNA de fita dupla
ssRNA	Vírus RNA de fita simples
RdRp	Polimerase de RNA RNA-dependente
FDA	Food and Drugs Administration
nM	Nanomolar
pH	Potencial Hidrogeniônico
RNPs	Ribonucleoproteína
HA	Hemaglutinina
NA	Neuraminidase
PB1	proteína básica 1
PB2	proteína básica 2
M1	Proteína da matriz M1
M2	Proteína do canal iônico M2
NP	Proteína nuclear
NS1	Proteína não estrutural 1
NS2	Proteína não estrutural 2
PA	Polimerase
RNA <sub>m</sub>	RNA mensageiro
RNA <sub>v</sub>	RNA viral
QSTR	Quantitative Structure-Toxicity Relationship
QSAR	Quantitative Structure-Activity Relationship
ADME-Tox	Absorção, distribuição, metabolismo e toxicidade
LogP	Coefficiente de partição
QSPR	Quantitative Structure-Property Relationship
IA	Inteligência artificial
VS	Triagem Virtual
EC <sub>50</sub>	Concentração necessária para ocasionar metade do efeito máximo

mm	Milimol
1D	Unidimensional
2D	Bidimensional
3D	Tridimensional
pEC <sub>50</sub>	Logaritmo negativo de EC <sub>50</sub>
ROC	Receiver Operating Characteristic
AUC	Área sob a curva
DOI	Digital Object Identifier
nDLH	Número de doadores de ligação de Hidrogênio
nALH	Número de aceptores de ligação de Hidrogênio
TPSA	Topological Polar Surface Area
RMSD	Root Mean Square Deviation
MCC	Coefficiente de correlação de Matthews
ACC	Acurácia
VPP	Verdadeiros positivos
VNP	Verdadeiros negativos
ABS %	Percentual de absorção
PM	Peso molecular
nV	Números de violações
Å	ångström
MUT	Mutagenicidade
CAR	Carcinogenicidade
ESR	Efeito no sistema reprodutor
IRR	Irritabilidade no tecido
TOX	Toxicidade geral
CYP450	Citocromo P450
nMET	Número de metabólitos formados para cada molécula
ID	ID dos metabólitos
ABD%	Percentual de abundância dos metabólitos
PDB	Protein Data Bank
Kcal/mol	Quilocalorias por mol
HIS	Histidina
ALA	Alanina

LEU	Leucina
GLY	Glicina
ARG	Arginina
TRP	Triptofano
ASG	Asparargina
ILE	Isoleucina
VAL	Valina
THR	Treonina
GLU	Glutamato
LYS	Lisina
ASP	Aspartato
MET	Metionina

## LISTA DE FIGURAS

Figura 1 - Representação esquemática da classificação das viroses de rna.....	24
Figura 2 - Arquitetura molecular do ácido siálico.....	27
Figura 3 - Estrutura viral do influenza A.....	28
Figura 4 - Arquitetura molecular do ácido siálico.....	30
Figura 5 - Exemplo geral dos fatores que geram um modelo de predição de atividade biológica.....	34
Figura 6 - Rota de biossíntese geral dos terpenos.....	37
Figura 7 - Rota de biossíntese geral dos compostos fenólicos.....	37
Figura 8 - Rota de biossíntese geral dos compostos nitrogenados.....	38
Figura 9 - Arquitetura molecular do alcalóide crinina.....	39
Figura 10 - Representação gráfica do desempenho do modelo de predição de atividade dos descritores do Dragon.....	53
Figura 11 - Representação gráfica do desempenho do modelo de predição de atividade dos descritores do CDK.....	53
Figura 12 - Representação gráfica do desempenho do modelo de predição de atividade dos descritores do VolSurf.....	54
Figura 13 - Gráfico de abundância dos 37 metabólitos oriundos da molécula HER01.....	76
Figura 14 - Estruturas em 2D de alguns dos metabólitos da molécula HERB01.....	77
Figura 15 - Estrutura 2D das 7 moléculas restantes + 5 fármacos controles.....	79
Figura 16 - Interações da HER03 com os resíduos de aminoácidos da M2.....	82
Figura 17 - Interações da HER04 com os resíduos de aminoácidos da HA.....	85
Figura 18 - Interações da HER03 com os resíduos de aminoácidos da NA.....	87
Figura 19 - Interações da HER03 com os resíduos de aminoácidos da PA.....	89

## LISTA DE GRÁFICOS

Gráfico 1 - Percentual de abundância dos 7 metabólitos da molécula HER01.....	76
Gráfico 2 - Resultados do docking molecular frente ao canal M2 do influenza A.....	81
Gráfico 3 - Resultados do docking molecular frente a hemaglutinina do influenza A.....	83
Gráfico 4 - Resultados do docking molecular frente a polimerase do influenza A.....	88

## LISTA DE TABELAS

Tabela 1 - Resultados dos parâmetros estatísticos de confiabilidade do modelo de predição utilizando os descritores do <i>Dragon</i> .....	51
Tabela 2 - Resultados dos parâmetros estatísticos de confiabilidade do modelo de predição utilizando os descritores do <i>CDK</i> .....	52
Tabela 3 - Resultados dos parâmetros estatísticos de confiabilidade do modelo de predição utilizando os descritores do <i>VolSurf</i> .....	52
Tabela 4 - Arquitetura molecular em 2D das 36 moléculas preditas como ativas.....	56
Tabela 5 - Resultado dos parâmetros da regra de Lipinski, valores de TPSA e percentual de absorção das moléculas.....	58
Tabela 6 - Resultado dos parâmetros da regra de Lipinski e percentual de absorção dos fármacos controle.....	63
Tabela 7 - Resultado das análises de toxicidade.....	65
Tabela 8 - Arquitetura molecular em 2D das 24 moléculas restantes. ....	70
Tabela 9 - Principais dados das moléculas frente ao metabolismo do CYP450.....	72
Tabela 10 - Arquitetura molecular das candidatas que não apresentaram metabólitos secundários abundantes tóxicos.....	78
Tabela 11 - Valores de energia (MolDockScore) das moléculas.....	80
Tabela 12 - Interações das moléculas com os resíduos de aminoácidos do M2.....	81
Tabela 13 - Interações das moléculas com os resíduos de aminoácidos da hemaglutinina.....	84
Tabela 14 - Interações das moléculas com os resíduos de aminoácidos da neuraminidase.....	86
Tabela 15 - Interações das moléculas com os resíduos de aminoácidos da polimerase.....	88

## 1. INTRODUÇÃO

As viroses de RNA tem sido uma grande problemática para todo o mundo, tendo em vista sua fácil replicação, mutação e transmissão (HAYLES, 2021). Especificamente, as duas primeiras décadas do século XXI vem sendo marcadas por pandemias de vírus respiratórios, a primeira declarada em 2009, através do vírus influenza A (H1N1) e atualmente, a segunda pandemia, declarada no dia 11 de março de 2020, pelo novo coronavírus 2019 (COVID-19), o SARS-CoV-2 (BEDFORD et al., 2019; WHO, 2020).

Diante do cenário atual da pandemia ocasionada pelo SARS-CoV-2 (COVID-19), faz-se necessário compactuar de que a literatura já relata estudos que demonstram que o influenza A tem uma capacidade “única” de agravar a infecção por SARS-CoV-2 e, portanto, a prevenção da infecção ocasionada pelo influenza A é de grande importância durante a pandemia de COVID-19, tendo em vista que a temporada do influenza A está se fundindo com a pandemia atual, potencialmente trazendo mais desafios e representando uma ameaça maior à saúde pública (BAI et al., 2021; HASHEMI et al., 2021).

O vírus Influenza A é um patógeno respiratório humano que causa infecções sazonais e endêmicas, além de pandemias periódicas (KRISCHUNS et al., 2021; FERNANDA & ANDRADE, 2019). Como um importante patógeno viral humano, a infecção pelo vírus influenza A pode provocar sérias complicações respiratórias, que se tornam mais importantes em grupos de risco, podendo levar esses indivíduos a óbito (LIU et al., 2021; CHENG et al., 2021).

A epidemia sazonal de tal vírus causa centenas de milhares de mortes em todo o mundo a cada ano, pontuada por pandemias ocasionais com mortes que chegam a milhões (ASADI et al., 2020; KRISCHUNS et al., 2021). Dito isto, o processo patológico provocado pela infecção do vírus influenza pode levar a quadros que variam de infecções assintomáticas a infecções respiratórias graves e que tal gravidade está intimamente relacionada com a interação com o hospedeiro, resposta imune e a virulência do patógeno (BAI et al., 2021).

Dessa forma, grupos de risco como pacientes idosos, crianças menores de dois anos, asmáticos e aqueles que apresentam patologias crônicas no trato respiratório, estão mais susceptíveis a complicações por influenza e que muitas vezes, vão a óbito por não possuírem fármacos antivirais eficazes (WANG et al., 2020; HARSHBARGER et al., 2021).

Um dos maiores desafios no controle do influenza A está atrelado a sua fácil replicação, transmissão, mutação e resistência aos medicamentos comumente utilizados, como por exemplo a amantadina e rimantadina e até mesmo, mais recentemente, o oseltamivir e zanamivir, em que

verifica-se altas taxas de resistência deste vírus para tais medicamentos, o que enfatiza ainda mais o planejamento e desenvolvimento de novos fármacos com atividades anti-influenza, que possam diminuir os mecanismos de resistência e possuir efetividade terapêutica (BISSO-ANDRADE, 2018; ROBLES et al., 2020).

Neste sentido, os fármacos com atividade em múltiplos alvos (do inglês *multitarget*), atuam como promissoras moléculas na busca e desenvolvimento de medicamentos mais eficazes e promissores, uma vez que, dentre as principais vantagens, tem a capacidade de bloquear mais de um alvo, obtêm melhor eficácia e perfis de resistência (NAMASIVAYAM et al., 2021; XU et al., 2021).

Com isso, a química farmacêutica dispõe de inúmeras ferramentas e métodos que agilizam e otimizam o processo da descoberta e desenvolvimento de novas moléculas atuantes em múltiplos alvos (TOROPOVA, 2017). O planejamento de tais moléculas através da utilização de ferramentas computacionais são de suma importância, tendo em vista que a utilização de técnicas computacionais, a citar pela triagem virtual híbrida, otimizam significativamente a busca por novos fármacos (TRIGUERO et al., 2016).

Adicionalmente, é possível citar a técnica do design de fármacos assistido por computador (CADD – *Computational Aided-Drug Design*), que reduz o custo, o tempo e o esforço necessário para desenvolver novos fármacos, acarretando assim, no aceleração da descoberta de novas moléculas com ação terapêutica (SALMAN et al., 2021; GURUNG et al., 2021).

Os estudos envolvendo as abordagens CADD têm sido aplicados no entendimento entre propriedades físico-químicas e biológicas, além disso, é possível conciliar os estudos de CADD com outras técnicas, coordenando e modelando os dados, norteando nas buscas de moléculas bioativas (YU & MACKERELL, 2017). Adicionalmente, o CADD possui dois tipos gerais de abordagens de desenho de fármacos: desenho de fármacos baseado em estrutura (*Structure-based*) e desenho de fármacos baseado em ligante (*Ligand-based*) (YU & MACKERELL, 2017; DAR et al., 2019).

Com isso, a realização de abordagens *in silico* constituem-se ferramentas valiosas e essenciais que contribuem para o sucesso do tratamento e para a erradicação de doenças, tendo em vista que, essas abordagens vem sendo utilizadas amplamente nos mais diversos campos da pesquisa, em que os estudos de triagem virtual, as abordagens baseadas em estrutura e ligante, têm sido aplicadas para identificar uma gama de moléculas que possam a vir a serem utilizadas na terapêutica, bem como para prever o modo de ligação de ligantes ativos e para estudar a dinâmica da proteína (LUISE & ROBAA, 2018; GURUNG et al., 2021).

## **2. FUNDAMENTAÇÃO TEÓRICA**

### **2.1 Vírus de RNA**

As doenças infecciosas são causadas por agentes biológicos como os vírus, bactérias e outros microrganismos patogênicos, os quais são as causas importantes de morbidade e mortalidade humana. Os microrganismos patogênicos podem ser propagados por meio da transmissão entre alimentos, fontes de água ou hospedeiros. Portanto, o diagnóstico por métodos confiáveis e precisos, além da disponibilidade de uma terapia medicamentosa eficaz são necessidade urgente para o manejo das doenças infecciosas (ZHU et al., 2021).

Segundo Klajman (2016), a relação entre os seres humanos e as doenças é tão antiga quanto à história humana e a despeito de toda a evolução tecnológica ocorrida, a população permanece fortemente atrelada ao mundo natural e aos seres microscópicos. A Organização Mundial da Saúde (OMS) (2020) (SCHUELER, 2020) conceitua o termo “pandemia” como a disseminação mundial de uma nova doença e a denominação passa a ser utilizada quando uma epidemia que afeta uma determinada região, se alastra por diferentes continentes com transmissão sustentada de pessoa para pessoa. Dessa forma, é necessário enfatizar que a gravidade da doença não é determinante e sim o seu poder de contágio, como também a sua proliferação geográfica (ARRAIS & SILVA, 2019).

Nessa perspectiva, faz-se necessário mencionar as definições de: epidemias e endemias. O termo epidemia é considerado quando ocorre o aumento do quantitativo de casos até um máximo de infecções sendo sucedido por uma diminuição dos mesmos (CDC, 2020; ROSENBERG, 2016). Já o conceito de endemia, tem como foco principal a regionalização, referindo-se a doenças que se encontram em uma determinada zona ou região de forma permanente durante um longo período (KALRA et al., 2015). Então, os conceitos citados diferem do termo pandemia, por esta ocorrer em todo o mundo ao mesmo tempo (ARRAIS & SILVA, 2019).

Segundo Senhoras (2020), epidemias fazem parte da realidade de um mundo cada vez mais globalizado, gerando uma série de consequências e vulnerabilidades biológicas aos Estados Nacionais, as quais eventualmente podem de forma rápida resultar em uma pandemia internacional, razão pela qual os fatores relacionados a conformação de agendas de cooperação internacional, transparência comunicacional e de respostas compartilhadas se tornam fatores essenciais para o sucesso do sistema de governança da Saúde Pública global, minimizando assim riscos epidemiológicos e consequências socioeconômicas.

Entre as pandemias já ocorridas no mundo, pode-se mencionar aquelas que são provocadas pelos vírus de RNA, ou seja, um vírus de RNA é um tipo de vírus que utiliza o RNA como o seu material genético. Neste sentido, a humanidade é frequentemente atingida pelos vírus, em parte devido aos grandes surtos das viroses de RNA das últimas décadas, a citar: Ebola, Zika, Influenza A e o Covid-19 (ARIS-BROSOU, PARENT, IBEH, 2019; FARIA et al., 2016; GIRE et al., 2014).

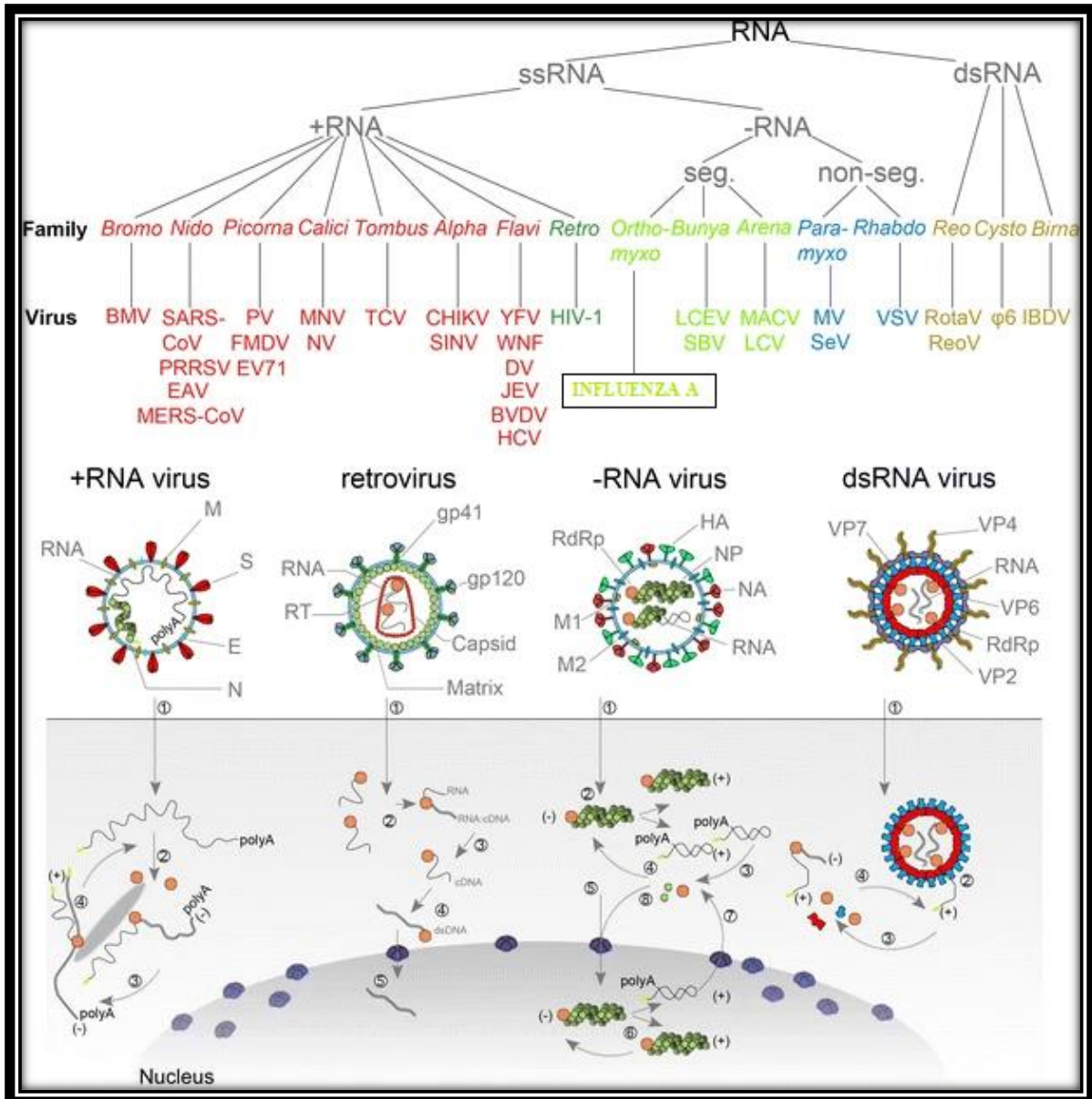
Tais vírus são conhecidos por terem algumas das taxas de mutação mais altas e diversas encontradas em qualquer replicador biológico (ARIS-BROSOU, PARENT, IBEH, 2019). Existem três classificações para vírus de RNA, sendo estas baseadas no genoma e no modo de replicação, as quais correspondem: vírus RNA de fita dupla (dsRNA); vírus RNA de fita simples (ssRNA) e retrovírus (AHLQUIST, 2006; DAS et al., 2021; FUKUHARA, 2019).

Os vírus de dsRNA contêm de uma a uma dúzia de moléculas de RNA diferentes, cada uma das quais codifica uma ou mais proteínas virais, conseqüentemente, este dsRNA roteia o desenvolvimento de um novo RNA viral (CUI et al., 2019; HISANO et al., 2018). Com relação aos vírus de RNA de fita simples (ssRNA), estes são caracterizados como de sentido positivo e de sentido negativo, ao qual o genoma de vírus ssRNA de sentido positivo é usado diretamente como mRNA, e os ribossomos hospedeiros traduzem este genoma em uma única proteína que é modificado pelo hospedeiro e proteínas virais para formar as várias proteínas necessárias para a replicação (COMAS-GARCIA, 2019; THÉBAUD et al., 2010; ZHANG et al., 2019). Um deles inclui RdRp (polimerase de RNA RNA-dependente), que copia o RNA viral para desenvolver uma cadeia dupla tipo replicativo (MAIDA et al., 2009; VENKATARAMAN et al., 2018).

O genoma de vírus ssRNA de sentido negativo deve ser copiado por uma replicase de RNA para formar RNA de sentido positivo (JAMIN & YABUKARSKI, 2017; LUO et al., 2020). A molécula de RNA de sentido positivo então atua como mRNA viral, e este mRNA é traduzido em proteínas por ribossomos hospedeiros. Para finalizar a classificação, os retrovírus apresentam um genoma de ssRNA, mas geralmente não são considerados vírus de RNA porque utilizam DNA intermediários para replicação (MILLER, 2001; TIAN et al., 2021).

Diante dessas considerações, a figura (1) representa melhor tais classificações e tipos para as viroses de RNA.

Figura 1. Representação esquemática da classificação das viroses de RNA.



Fonte: Adaptado de Velthuis (2014).

## 2.2 Influenza – Aspectos gerais

Os vírus influenza, de nome *Myxovirus influenzae*, são pertencentes à família *Orthomyxoviridae*, que são vírus de RNA peculiares, que contém um genoma RNA segmentado e fita negativa simples (HUTCHINSON, 2018; FERNANDA & ANDRADE, 2019; LIU ET AL., 2021). Os vírus como o influenza A, possuem a característica de codificar seus próprios capsídeos que envolvem seus genomas, sendo tal vírus composto por inúmeras moléculas de RNA que contém toda a informação genética necessária para a síntese de novos vírus (DAS et al., 2021; NOGUEIRA & PONCE, 2021).

Os vírus do influenza são classificados em três tipos: A, B e C. O tipo A infecta humanos e animais e está envolvido em episódios epidêmicos e pandêmicos. O tipo B infecta apenas humanos e está ligado a surtos moderados. Já o tipo C é mais estável e acomete humanos e suínos, causando doença subclínica, sem potencialidade epidêmica (COSTA & MERCHAN-HAMANN, 2016).

Especificamente, o influenza A causa diversas manifestações clínicas que variam de quadros leves, graves e até ao óbito (WANG et al., 2021). O potencial pandêmico fez com que o influenza A se destacasse das demais doenças infecciosas, dentre as pandemias oriundas de tal vírus, percebe-se que ocorreram em vários períodos da história e compreendem respectivamente: a gripe espanhola de 1918 a 1920; a gripe asiática ocorrida entre 1957 a 1958; a Gripe de Hong Kong de 1968 a 1969; a gripe russa de se estendeu de 1977 a 1978; a gripe aviária ocorrida entre os anos de 2003 a 2004 e por fim, a gripe H1N1 prevalente no ano de 2009 (COSTA & MERCHAN-HAMANN, 2016).

Diante deste cenário pandêmico, vale destacar que as duas primeiras décadas do século XXI vem sendo marcadas por pandemias de vírus respiratórios, sendo a primeira declarada em 2009, através do vírus influenza A (H1N1) e atualmente, a segunda, declarada no dia 11 de março de 2020, pelo novo coronavírus 2019 (COVID-19), o SARS-CoV-2 (BEDFORD et al., 2019; WHO, 2020).

Neste sentido, o influenza A causa epidemias anuais e pandemias ocasionais de infecções do trato respiratório que produzem um amplo espectro de gravidade clínica da doença em humanos (FLERLAGE et al., 2021; HARSHBARGER et al., 2021). De um modo geral, esse vírus causa vários graus de doença respiratória, uma vez que replicam seus genomas dentro do núcleo da célula hospedeira e podem infectar uma variedade de células no trato respiratório, incluindo células epiteliais ciliadas, células alveolares tipo I e II e células do sistema imunológico (FLERLAGE et al., 2021).

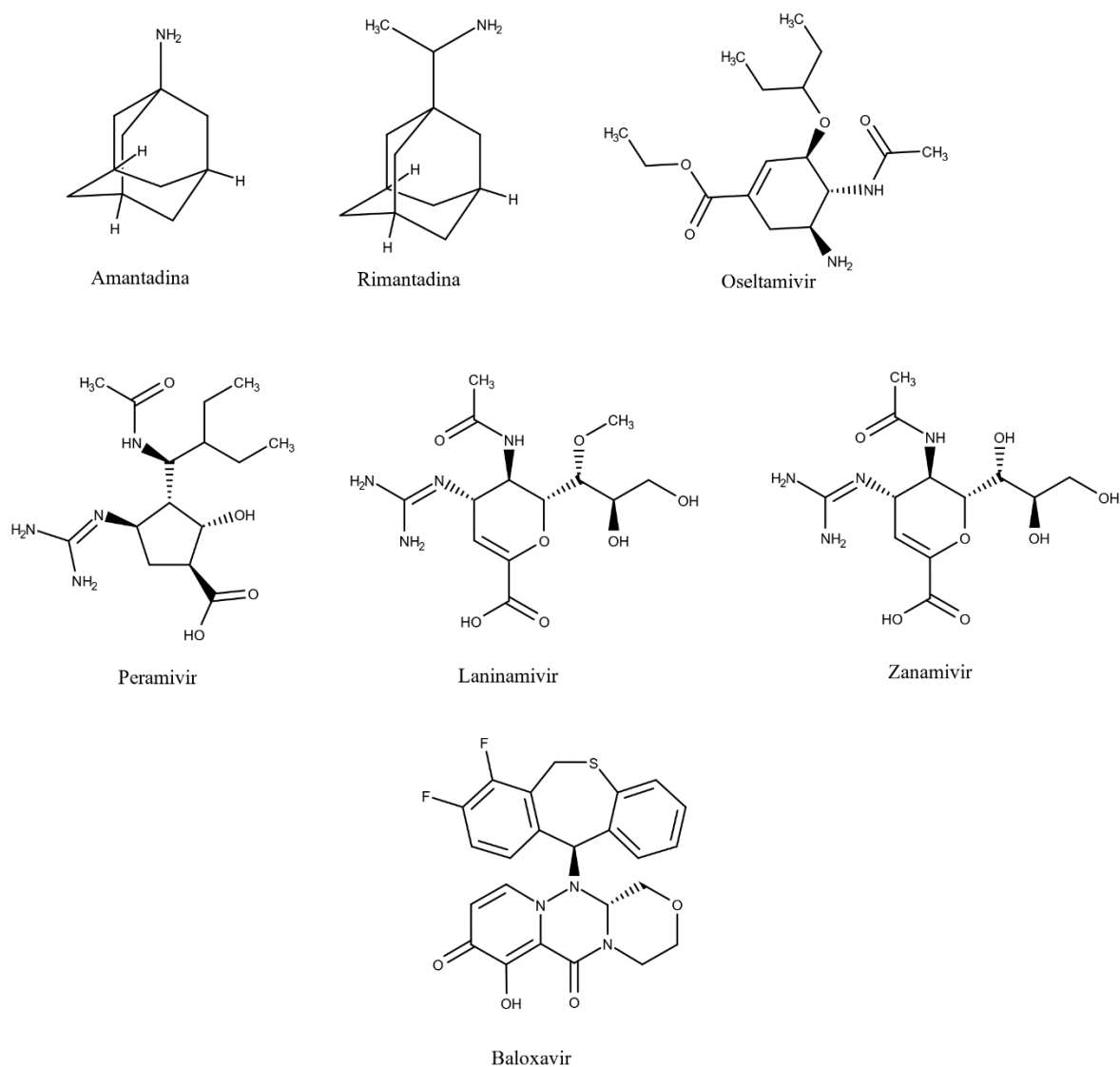
O efeito final nas células infectadas depende do tipo específico de célula (CHEN et al., 2021; PFLUG et al., 2017). O tropismo do vírus ocorre devido à capacidade dos vírus influenza de se ligarem a diferentes isoformas de ácido siálico (Figura 1) presentes nas células hospedeiras (CLEMENTI et al., 2021; MATROSOVICH et al., 2004).

Além da amantadina e rimantadina, qual atuam inibindo o canal M2 do vírus e assim, impedem o influxo de prótons que propiciam o pH (potencial hidrogeniônico) adequado para a replicação viral, atualmente, para o tratamento do influenza A, as classes de antivirais que possuem aprovação para uso, recomendada pelo CDC (Centers for Disease Control and Prevention) e aprovada pelo FDA (Food and Drugs Administration) são: Oseltamivir,

Zanamivir, Peramivir e Laninamivir, que agem inibindo a neuraminidase do vírus, importante para o seu ciclo replicativo (KOONIN & PATEL, 2018).

Tanto o oseltamivir, quanto o zanamivir têm aprovação para uso global, contudo, o peramivir somente é aprovado na China, Japão, Coreia do Sul e Estados Unidos e o laninamivir unicamente no Japão (ISON, 2017; DAVIDSON, 2018). Adicionalmente, tem-se também o baloxavir, aprovado nos Estados Unidos e no Japão em 2018, atuando na inibição da polimerase dependente de RNA do vírus, impedindo a transcrição do genoma viral (ABED, SAIM-MAMOUN, BOIVIN, 2021; SÁNCHEZ-ESPINOZA et al., 2020). A figura (2) demonstra a arquitetura molecular em 2D desses fármacos.

Figura 2. Arquitetura molecular dos fármacos utilizados para o tratamento do influenza A.



### 2.3 Estrutura e organização viral do influenza

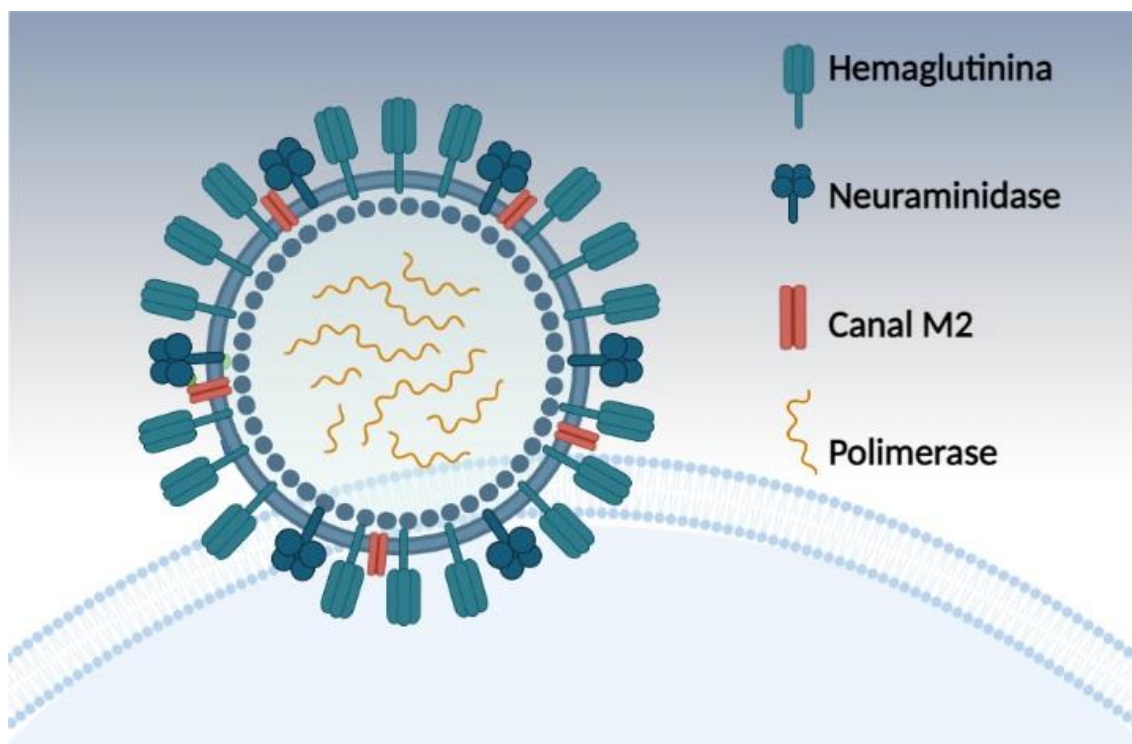
Os vírus nada mais são do que partículas de dimensões subcelulares e estas partículas virais, ou viriões, conseguem entrar na célula hospedeira, iniciando o ciclo de replicação viral, que consiste na síntese de todas as moléculas necessárias para a geração dos novos viriões descendentes (NOGUEIRA & PONCE, 2021). Especificamente, o Influenza é um vírus pleomórfico, ou seja, não apresenta uma estrutura fixa, apesar de poderem ser esféricas, com um diâmetro de 100 nM, ou filamentosas, de 80 a 120 nM (HUTCHINSON, 2018; NOGUEIRA & PONCE, 2021).

O genoma do vírus influenza A consiste em oito segmentos de RNA de cadeia simples, de sentido negativo, que são encapsidados como complexos individuais de ribonucleoproteína em forma de bastonete (RNPs). Cada RNP contém um RNA viral, uma polimerase viral e várias cópias da nucleoproteína viral (NP) (ZHENG & TAO, 2013; FERNANDA & ANDRADE, 2019).

Cada segmento de RNA é compactado juntamente com proteínas virais (de suma importância para o processo replicativo), que em conjuntos formam as nucleocápsides estruturalmente independentes. Especificamente, o genoma viral tem aproximadamente 13kb e codifica 13 proteínas: as glicoproteínas de superfície hemaglutinina (HA) e neuraminidase (NA); a proteína da matriz M1 (M1) e proteína do canal iônico M2 (M2); a proteína nuclear (NP); o complexo de RNA polimerase (PB1, PB2, PA) e proteínas não estruturais (NS1, NS2) (BAUMEISTER & PONTORIERO, 2017; FERNANDA & ANDRADE, 2019).

Na superfície do envelope viral estão inseridas as glicoproteínas hemaglutinina, neuraminidase e a proteína do canal M2 (BAUMEISTER & PONTORIERO, 2017). Internamente ao envelope viral encontra-se um arranjo de proteínas M1 e as outras proteínas virais não estruturais e enzimas e o material genético fragmentado em 8 complexos de ribonucleoproteínas (BAUMEISTER & PONTORIERO, 2017). Dentre essas proteínas virais, pode-se destacar quatro principais para o desencadeamento e manutenção do processo replicativo: as proteínas de superfície hemaglutinina (HA) e neuraminidase (NA), proteína do canal iônico M2 e a polimerase (PA), conforme demonstra a figura (3).

Figura 3. Estrutura viral do influenza A.



Fonte: Próprio autor, 2021.

## 2.4 Ciclo viral

A princípio, o ciclo de infecção viral começa com a ligação do vírus aos receptores na membrana da célula hospedeira (NIE et al., 2021). Para o vírus influenza A, duas glicoproteínas foram relacionados ao comportamento de ligação na membrana celular, a hemaglutinina e a neuraminidase. O vírus se adsorve na membrana celular do hospedeiro utilizando a HA, que se liga aos resíduos de ácido siálico presente nos glicoconjugados da superfície da célula, desencadeando a entrada viral (DOU et al., 2018; NIE et al., 2021).

Posteriormente a entrada do vírus na célula, há a participação da proteína M2 que forma o canal iônico, permitindo o influxo de prótons para a partícula viral, promovendo diminuição do pH, que faz com que a matriz M1 se desintegre e se solte das RNPs, o envelope viral se dissocia e as RNPs vão para o núcleo da célula (DOU et al., 2018).

No núcleo da célula há a polimerase viral, uma enzima que apresenta atividade RNA polimerase dependente de RNA e basicamente, transcreve o genoma do RNA viral no núcleo da célula infectada (KRISCHUNS et al., 2021). Além disso, a mesma é essencial para o ciclo replicativo do vírus, pois permite a síntese de novas moléculas de RNA a partir do genoma viral, função que é exclusiva dos genomas de vírus de RNA (KRISCHUNS et al., 2021; NOGUEIRA & PONCES, 2021).

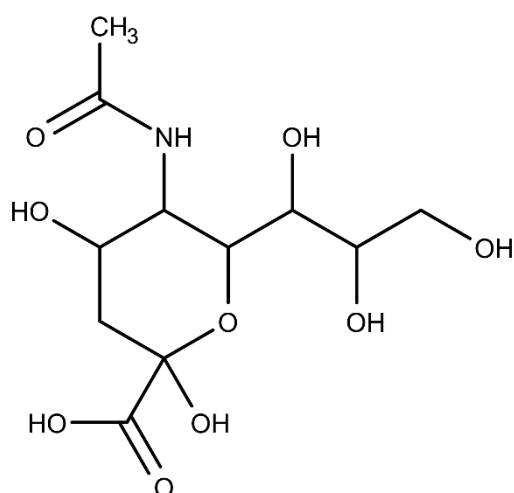
Neste sentido, o complexo da polimerase é formado por um heterotrímero que inclui a proteína ácida, proteína básica 1 (PB1) e proteína básica 2 (PB2), complexo este, funcional para a transcrição e replicação viral, tendo em vista que participa da clivagem de RNA mensageiro (RNAm) do hospedeiro e transcreve o RNA<sub>v</sub> (-) (RNA viral negativo) em RNAm (+), havendo assim a formação de novos vírions (ZHENG & TAO, 2013; BAUMEISTER & PONTORIERO, 2017).

Quando os vírions formados estão prontos para sair da célula e infectar as demais, há então a atuação da neuraminidase. Para que ocorra a propagação da infecção virótica, se faz necessário a quebra da ligação do vírus com a membrana celular, hemaglutinina/ácido siálico, o que é efetivamente feito pela enzima neuraminidase viral. A neuraminidase atua clivando o ácido siálico da superfície da célula, ocasionando a liberação dos vírions descendentes (DOU et al., 2018).

Ou seja, o papel da neuraminidase é de suma importância para que os vírions não se liguem ao ácido siálico da mesma célula que estão saindo, tendo em vista que os mesmos são compostos pela hemaglutinina e tal enzima, já possui afinidade pelo ácido siálico, e assim, a neuraminidase cliva o ácido siálico para que então os vírions possam sair da célula, impedindo assim, uma reinfeção e com isso, os vírions darão continuidade ao ciclo viral, infectando as demais células, através da ligação com o ácido siálico de tais (DOU et al., 2018; FERNANDA & ANDRADE, 2019).

O conhecimento da base molecular da infecção virótica levou à pesquisa de substâncias que pudessem competir com o ácido siálico (Figura 4) pela ligação no sítio ativo da neuraminidase. Assim, foi sintetizado o "Oseltamivir e Zanamivir", fármacos que se ligam mais efetivamente a neuraminidase que o próprio ácido siálico, impedindo a continuidade do processo de infecção viral.

Figura 4. Arquitetura molecular do ácido siálico.



## Ácido Siálico

Fonte: Próprio autor, 2021.

### 2.5 Química medicinal

Na posteridade, o desenvolvimento de moléculas ativas com fins terapêuticos deve tornar-se cada vez mais maior, a fim de manter a esperança na luta contra os agentes patológicos, com a produção de fármacos mais seguros e eficazes, tendo em vista que, os novos ataques de microrganismos patogênicos têm potencializado o surgimento de novas doenças em todo o mundo (BEZERRA, OLIVEIRA et al., 2018; QUEIROZ, FREIRE, 2018).

É notório o grande investimento que as indústrias farmacêuticas aplicam na busca de novas alternativas que auxiliem no desenvolvimento de novos fármacos (SILVA, 2020). O planejamento e desenvolvimento de novos fármacos é um processo de grande complexidade, o qual requer tempo e investimento de diferentes setores como, por exemplo, do setor industrial (OLIVEIRA et al., 2018). As indústrias evoluíram bastante e continuam crescendo a cada dia, principalmente devido os avanços tecnológicos, bem como no campo de pesquisas farmacêuticas, que tem passado por diversas aprimorações, decorrentes dos avanços tecnológicos (GUERRA, 2019).

Neste sentido, um grande número de tecnologias computacionais tem sido aplicado no planejamento e desenvolvimento de novos candidatos a fármaco, e uma das áreas que tem se destacado é a Química Medicinal, que com ajuda de técnicas de triagem virtual (virtual screening) e da modelagem molecular, tem direcionado a síntese de diversos compostos farmacologicamente ativos e ótimos candidatos a fármacos (GOMES, 2019).

O avanço da tecnologia computacional e ao mesmo tempo a concepção e entendimento dos alvos patológicos, seus sítios ativos, catalíticos e de técnicas moleculares, estimularam o surgimento de modelos preditivos que norteiam a necessidade da realização de experimentos *in vitro* e *in vivo* (SILVAb, 2020). Neste sentido, as abordagens *in silico* tem sido de vital importância nas pesquisas de desenvolvimento de fármacos, sendo realizadas com o auxílio de métodos computacionais e matemáticos (VAVOUGIOS et al., 2018).

A química medicinal é uma das áreas que visa a busca de novos fármacos através da conciliação de áreas como a farmácia, química e quimioinformática (MASCARENHAS et al., 2019). A química medicinal baseia-se em modelos como a Quantitative Structure-Toxicity Relationship (QSTR, tradução: Relação Quantitativa Estrutura-Toxicidade) e Quantitative Structure-Activity Relationship (QSAR, tradução: Relação Quantitativa Estrutura-Atividade) para estipular uma relação matemática confiável entre propriedades toxicológicas, físico-químicas e farmacológicas de um determinado composto (MASCARENHAS et al., 2019; ALVES et al., 2018).

As abordagens *in silico* são vitais em pesquisas que buscam selecionar promissores candidatos a fármacos, uma vez que avaliam suas características físico-químicas e seu potencial terapêutico (AMIN et al., 2021; SYDOW et al., 2019). Essas abordagens tem sido utilizadas para reduzir os testes clínicos, tendo em vista que introduzem conhecimento e racionalidade a cada etapa do processo de descoberta de um novo composto terapêutico, embasando-se em modelos biorreguladores humanos para gerar informações quanto às características toxicológicas, farmacocinéticas e farmacodinâmicas dos alvos analisados, ou seja, características relacionadas a absorção, distribuição, metabolismo e toxicidade (ADME-Tox) (FOWLER et al., 2017; ALVES et al., 2018).

Com isso, estas abordagens contribuem na análise e predição de diversas propriedades, o que permite entender como elas influenciam na ação farmacológica e como podem ser moduladas para evitar a eliminação de bons candidatos devido a características farmacodinâmicas, farmacocinéticas ou toxicológicas desfavoráveis (FOWLER et al., 2017).

Diversas informações que são vitais durante a análise de um candidato a fármaco podem ser obtidas utilizando os métodos *in silico*, a citar: percentual de absorção teórica por via oral, massa molecular, parâmetros toxicológicos (concentração necessária para matar 50% dos organismos, mutagenicidade, imunotoxicidade) coeficiente de partição (logP), dentre outros (MASCARENHAS et al., 2019; MEDEIROS et al., 2020).

## 2.6 Quimioinformática

A quimioinformática é uma área que envolve a química e a tecnologia computacional com o intuito de resolver problemas relacionados a química das moléculas. Essas técnicas são utilizadas intensamente pelas indústrias farmacêuticas no processo de descoberta de fármacos, principalmente na fase de planejamento racional. Esses métodos podem também ser utilizados de várias formas em indústrias químicas e centros acadêmicos. (SILVAb, 2020).

O termo quimioinformática foi descrito por Frank Brown em 1998, definindo-a como “mistura de recursos de informação para transformar dados em informação e informação em conhecimento, no intuito de tomar decisões melhores e mais rápidas na área de identificação e otimização de compostos líderes” (ALVES et al., 2018; BRONW, 1998). A aplicação mais clássica da quimioinformática corresponde no recurso à mecânica molecular ou à mecânica quântica para o cálculo de propriedades termodinâmicas de moléculas e reações (SILVAb, 2020).

Uma gama de programas tem sido aplicados no design de fármacos, uma vez que são utilizados para os mais diversos fins, como por exemplo para problemas de otimização, sendo primordiais desde a análise conformacional até a seleção de descritores moleculares mais adequados ao estabelecimento de QSAR. A análise para extrair conhecimentos relevantes destas bases de dados envolve algoritmos e software para alinhamento de sequências, pesquisa de semelhanças, estimativa de árvores filogenéticas, previsão estrutural e inferência funcional (ALVES et al., 2018).

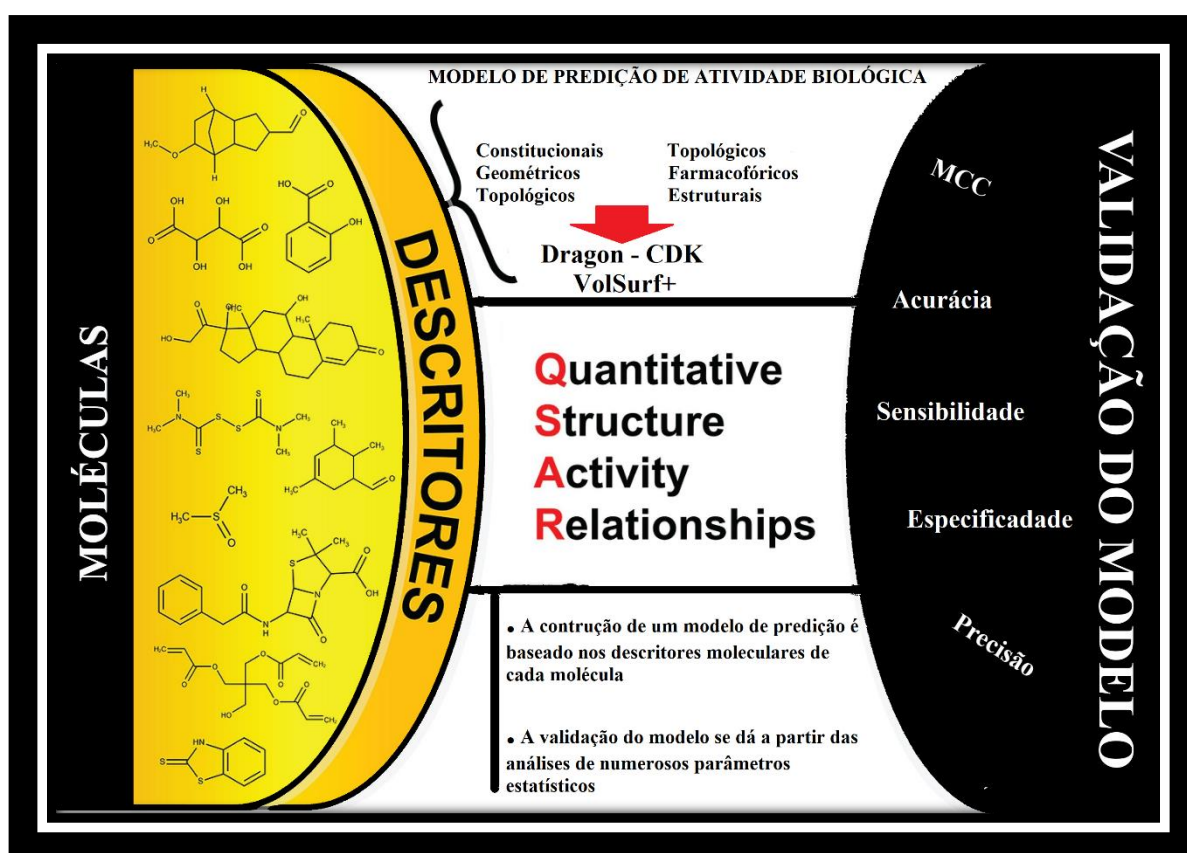
Os grandes avanços científicos dos últimos anos, levaram as indústrias farmacêuticas a procurarem estratégias de mecanização e robotização de vários processos ao longo da cadeia produtiva. A aplicação em pesquisa e desenvolvimento de fármacos é recente e tem sido explorada para prever as propriedades moleculares (QSPR - Quantitative Structure-Property Relationship) e atividade farmacológica (QSAR), análise de rotas sintéticas, e planejamento de novos compostos (COLEY et al., 2017). Além dessas abordagens, outras são significativamente utilizadas, tais como: inteligência artificial (IA), triagem virtual (VS), reaproveitamento de drogas, dentre outras (LINHARES et a., 2021).

De modo sucinto, um modelo de QSAR relaciona a estrutura química com a propriedade biológica de uma determinada molécula, através do uso de vários métodos quimioinformáticos e estatísticos, a fim de determinar as propriedades biológicas baseando-se em sua estrutura e propriedades químicas (ALVES et al., 2018). Tendo como base tais conhecimentos, a presente pesquisa associou, primariamente, métodos de QSAR e análise estatística, abordando uma triagem virtual híbrida (que consiste em uma abordagem baseada no ligante juntamente com

uma abordagem baseada na estrutura (REN et al., 2017; SHANG et al., 2017; SANTOS et al., 2018; LAM et al., 2018).

A triagem virtual basicamente consiste na classificação de substâncias com base nas características que melhor explicam a atividade biológica a ser observada, como é observado em um modelo de predição de atividade biológica, que utiliza inúmeras características moleculares (descritores) de moléculas conhecidas para prever a atividade das não conhecidas, conforme demonstra a figura (5) (SYDOW et al., 2019).

Figura 5. Exemplo geral dos fatores que geram um modelo de predição de atividade biológica.



Fonte: Próprio autor, 2021.

Deste modo, é possível obter as moléculas com melhores perfis farmacológicos *in silico*, através de uma gama de análises, a citar: predição de atividade biológica, biodisponibilidade, absorção, riscos de toxicidade, a interação ligante-receptor e a estabilidade no sítio ativo (PANDEY et al., 2017; RUIZ-TORRES et al., 2017; HANSMANN et al., 2018; CARACCILO, FAROKHZAD, MAHMOUDI, 2017).

De fato, as abordagens *in silico* permitem a triagem virtual de milhões de compostos em um tempo acessível, reduzindo assim os custos iniciais de identificação de acertos e melhorando

as chances de encontrar os candidatos a fármacos desejados. No momento, várias técnicas de modelagem molecular estão disponíveis para facilitar as tarefas de descoberta de fármacos, a maioria delas sendo classificadas em abordagens baseadas na estrutura e baseadas no ligante (AMIN et al., 2021).

Métodos baseados em estrutura contam com informações derivadas do conhecimento da estrutura 3D de um alvo de interesse, e permitem classificar bancos de dados de moléculas de acordo com a complementaridade estrutural e eletrônica dos ligantes para um determinado alvo (PINZI & RASTELLI, 2019).

Nesse contexto, o docking molecular está entre um dos mais populares e bem sucedidos métodos *in silico* baseados em estrutura, que ajudam a prever as interações que ocorrem entre moléculas e alvos biológicos (PINZI & RASTELLI, 2019; AMIN et al., 2021). Este processo é geralmente realizado primeiro prevendo a orientação molecular de um ligante dentro de um receptor e, em seguida, estimando sua complementaridade através do uso de uma função de pontuação (PINZI & RASTELLI, 2019).

## 2.7 Plantas medicinais

As plantas com fins medicinais são a maior fonte de moléculas farmacologicamente ativas e estas, consistem toda a planta ou partes da mesma que possuem compostos responsáveis pela ação terapêutica (BRASIL, 2017). O uso dessas plantas é difundido por todo o planeta, através de conhecimentos passados de geração em geração e seu uso é feito, principalmente, por aqueles que buscam complementar o tratamento de uma determinada enfermidade ou até mesmo como uma alternativa ao tratamento medicamentoso convencional (BORTOLUZZI, SCHMITT, MAZUR, 2019; BASSO, LOCATELLI, 2020).

O uso de plantas medicinais já ocorre há muitos anos e muitas destas plantas possuem potenciais farmacológicos, devido principalmente por componentes químicos originados do metabolismo secundário das mesmas, tais como os compostos fenólicos, representados por exemplo, pelos flavonóides e taninos (ALVES et al., 2019).

As espécies das famílias *Solanaceae* por exemplo, possuem inúmeras propriedades farmacológicas descritas, a citar: antivirais, fungicida, inseticida, bactericida, dentre outros (BEZERRAa, 2018; CHOWANSKI et al., 2016; SAMPAIO et al., 2019). Por outro lado, dentre as espécies da família *Fabaceae*, tem-se a *Erythrina velutina*, nativa da caatinga na região semiárida do nordeste brasileiro e alvo de diversos estudos já relatados na literatura que salientam suas propriedades terapêuticas, a citar: propriedades anti-inflamatórias e

anticoagulantes (MACHADO et al., 2013), gastroprotetoras (LIMA et al., 2017), neuroprotetor (RODRIGUES et al., 2017), analgésico (SILVA et al., 2020), dentre outros.

Frente a isto, pesquisas que se empenham na busca por novos compostos terapêuticos a partir de plantas ou no melhoramento de fitoterápicos já existentes vem se destacando e continuam a desempenhar um papel importante nos dias de hoje (BARROS, 2017). De acordo com Newman e Cragg (2016), o campo dos produtos naturais é envolvido em aproximadamente 50,6% de todos os novos fármacos aprovados pelo FDA e organizações semelhantes.

Contudo, mesmo os medicamentos gerados a partir de plantas tradicionais que são amplamente utilizados na medicina popular, existem ainda inúmeras plantas que não foram minuciosamente investigadas e precisam ser mais exploradas, a fim da descoberta de novos agentes terapêuticos (BARROS, 2017).

A abordagem etnofarmacológica demonstra ser bastante eficaz nas pesquisas envolvendo plantas medicinais para a produção de novos fármacos (LIRA, SOUSA, LINS, 2020). A conciliação de informações sobre as plantas ao seu modo de uso com determinado efeito terapêutico é estudada na etnofarmacologia e etnobotânica, possibilitando a triagem de espécies com base nos saberes tradicionais e adquirindo informações úteis essenciais nas mais diversas pesquisas (GOIS et al., 2016).

## **2.7 Metabólitos secundários**

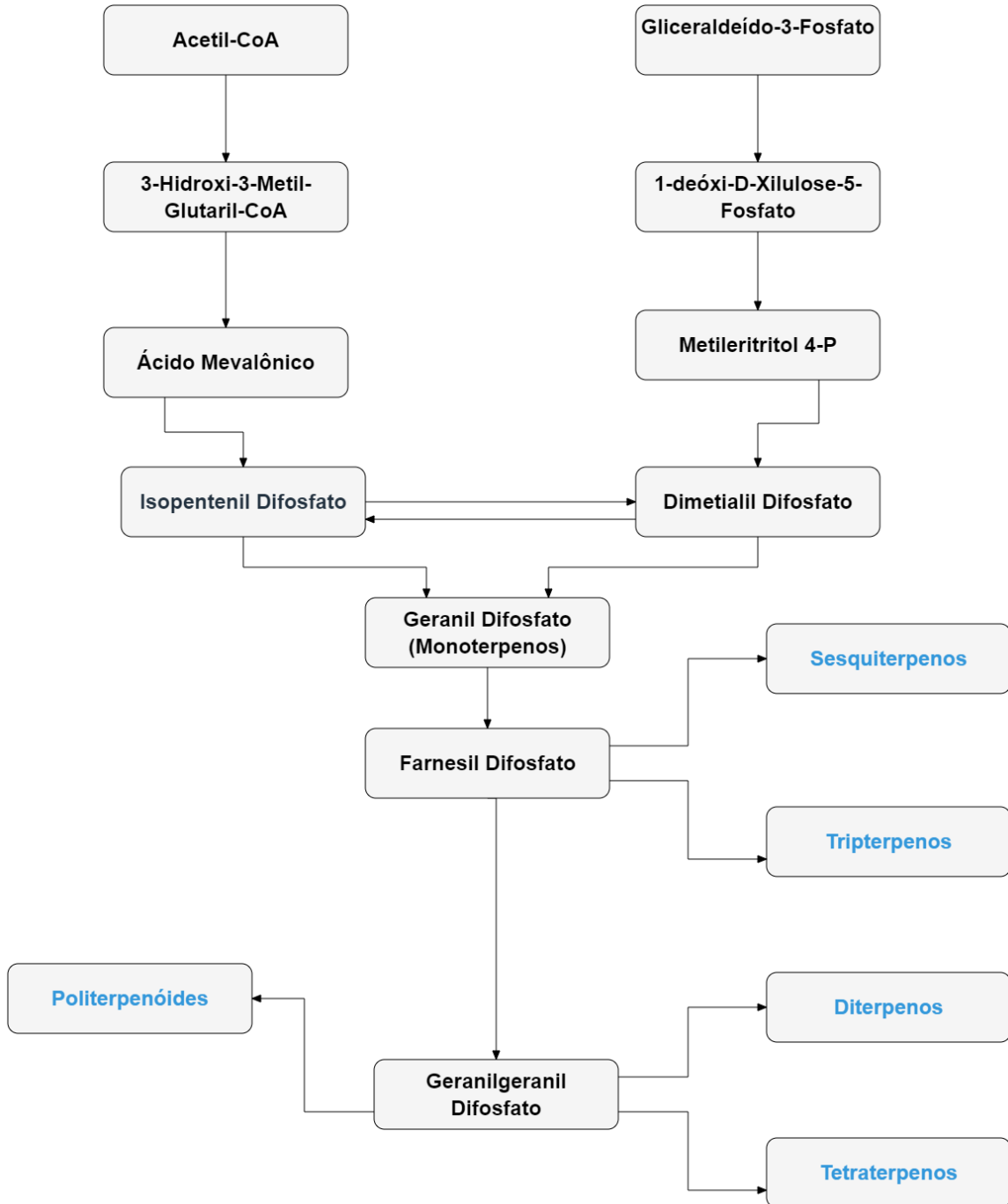
As plantas possuem um metabolismo secundário, não essencial à vida, que funciona com a produção de uma grande diversidade de produtos químicos. Os compostos provindos deste metabolismo são denominados metabólitos secundários, compostos naturais produzidos a partir da necessidade de proteção a estresses abióticos e bióticos, atração de agentes polinizadores, comunicação, dentre outros (BORGES, AMORIM, 2020; JASSBI et al., 2017; SILVA et al., 2020);

Os metabólitos secundários são principalmente, se não os únicos responsáveis por causar uma determinada ação terapêutica de uma planta e de acordo com Borges & Amorim (2020), esses compostos são divididos em três grupos principais:

- ✓ Terpenos - provindos da rota do ácido mevalônico e da rota do metileritritol fosfato, conforme demonstra a figura (6).
- ✓ Compostos fenólicos - provindos da rota da fenilalanina e ácido malônico, conforme demonstra a figura (7).

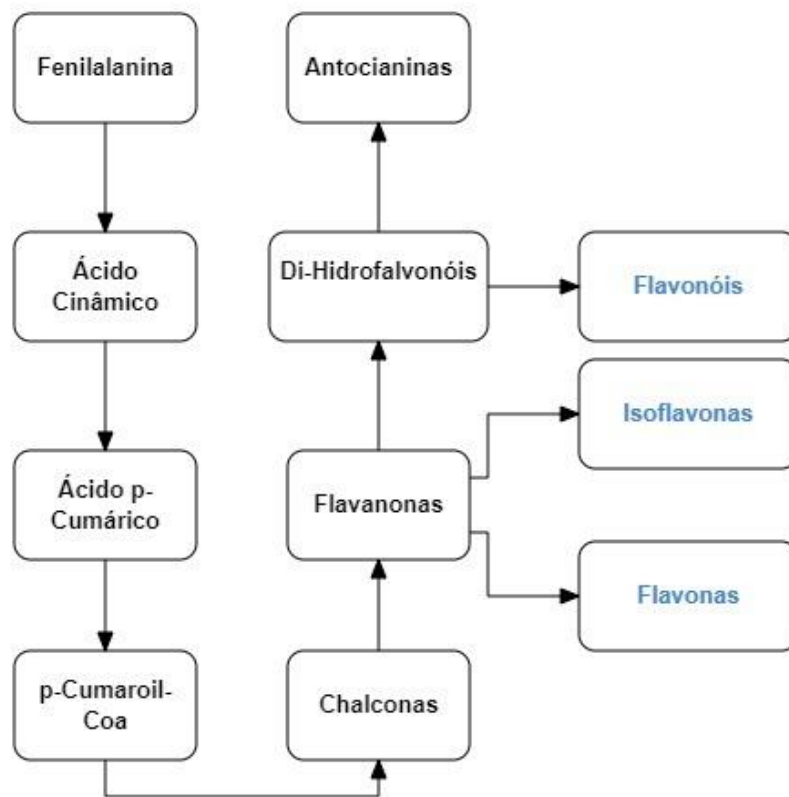
- ✓ Compostos nitrogenados - provindos do ciclo do ácido tricarbóxico e do ácido químico, conforme demonstra a figura (8).

Figura 6. Rota de biossíntese geral dos terpenos.



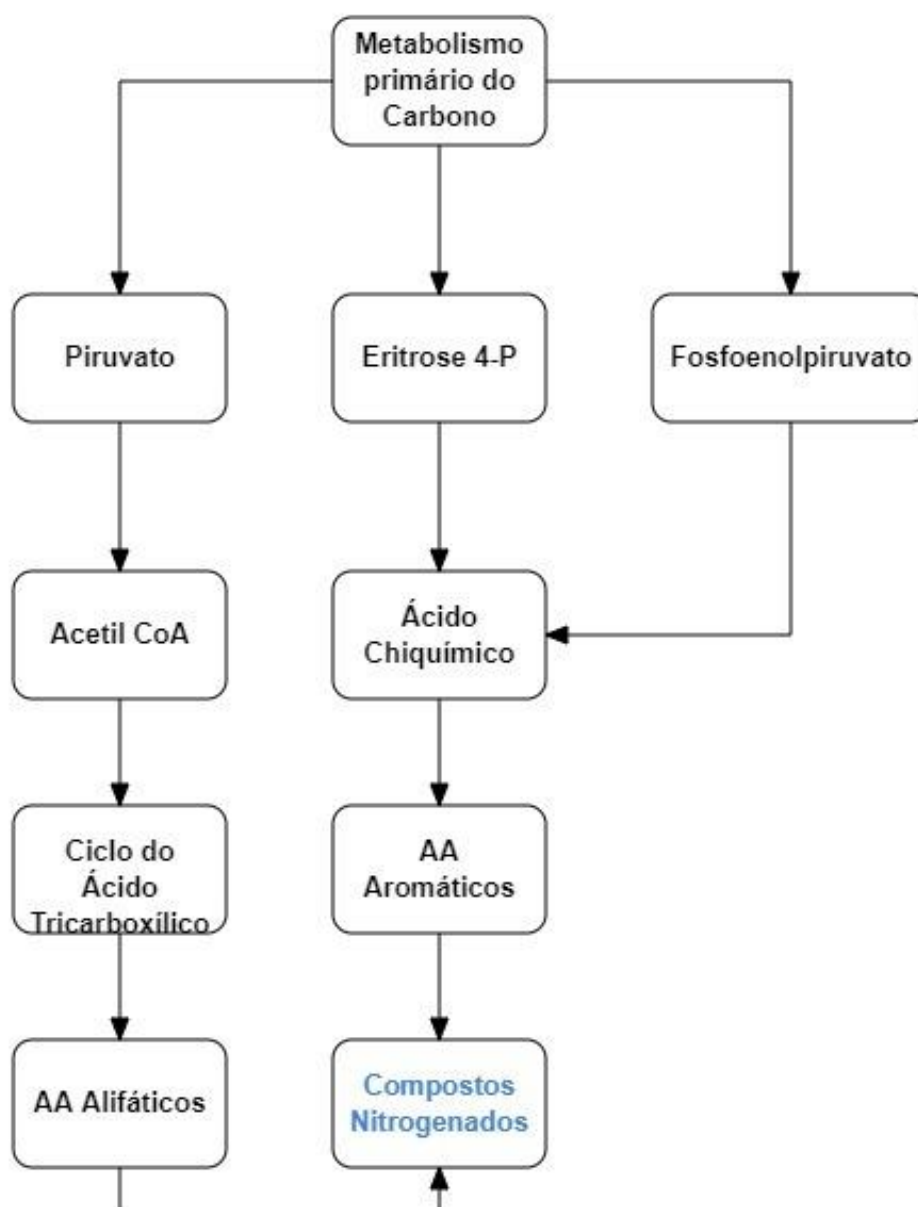
Fonte: Próprio autor, 2021.

Figura 7. Rota de biossíntese geral dos compostos fenólicos.



Fonte: Próprio autor, 2021.

Figura 8. Rota de biossíntese geral dos compostos nitrogenados.

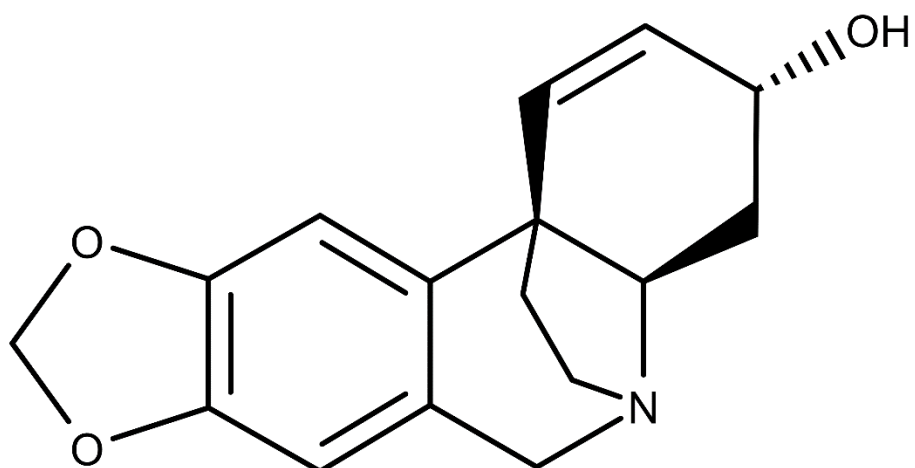


Fonte: Próprio autor, 2021.

Inúmeras espécies de plantas possuem metabólitos secundários farmacologicamente ativos, a citar: quinonas, flavonóides, saponinas, alcalóides, carotenóides e terpenóides, e estes, com inúmeras propriedades terapêuticas, inclusive antivirais, contribuindo para saúde da população (POUTARAUD; MICHELOT-ANTALIK; PLANTUREUX, 2017).

A exemplo, temos os alcalóides da família *Amaryllidacea*, relatados por suas propriedades antivirais, a citar os alcalóides crinina (figura 9) e licorina (GONZÁLEZ-CHAVARRO et al., 2020).

Figura 9. Arquitetura molecular do alcalóide crinina.



Fonte: Próprio autor, 2021.

Diante deste contexto, a multiplicidade molecular proveniente das plantas é imensa e tem sido bastante aproveitada na química medicinal, principalmente na obtenção de moléculas ativas, tendo em vista que os metabólitos secundários podem ser ativos em sistemas biológicos, isso já comprovado através dos inúmeros fármacos já em comercialização e que são metabólitos secundários (ROSALES, 2020).

Neste sentido, vários compostos oriundos do metabolismo secundário vegetal fazem parte do arsenal terapêutico atual, tais como os alcalóides: vimblastina (antineoplásico), vincristina (antineoplásico), morfina (analgésico), codeína (analgésico) e atropina (anticolinérgico) (BARROS, 2017; BAÑOS, GONZALÉZ-CHAVARRO et al ., 2020; GUARDIOLA, 2018; FIELDING et al., 2020). Assim, vários fármacos são derivados ou até os próprios metabólitos secundários isolados provindos das plantas, o que torna essencial o desenvolvimento de pesquisas envolvendo métodos rápidos, eficazes e de baixo custo frente a estes promissores compostos naturais (BARROS, 2017).

Dentro destas perspectivas, as moléculas naturais nos fornecem informações estruturais úteis na busca de novos compostos bioativos que possam vir a alcançar os requisitos necessários para ser um bom fármaco, tais como propriedades relacionadas a atividade, seletividade, toxicidade, tempo de meia vida, dentre outros (BARACUHY et al., 2016; MEDEIROS et al., 2020).

## 3 OBJETIVOS

### 3.1 Objetivo geral

Propor moléculas naturais com potencial atividade anti-influenza A, através de uma triagem virtual híbrida predita por uma análise de consenso de três modelos de predição de atividade biológica.

### 3.2 Objetivos específicos

- Buscar em um banco de dados contendo moléculas bioativas e conhecidas contra o influenza A, oriundas do ChEMBL database para a construção do modelo de predição;
- Buscar um conjunto de moléculas naturais oriundas do *Sistemat X database*, para serem as moléculas teste;
- Otimizar as estruturas químicas das moléculas bioativas conhecidas e das moléculas naturais teste, através do programa *Standardizer version 21.9*;
- Gerar os descritores moleculares de ambas, utilizando os programas *Dragon 5.0*, *CDK* e *VOLSURF*;
- Construir um modelo de predição de atividade, através do *software* KNIME Analytics Platform, para cada conjunto de descritores moleculares oriundas dos 3 programas citados anteriormente;
- Realizar uma análise por consenso para eleição das moléculas naturais de melhores perfis dentre as moléculas obtidas do *SistematX*;
- Investigar das interações dos compostos e metabólitos em isoformas do citocromo P450 através do programa *MetaSite*;
- Realizar análises da regra dos cinco de Lipinski, bem como calcular o percentual de absorção e biodisponibilidade por via oral;
- Realizar a ancoragem (*docking*) molecular, observando a interação dos compostos em estudo com quatro principais alvos biológicos importantes para o influenza A;
- Selecionar as moléculas mais promissoras de acordo com a triagem virtual híbrida (com base no ligante e na estrutura do alvo).

## 4. METODOLOGIA

A metodologia desta pesquisa pode ser dividida em 6 grandes etapas de trabalho: a) construção de três modelos de predição de atividade, b) análise por consenso dos três modelos de predição, c) análise da biodisponibilidade oral pela regra de Lipinski e taxa de absorção, d) análise do metabolismo das principais moléculas ativas, e) realização dos riscos de toxicidade, f) realização do docking molecular frente aos quatro alvos enzimáticos.

### 4.1 Conjunto e tratamento de dados das moléculas ativas

Inicialmente foi obtido um conjunto de dados a partir de moléculas ativas disponibilizadas gratuitamente no banco de dados de estruturas químicas ChEMBL (<https://www.ebi.ac.uk/chembl>) sendo usado como palavra chave “Influenza A” como alvo e foram consideradas as moléculas que apresentavam EC<sub>50</sub> (Concentração necessária para ocasionar metade do efeito máximo) (nM), como caráter de seleção.

Adicionalmente, para poder utilizar as planilhas de um banco de dados de estruturas, fez-se necessário realizar um conjunto de tratamentos de dados, que são procedimentos que visam garantir que em cada linha, coluna e célula da planilha estejam preenchidas com as informações destinadas para aquela molécula, a fim de garantir uma linearidade de dados e evitar erros nos programas subsequentes.

Para alcançar tal objetivo, todo o procedimento de tratamento foi feito utilizando o software *Microsoft Office Excel*, em que se pode: organizar e estruturar a planilha, remover as duplicatas, padronizar as unidades química envolvidas, principalmente de um conjunto de dados referente a mesma propriedade química, no caso, a EC<sub>50</sub> (nM), assim, o uso de conversões de unidades muitas vezes é empregado em planilhas oriundas de banco de dados de moléculas.

Vale salientar que toda e qualquer planilha com dados que irão ser utilizados em quimioinformática, a estruturação, organização, remoção das duplicadas, remoção de células vazias e a padronização das unidades são fundamentais para o êxito das análises *in silico*, tendo em vista que a qualidade dos dados de saída para cada análise computacional depende da qualidade dos dados de entrada.

Além disso, este conjunto de moléculas e dados foram agrupados em uma única planilha utilizando o Microsoft Office Excel 365, ao qual foi posteriormente tratadas para obter melhor linearidade e qualidade de dados, a citar: como caráter de exclusão, foi selecionado apenas as moléculas com peso molecular de até 900 mm (milimol), uma vez que moléculas grandes e pesadas não é de interesse, tendo em vista que poderá ocasionar em uma redução da

sua permeabilidade pela membrana, já que pressupõe que será uma molécula grande e pesada, tornando difícil sua passagem pela membrana.

Após o tratamento da planilha, por fim, o conjunto de dados foram utilizados na criação do modelo de predição de atividade biológica (mPRECT).

## **4.2 Conjunto e tratamento de dados das moléculas candidatas**

O conjunto de moléculas naturais a passarem pela triagem virtual, foram provenientes de um banco de dados da plataforma *SistematX data base* (SCOTTI et al., 2018). O banco de dados contendo as moléculas naturais passaram pelo mesmo tratamento citado anteriormente, para as moléculas bioativas conhecidas.

## **4.3 Otimização da arquitetura molecular**

As arquiteturas moleculares em 2D e 3D otimizadas das 803 moléculas bioativas oriundas do ChEMBL e das 989 moléculas naturais a serem testadas oriundas do *SistematX* foram obtidas a partir de seus respectivos códigos de SMILES através do software *Standardizer 18.21.0* da *ChemAxon*®, agrupadas em um arquivo no formato “SDF” contendo os dados e coordenadas espaciais de cada estrutura.

As estruturas químicas é a base para qualquer procedimento quimioinformático, a partir dessas estruturas é possível realizar o cálculo de descritores moleculares com base no objetivo de cada estudo, além da padronização das unidades químicas, as padronizações das estruturas químicas podem fazer com que um pesquisador de dados ganhe bastante tempo de análises que possivelmente dariam erradas ou gerariam dados sem qualidades.

Com isto, tal padronização consistiu, especificamente nesta pesquisa: na adição de hidrogênios que poderiam estarem ausentes nas estruturas, conversão para 2D e 3D, a aromatização, remoção de prováveis solventes, a remoção de sais e outros diversos procedimentos que o *software Standardizer* disponibiliza.

## **4.4 Descritores Moleculares**

Para a construção de um modelo de predição, faz-se necessário o uso de descritores moleculares, que são características estruturais e químicas de cada molécula envolvida. Após a

otimização das estruturas químicas de todas as moléculas envolvidas, feitas pelo *Standardizer*, o próximo passo foi gerar os descritores moleculares de tais moléculas.

Neste sentido, para que uma estrutura química de uma determinada molécula seja compreendida e processada por um determinado programa, ela precisa ser descrita em uma sequência numérica única (ALVES et al., 2018). Os descritores moleculares descrevem propriedades que podem ser calculadas a partir de uma representação unidimensional (1d), embasados nas propriedades físico-químicas e na fórmula molecular (logP e massa molecular, por exemplo), bidimensionais (2D), que descrevem propriedades que podem ser representadas como ligações e átomos (número de átomos de nitrogênio, número de ligações, dentre outros) e tridimensionais (3D), que dependem da conformação das moléculas (volume molecular, área da superfície, dentre outros), podendo serem classificados, mediante sua natureza, em: descritores constitucionais, descritores híbridos, descritores topológicos, descritores geométricos e descritores eletrônicos (ALVES et al., 2018; RODRIGUES et al., 2021; SILVAd, 2020). Cada tipo com seus descritores moleculares, e cada um com seu significado particular (SILVAd, 2020).

A utilização de descritores moleculares permite analisar o grau de variabilidade estrutural existente em um conjunto de moléculas, o que é comumente útil no design de bibliotecas combinatoriais, além disso, permite estabelecer parâmetros de semelhança estrutural em um conjunto de moléculas, o que justifica sua ampla utilização no campo dos modelos de predição de atividade biológica (SILVAd, 2020).

Diante de tal importância, os descritores são calculados a partir da seleção de um arquivo de entrada, que geralmente é um arquivo no formato “SDF” contendo todas as estruturas químicas otimizadas a serem analisadas e então, o cálculo dos descritores é iniciado. Existem muitos programas para geração de descritores moleculares, dos mais simples aos mais complexos, como exemplo o aplicativo gratuito em *Java CDK Descriptor Calculator 1.4.8*, o software não livre *Dragon 5.0* e o software não livre *Volsurf+*.

Assim, as estruturas químicas otimizadas provenientes do programa *Standardizer* foram salvas no formato “SDF” e então, foram submetidas aos programas listados acima (*CDK*, *Dragon* e *VolSurf*), a fim de gerar os descritores moleculares, uma vez que são imprescindíveis para a construção de um modelo de predição. Diante disto, foram realizados três modelos preditivos, cada um com base nos descritores fornecidos pelos três programas. E com isso, é possível então obter uma predição por consenso, ao cruzar os resultados das três predições, fazendo com que o método de predição de atividade ganhe maior confiabilidade e maior robustez a metodologia (MOREIRA, 2014).

## 4.5 Modelo de predição de atividade biológica – mPRECT

### 4.5.1 *KNIME Analytics Platforms 3.6*

Os modelos de predição da atividade biológica quimioinformáticos são derivados de um processo de aprendizado supervisionado, em que o modelo é treinado a partir de uma base de dados de moléculas conhecidas, a base de treinamento do modelo. Dito isto, os modelos foram feitos a partir do *software* estatístico *KNIME Analytics Platforms 3.6* (SYDOW et al., 2019).

Este é uma ferramenta de quimioinformática embasada em um mapa auto organizável que dispõe de uma gama de vantagens, dentre elas está a interoperabilidade entre outros *softwares*, integrando diversos nós de dados em conjunto com inúmeros outros *softwares*, que juntos formam o modelo de predição (FILLBRUNN et al., 2017; DIETZ et al., 2020).

O *KNIME* por si só, já vem com uma vasta série de catálogos de nós de processamento, disponibilizando uma infinidade de extensões e algoritmos para as mais diversas áreas de pesquisa e com isso, há uma integração de ferramentas e fornecimento de uma estrutura de dados previamente conectadas e compartilhada, permitindo que os nós de diferentes áreas sejam interconectados para criar fluxos de trabalho (workflow) de grande complexidades (FILLBRUNN et al., 2017; MASSARO et al., 2018).

Com isso, o modelo de predição proposto pelo *KNIME* consiste na aplicabilidade de inteligência artificial e aprendizado de máquina, no qual o modelo em si, orienta a máquina (computador) a classificar cada estrutura molecular cuja atividade deseja-se predizer, com base na similaridade de suas características estruturais, ou seja, dos descritores moleculares (FILLBRUNN et al., 2017; MARTINS, 2017; SYDOW et al., 2019).

Neste sentido, com base na comparação dessas características estruturais, a busca de similaridades ocorre a partir das moléculas classificadas como ativas no momento da construção do modelo e uma vez a máquina sabendo das características de tais moléculas ativas e inativas, por comparação, ela pode classificar moléculas cuja predição de uma determinada atividade biológica não tenha sido relatada na literatura (ALVES et al., 2018; MARTINS, 2017; FILLBRUNN et al., 2017; PICCIRILLO & AMARAL, 2018).

A similaridade é analisada com base dos descritores moleculares das moléculas ativas e comparados com os das moléculas cuja atividade é desconhecida, ou seja, as moléculas naturais. Adicionalmente, houve uma divisão na tabela das moléculas ativas utilizadas na construção do modelo, com base no seguinte critério: divide-se o conjunto das moléculas ativas utilizadas para a construção do modelo ao meio, as moléculas com  $pEC_{50}$  ( $-\log(EC_{50})$ ) mais altos são atribuídas

a atividade “A” (ativas) e as demais como “I” (inativas), assim a máquina é capaz de classificar qualquer outra estrutura que esteja dentro do domínio de aplicabilidade do modelo.

#### 4.5.2 *Random Forest*

Na presente pesquisa foi utilizado o *Random Forest* como algoritmo de classificação integrado com o aprendizado de máquina, tendo em vista sua precisão, eficiência e robustez (FALCÓN-CANO et al., 2020; FASHOTO et al., 2018; PETKOVIC et al., 2018). O *Random Forest* engloba um conjunto de árvores de decisões que analisam as moléculas randomicamente, sendo possível classificar cada molécula através de uma análise de consenso entre as diversas árvores de decisão (DOGRU & SUBASI, 2018; PETKOVIC et al., 2018; FALCÓN-CANO et al., 2020).

Para isso, o banco contendo as moléculas ativas contra o influenza A utilizadas para construir o modelo de predição foi particionado aleatoriamente em treino (80%) e teste (20%) no qual a classificação final é dada pela maioria de ocorrência de uma classificação em cada molécula (FORTI, 2018; VAZ, 2017; PETKOVIC et al., 2018). Posteriormente ao particionamento deste banco de dados, a máquina começa a classificar cada estrutura química pela similaridade com as estruturas das moléculas inicialmente admitidas como ativas, quanto mais similar com as moléculas ativas e dentro do domínio de aplicabilidade, maior a probabilidade da estrutura que está tendo a atividade ou propriedade predita ser também ativa (ALMEIDA, 2018; LINS, 2018; MARTINS, 2017).

Neste sentido, para a predição ser considerada aceitável, a molécula cuja predição está sendo feita deverá estar inserida no domínio de aplicabilidade, uma vez que esse domínio garante que as características das moléculas testadas são conhecidas pelo modelo criado. Como o modelo toma como base na comparação por similaridade, o domínio de aplicabilidade garante que a molécula cuja predição está sendo feita, apresente características estruturais representadas no espaço químico das moléculas utilizados na construção do modelo (as moléculas ativas contra o influenza A) e quando as moléculas testadas e as moléculas do modelo estão no mesmo espaço químico, aumenta-se a confiabilidade da predição e quanto maior a porcentagem da aplicabilidade mais confiável é a predição para a molécula analisada. (ALVES et al., 2017; RIBEIRO, 2017).

#### 4.5.3 *Wecka Predictor 3.7*

Especificamente, a extensão utilizada para a predição da atividade foi o *Weka Predictor* 3.7, para realizar a classificação das moléculas naturais cuja atividade anti influenza A é desconhecida (FILLBRUNN et al., 2017; ROSA et al., 2018). Somando o aprendizado de máquina do *Random Forest* e a classificação da *Wecka Predictor*, o mapa auto organizável criado no *workflow* do *KNIME* vem ganhando as atenções de muitos pesquisadores na área de quimioinformática pela sua eficiência na predição e facilidade de configuração (ROSA et al., 2018; ABDELJOUAD, BRAHAMI, MATTA, 2020).

Para avaliar a confiabilidade do modelo serão considerados, com base na tabela de confusão gerada pelo modelo, a acurácia acima de 70% e o coeficiente de correlação de Matthews acima de 50%, estes são parâmetros estatísticos que garantem a qualidade das predições realizadas. Outros parâmetros também serão analisados na discussão dos dados e da eficiência preditiva do modelo, que são: as curvas ROC (*Receiver Operating Characteristic*) do teste e da validação interna e o domínio de aplicabilidade.

#### 4.5.4 Curva ROC

Um dos métodos de avaliação mais importantes para verificar o desempenho de qualquer modelo de predição, é através da análise da curva ROC, um gráfico que mostra o desempenho do modelo em todos os limites de classificação, a qual permite comparar o desempenho do modelo através da escolha de critérios de classificação em bons e maus pagadores, de acordo com a escolha de diferentes pontos de corte ao longo das amplitudes dos escores observadas para os modelos obtidos (HOO et al., 2017; ZWEIG & CAMPBELL, 1993; ZENG et al., 2021).

Para tanto, os gráficos das curvas ROC são criadas plotando-se a sensibilidade (verdadeiro-positivo) no eixo y contra 1 – especificidade (verdadeiro-negativo) no eixo x, conseqüentemente, a curva ROC deve ser interpretada de forma que, quanto mais a curva estiver distante da diagonal principal, melhor o desempenho do modelo associado a ela. Esse fato sugere que, quanto maior for a área entre a curva ROC produzida e a diagonal principal, melhor é o desempenho geral do modelo (GARCÍA, FERREIRA, PATINO, 2021; RUISÁNCHEZ et al., 2021).

Adicionalmente, a área sob a curva (AUC) é um parâmetro útil frequentemente usado para comparar o desempenho geral de curvas diferentes. Um valor próximo a 1 mostra a capacidade máxima de classificação, enquanto a linha diagonal (AUC = 0,5) indica a

capacidade mínima de aceitabilidade (GARCÍA, FERREIRA, PATINO, 2021; HOO et al., 2017; RUISÁNCHEZ et al., 2021).

Diante disto, o fluxo de trabalho do modelo de predição de atividade utilizado nesta pesquisa pode ser facilmente encontrado através do seguinte DOI (Digital Object Identifier): 10.6084 / m9.figshare.16777612.

#### **4.6 Regra dos Cinco de Lipinski e %ABS**

Tendo em vista que a maioria dos fármacos utilizados para o tratamento do influenza A apresentam-se na forma de comprimidos e são administrados por via oral, é de suma importância utilizar a regra de Lipinski, uma vez que dispões de parâmetros utilizados para avaliar o padrão de biodisponibilidade oral teórica das moléculas alvo, baseando-se na “Regra dos Cinco” desenvolvida por Lipinski et al. (1997), calcula o coeficiente de partição (LogP), a massa molecular em dáltons, o número de doadores de ligação de Hidrogênio (nDLH) e o de aceptores (nALH), além de dar o valor da Topological Polar Surface Area (TPSA, traduzido para o português: Área da Superfície Polar Topológica), necessária para calcular o percentual de absorção oral teórico (%ABS), a partir da equação:  $\%ABS = 109 - (0.345 \times TPSA)$ .

Para estimar as violações a regra de Lipinski foi usado o software *Molinspiration Property Calculator* (MOLINSPIRATION, 2021), que possibilita o cálculo de importantes propriedades físico-químicas da molécula. Ao utilizar a taxa de absorção como parâmetro de exclusão na triagem virtual, apenas são considerados as moléculas que apresentaram %ABS igual ou superior a menor taxa dentre os fármacos usados como controle:

- Amantadina - 100 %;
- Rimantadina - 100%;
- Baloxavir - 83.12%;
- Oseltamivir - 77,72%;
- Zanamivir - 39,75%.

#### **4.7 Predição dos riscos de toxicidade**

As moléculas foram submetidas à análises de suas prováveis toxicidades, a citar: carcinogenicidade, tumorigenicidade, citotoxicidade, efeito no sistema reprodutor e irritabilidade na pele. Para isto, foi utilizado o software *OSIRIS DataWarrior 6.0*, uma vez que neste programa, os riscos de toxicidade é predito pela comparação entre um banco de dados interno, no qual o programa busca fragmentos conhecidos nas moléculas analisadas e esses

fragmentos quando encontrados, geram um alerta de toxicidade em quatro parâmetros selecionados: mutagenicidade, carcinogenicidade, efeito no sistema reprodutor e irritabilidade na pele (GUREVICH et al., 2021; RAO & SHETTY, 2021)

Neste sentido, mediante a toxicidade geral, foi atribuída então o critério “Não” quando uma molécula não apresenta alertas para nenhum dos quatro parâmetros analisados pelo OSIRIS. E com isso, foi utilizado como critério de exclusão, as moléculas que apresentaram risco de toxicidade em pelo menos um dos parâmetros citados anteriormente, pois assim, há uma redução da probabilidade de a molécula apresentar efeito tóxico, o que não é viável para um candidato a fármaco.

#### 4.8 Metabolismo

O *software* não livre *MetaSite*<sup>TM</sup> 6.0 da *Molecular Discovery* foi utilizado para realizar as predições do metabolismo *in silico* (METASITE, 2021). Este, forneceu um resumo das informações sobre os metabólitos previstos, como suas estruturas, fórmulas moleculares, porcentagem de abundância metabólica, dentre outros (ODOARDI et al., 2021). Especificamente, o *MetaSite* prediz as possíveis transformações sofridas pelas moléculas em um ou mais citocromos (proteínas responsáveis em transformar através de reações químicas as substâncias ingeridas para promover sua eliminação através da polarização da estrutura), em alguns regiões específicas do corpo como pele, cérebro, fígado dentre outros (CECE-ESENCAN et al., 2016; RADCHENKO et al., 2017; ODOARDI et al., 2021).

Vale salientar que neste estudo, foram considerados apenas as estruturas químicas dos metabólitos secundários provenientes do metabolismo do fígado, e estes, foram submetidos a predição dos riscos de toxicidade, comparados com as suas respectivas abundâncias metabólicas.

Cada metabólito secundário teve seus riscos de toxicidade preditos, através do *software OSIRIS Data Warrior*, pois os bioativos que apresentavam parâmetros citotóxicos como alta mutagenicidade e/ou alta carcinogenicidade em seus metabólitos com abundância acima de 70% era desconsiderado para a seleção das moléculas anti-influenza A de melhor potencial a candidatos a fármacos, tendo em vista que, compreende-se para esta pesquisa que um candidato a fármaco deve apresentar baixa ou nenhum tipo de toxicidade, parâmetro que serve também em seus metabólitos secundários mais abundantes.

#### 4.9 Docking Molecular

Para a realização do *docking molecular* foi utilizado o *software* não livre *Molegro Virtual Docker 6.0*, no qual as moléculas foram importadas juntamente com a proteína oriunda do site *Protein Data Bank* ([www.rcsb.org](http://www.rcsb.org)). Os *templates* foram criados baseados nos inibidores complexados juntos as proteínas estudadas: canal M2 (PDB ID - 3BKD, STOUFFER et al., 2008), hemaglutinina (PDB ID - 1RUY, GAMBLIN et al., 2004), neuraminidase N1 (PDB ID - 2HTY, RUSSEL et al., 2006) e polimerase (PDB ID - 6FS7, OMOTO et al., 2018).

Ainda, além da realização do *docking molecular*, foi realizado o *redocking* com o ligante complexado junto a proteína com as mesmas condições computacionais que o *docking* original, a fim da descoberta do valor do RMSD (*Root Mean Square Deviation*), que é o valor do desvio quadrático médio entre os mesmos átomos do ligante complexado junto a proteína e o mesmo ligante importado, uma vez que, com ele é possível validar o protocolo usado para o acoplamento, quanto mais baixo o RMSD melhor a precisão do encaixe molecular realizado, comumente adotados valores menores que 3.0 Å (LOO et al., 2018; NEVEU et al., 2018; PADHORNY et al., 2018; PAGADALA et al., 2017).

Na interpretação dos resultados do *docking*, além das energias é possível analisar os tipos de interações, que no *Molegro* são três: hidrogênio (linhas tracejadas azuis), eletrostáticas (linhas tracejadas verdes) e estéricas (linhas tracejadas vermelhas). Ainda mais, é possível ver os resíduos de aminoácidos envolvidos em cada interação entre os ligantes e os respectivos receptores, com isso, tenta-se identificar possíveis resíduos comuns entre os análogos e os controles e/ou inibidores complexados juntos a cada proteína na busca de relacionar a importância dessas interações para a atividade biológica.

## 5. RESULTADOS E DISCUSSÕES

### 5.1 Modelo de predição de atividade biológica

A princípio, tanto as moléculas ativas quanto as moléculas candidatas, foram submetidas a inúmeros programas computacionais, para otimização estrutural, geração de descritores moleculares e banco de dados, bem como o tratamento das planilhas para submissão nos modelos de predição. Após isto, por fim, o conjunto de dados apresentou um total de 803 moléculas bioativas anti-influenza A registrados na literatura (disponíveis do ChEMBL).

As 803 moléculas com atividade conhecida frente ao influenza A, em conjunto com as 989 moléculas naturais com atividade desconhecida e candidatas a fármacos, foram submetidas ao modelo de predição de atividade biológica, desenvolvido no *software KNIME Analytics Platform*. Vale salientar que foram construídos três modelos de predição, cada qual utilizando descritores moleculares fornecidos pelos *softwares Dragon, CDK e Volsuf*, a fim de realizar uma análise por consenso, garantindo um maior percentual de confiabilidade para tais moléculas.

A princípio, para levar em consideração a predição do modelo, é necessário que tal apresente uma confiabilidade de predição, e para isto, os dados estatísticos dos modelos de predição foram capazes de confirmar que tais modelos apresentaram confiabilidade e isto pode ser confirmado a partir dos valores dos dados estatísticos oriundos dos três modelos de predição, utilizando os descritores do *Dragon, CDK e Volsuf*, respectivamente, como demonstrado nas tabelas (1,2,3).

Tabela 1. Resultados dos parâmetros estatísticos de confiabilidade do modelo de predição utilizando os descritores do *Dragon*.

Parâmetros de confiabilidade do modelo de predição					
	Precisão	Sensibilidade	Especificidade	MCC	ACC
<b>Validação interna</b>	0,837	0,790	0,848	0,640	0,819
<b>Teste</b>	0,847	0,762	0,864	0,630	0,814

Legendas: MCC: coeficiente de correlação de matthews; ACC: acurácia.

Fonte: Próprio autor, 2021.

Tabela 2. Resultados dos parâmetros estatísticos de confiabilidade do modelo de predição utilizando os descritores do *CDK*.

<b>Parâmetros de confiabilidade do modelo de predição</b>					
	Precisão	Sensibilidade	Especificidade	MCC	Acurácia
<b>Validação interna</b>	0,803	0,790	0,808	0,598	0,798
<b>Teste</b>	0,833	0,750	0,852	0,605	0,801

Fonte: Próprio autor, 2021.

Tabela 3. Resultados dos parâmetros estatísticos de confiabilidade do modelo de predição utilizando os descritores do *VolSurf*.

<b>Parâmetros de confiabilidade do modelo de predição</b>					
	Precisão	Sensibilidade	Especificidade	MCC	Acurácia
<b>Validação interna</b>	0,773	0,777	0,774	0,601	0,776
<b>Teste</b>	0,829	0,787	0,840	0,628	0,814

Fonte: Próprio autor, 2021.

Nas tabelas (1,2,3), encontra-se os valores de precisão, sensibilidade, especificidade, MCC e ACC. Discutindo primeiramente os dados estatísticos do grupo de teste dos três modelos, a acurácia (ACC) obteve valores que variaram entre 0,801-0,814, sendo considerado ótimos valores, tendo em vista que tal parâmetro refere-se ao grau de observância dos dados e o quanto cada predição aproxima-se da atividade biológica real das moléculas analisadas, um valor satisfatório de acurácia dos modelos também pode ser calculado para a validação interna cruzada (*cross validação*), que apresentou valores entre 0,776-0,819.

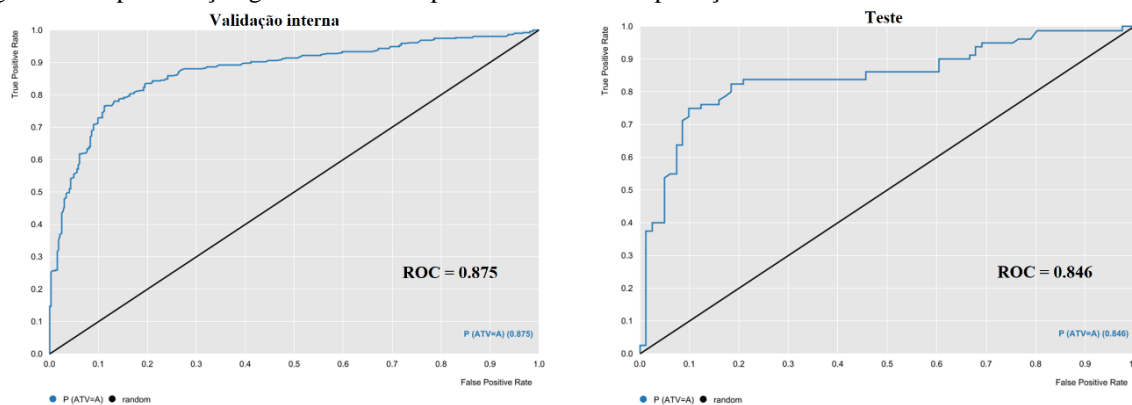
Além da acurácia, outros parâmetros estatísticos foram calculados para comprovar o desempenho e confiabilidade do modelo, como por exemplo, o coeficiente de correlação de Matthews (MCC), muito relevante quando se fala de aprendizado de máquinas, tendo em vista que expressa a correlação total entre os dados observados e os dados preditos, dentre outras

palavras, ele avalia de modo geral a confiabilidade do modelo. Neste sentido, obteve-se para o grupo de teste um MCC com valores de 0,605-0,630 e para a validação interna 0,598-0,640.

Adicionalmente, outros parâmetros que em conjunto garantem a confiabilidade de predição do modelo, são a precisão, sensibilidade e especificidade, em que ambos apresentaram também, ótimos valores, sendo isto demonstrado nas tabelas (1,2,3). Frente a isto, todos os valores apresentados demonstram o desempenho e confiabilidade satisfatória dos modelos, tendo em vista que tais valores estão acima de 70% (0,7), tornando os modelos desenvolvidos para a predição da atividade anti-influenza A neste estudo, aceitável, uma vez que tais modelos apresentaram uma ótima capacidade de classificação dos dados (sensibilidade e especificidade), ótima proporção de verdadeiros positivos (VPP) e boa proporção de verdadeiros negativos (VNP), ótima predição do quanto se aproxima da atividade biológica (ACC) e boa correlação entre os dados observados e preditos (MCC).

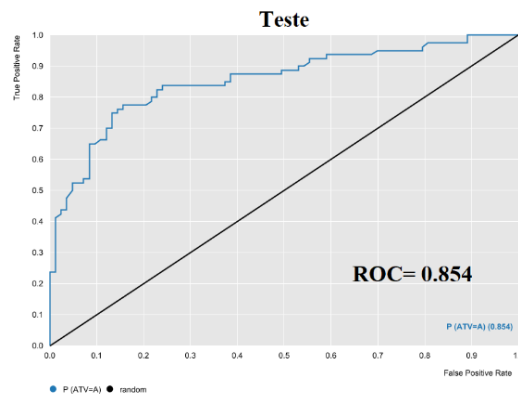
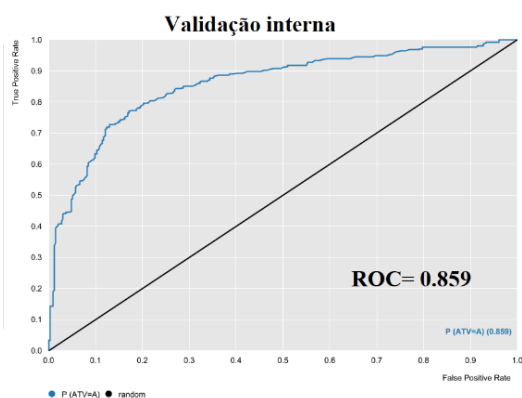
Com tais valores, é comprovado que os modelos são capazes de prever, graças a representação gráfica do desempenho, tanto do teste quanto da validação interna, por meio da curva ROC, representada nas figuras (10, 11, 12), a partir dos modelos gerados pelos descritores *Dragon*, *CDK* e *VolSurf*.

Figura 10. Representação gráfica do desempenho do modelo de predição de atividade dos descritores do *Dragon*.



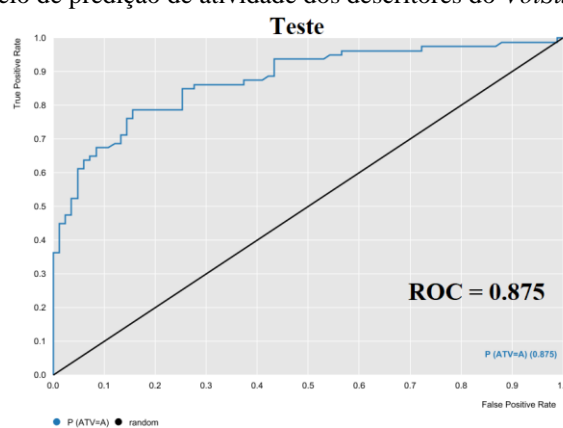
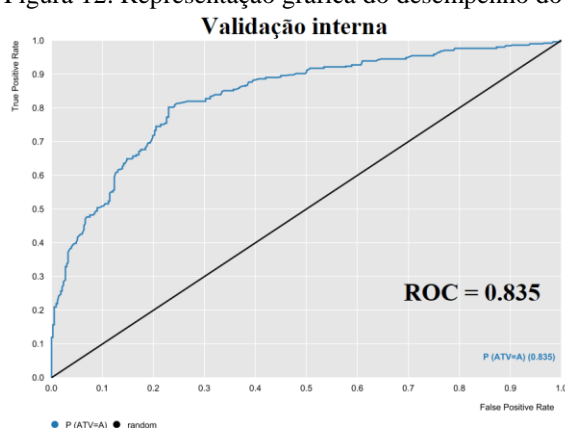
Fonte: Próprio autor, 2021.

Figura 11. Representação gráfica do desempenho do modelo de predição de atividade dos descritores do *CDK*.



Fonte: Próprio autor, 2021.

Figura 12. Representação gráfica do desempenho do modelo de predição de atividade dos descritores do *VolSurf*.



Fonte: Próprio autor, 2021.

Após a decisão de confiabilidade do modelo segundo os dados estatísticos, as 989 moléculas naturais com atividade desconhecida foram então submetidas aos modelos de predição de atividade biológica discutido acima. Em um modelo de predição, o pesquisador força a classificação de moléculas com atividade desconhecida, a partir de semelhanças no ambiente químico de moléculas com atividade conhecida pré-classificadas.

A pré-classificação foi realizada da seguinte forma: após o tratamento do banco de dados contendo as 803 moléculas de atividade conhecida, foram calculados o  $pEC_{50}$  ( $= -\log(EC_{50})$ ) para cada molécula da planilha tratada, salientando que o banco de dados já traz os valores de  $EC_{50}$  para cada molécula, contudo, para facilitar o tratamento de dados, foi transformado para  $pEC_{50}$ , que pode ser definido como o logaritmo negativo da  $EC_{50}$ .

Posteriormente, esta planilha foi então ordenada de forma decrescente de  $pEC_{50}$  (pois quanto maior o valor, mais potente é a molécula, já que é o inverso da  $EC_{50}$ ), por fim, as moléculas a partir da mediana do conjunto de dados foi classificado como “Ativo” e abaixo da mediana como “Inativo”. Especificamente, foram consideradas como ativas, as moléculas com

pEC<sub>50</sub> igual ou maior a 5,14 nM e como inativas, as que continham valores de pEC<sub>50</sub> igual ou menor a 5,07 nM.

Neste sentido, como cada planilha contém os descritores moleculares gerados pelo *Dragon*, *CDK* e *VolSurf*, o modelo em si poderá classificar as moléculas com base na semelhança dos ambientes químicos estruturais entre as moléculas pré-classificadas com base na atividade conhecida e as moléculas cuja atividade é desconhecida. Tendo conhecimento disto, faz-se necessário citar o “domínio de aplicabilidade”, que é de suma importância para um modelo de predição, uma vez que este preserva todas as características estruturais das moléculas que formam o modelo, assim, aquelas moléculas que são aprovadas pelo domínio possuem a garantia que o modelo foi capaz de classificá-la tendo como base as características das moléculas que constituem o modelo.

Por exemplo, um certo modelo de predição para uma atividade desejada foi desenvolvido utilizando moléculas grandes, ramificadas, cíclicas e com diversos heteroátomos, este modelo então, não será capaz de classificar moléculas que apresentem anéis ramificados com apenas um heteroátomo, ou então que apresente estrutura simples e alifática. Tal falta de classificação pode ser comprovada pelos baixos valores estatísticos, indicando que o modelo não apresenta confiabilidade e baixa capacidade preditiva por exemplo.

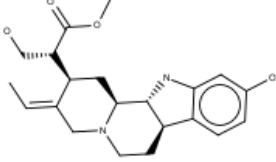
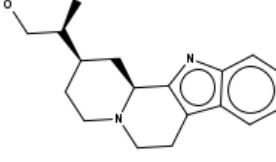
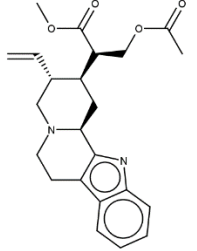
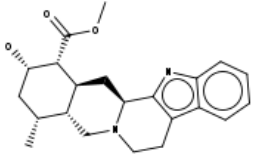
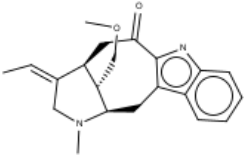
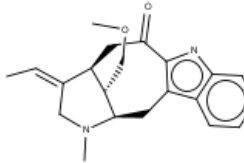
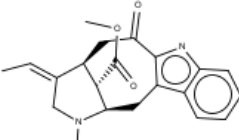
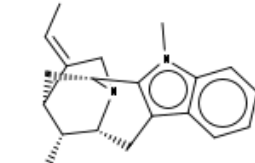
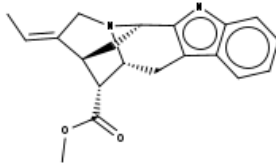
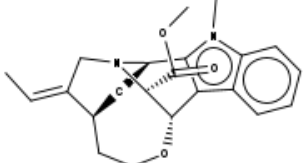
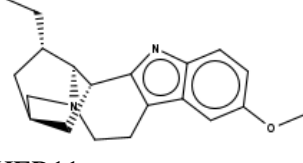
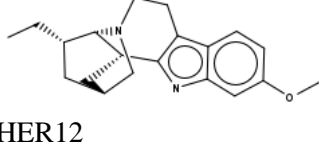
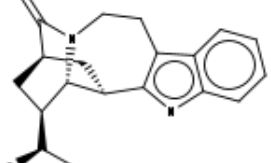
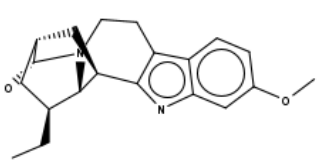
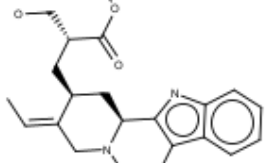
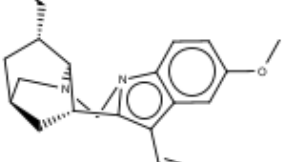
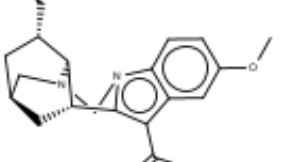
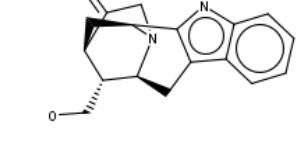
Ainda, essa falta de capacidade do modelo em classificar esse conjunto de moléculas pode ser facilmente explicada pelo fato de que as características das moléculas cuja atividade quer ser predita não estarem dentro do domínio de aplicabilidade e para contornar esse problema, comumente são inseridas na criação do modelo, moléculas com características estruturais com um mínimo de similaridade com as das moléculas cuja atividade é conhecida.

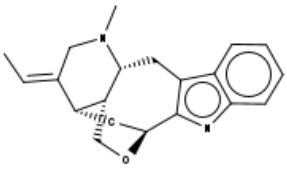
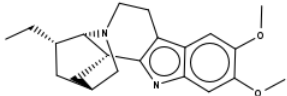
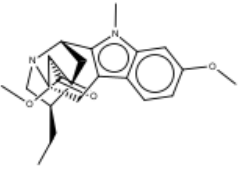
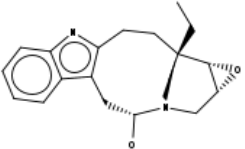
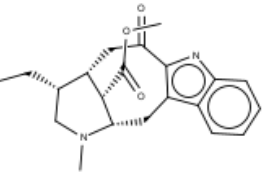
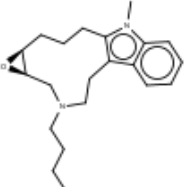
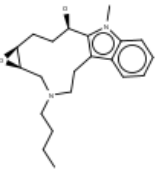
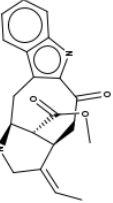
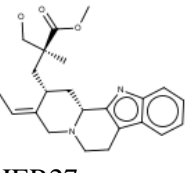
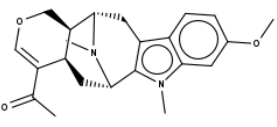
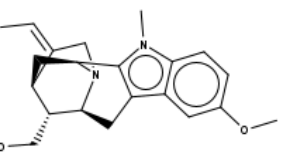
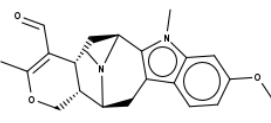
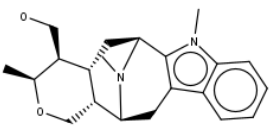
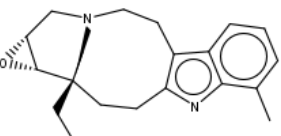
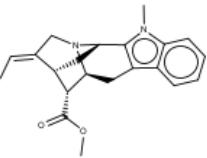
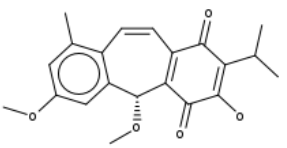
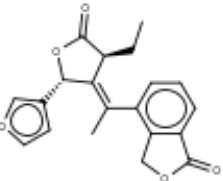
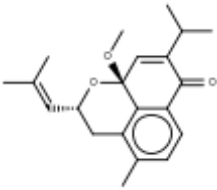
Frente a isto, das 989 moléculas cujas atividade frente ao influenza A são desconhecidas, 54 foram consideradas confiáveis pelo domínio de aplicabilidade do modelo utilizando os descritores do *Dragon*, 279 para o *CDK* e 310 moléculas ativas para o *VolSurf*, e apenas para estas a classificação da atividade foi considerada. Com o resultado da predição dos três modelos e para aumentar ainda mais a confiabilidade da predição, foi feita uma análise por consenso, em que foi considerado apenas as moléculas classificadas como ativas nos três modelos de predição.

Sendo assim, das 989 moléculas iniciais e após a análise por consenso dos três modelos de predição, apenas 36 foram classificadas como ativas e somente estas, foram submetidas as demais etapas da triagem virtual. Além disto, é necessário mencionar que, 95% das moléculas naturais preditas como ativas na análise por consenso, são alcalóides, classe esta já consagrada na literatura por suas propriedades terapêuticas, como já citado. Segue abaixo as estruturas das

moléculas classificadas como ativas (Tabela 4) nomeadas de HER01-36, com as quais foi dada continuidade a triagem virtual.

Tabela 4. Arquitetura molecular em 2D das 36 moléculas previstas como ativas.

 <p>HER01</p>	 <p>HER02</p>	 <p>HER03</p>
 <p>HER 04</p>	 <p>HER 05</p>	 <p>HER 06</p>
 <p>HER 07</p>	 <p>HER 08</p>	 <p>HER 09</p>
 <p>HER10</p>	 <p>HER11</p>	 <p>HER12</p>
 <p>HER13</p>	 <p>HER14</p>	 <p>HER15</p>
 <p>HER16</p>	 <p>HER17</p>	 <p>HER18</p>

 HER19	 HER20	 HER21
 HER22	 HER23	 HER24
 HER25	 HER26	 HER27
 HER28	 HER29	 HER30
 HER31	 HER32	 HER33
 HER34	 HER35	 HER36

OBS: Os hidrogênios foram explícitos para cada molécula.

Fonte: Próprio autor, 2021.

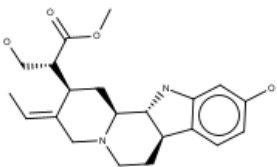
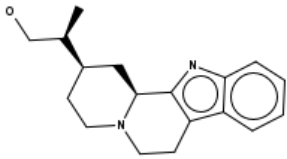
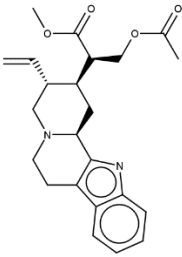
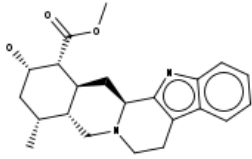
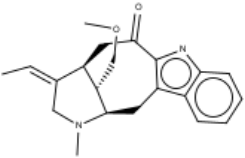
## 5.2 Análise da regra de Lipinski, TPSA e Taxa de absorção (%ABS)

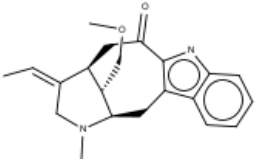
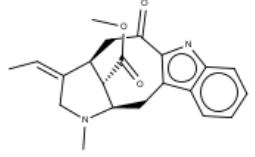
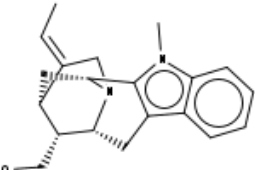
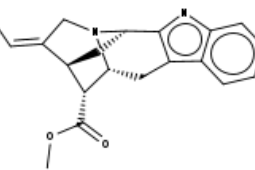
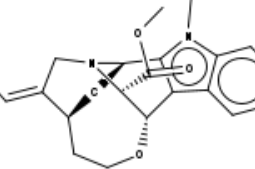
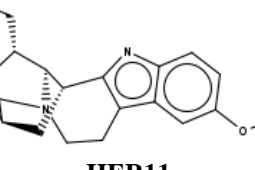
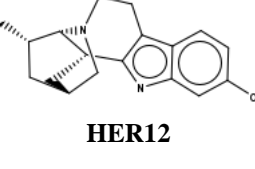
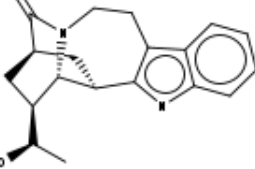
Sabendo que a maioria dos fármacos atualmente utilizados para o tratamento do influenza A são administrados por via oral, faz-se necessário analisar a biodisponibilidade oral e a taxa de absorção das moléculas candidatas. Deste modo, foi utilizado a regra dos cinco de Lipinski a fim de prever a biodisponibilidade oral das moléculas candidatas e para isso, as moléculas devem atender os seguintes requisitos: o número de doadores de ligação de

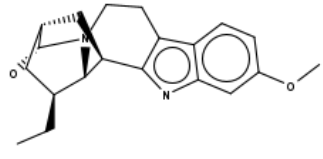
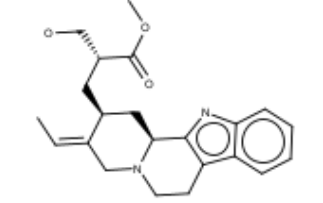
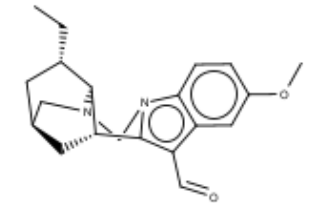
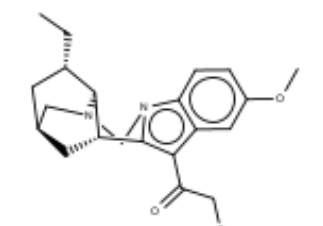
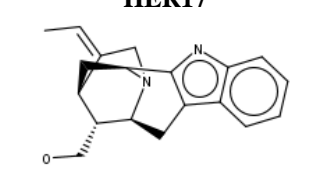
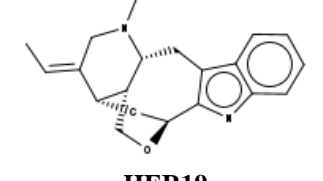
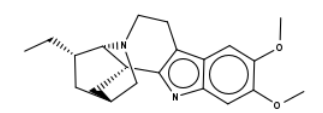
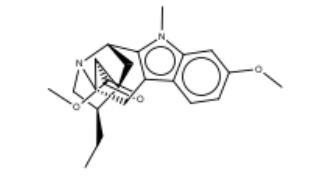
hidrogênio no ligante não deve ser maior que cinco, o número de aceptores de ligação de hidrogênio não deve ser maior do que dez, o logP não deve ser maior que cinco e o peso molecular não deve ultrapassar 500 g/mol (LIPINSKI et al., 1997).

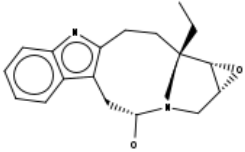
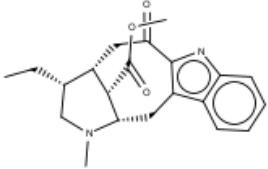
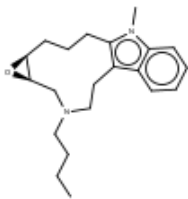
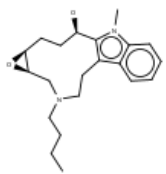
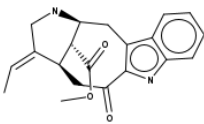
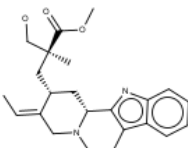
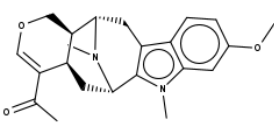
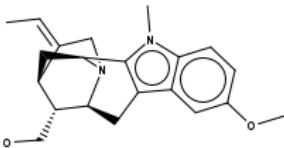
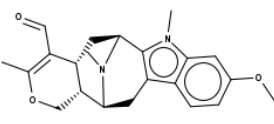
Neste sentido, os resultados diante das 36 moléculas previstas como ativas são demonstrados na tabela (5).

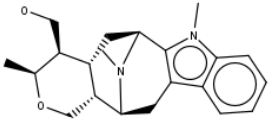
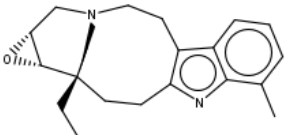
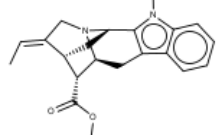
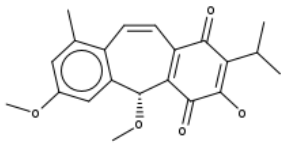
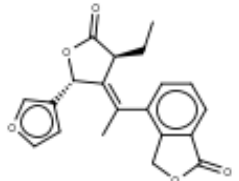
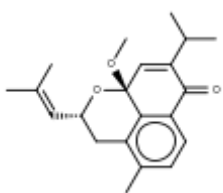
Tabela 5. Resultado dos parâmetros da regra de Lipinski, valores de TPSA e percentual de absorção das moléculas.

Moléculas	LogP	TPSA	ABS %	nDLH	nALH	P.M	nV
 <b>HER01</b>	2,15	82,03	80,70 %	3	6	372,46	0
 <b>HER02</b>	3,24	39,26	95,46 %	2	3	284,40	0
 <b>HER03</b>	4,18	71,64	84,29 %	1	6	396,49	0
 <b>HER04</b>	3,69	65,56	86,38 %	2	5	368,48	0
 <b>HER05</b>	2,98	56,33	89,57 %	2	4	324,42	0

	3,60	45,33	93.36 %	1	4	338,45	0
<b>HER06</b>							
	3,48	62,41	87.47 %	1	5	352,43	0
<b>HER07</b>							
	3,34	28,40	99.20 %	1	3	308,43	0
<b>HER08</b>							
	3,77	45,33	93.36 %	1	4	322,41	0
<b>HER09</b>							
	3,50	43,71	93.92 %	0	5	366,46	0
<b>HER10</b>							
	4,27	28,26	99.25 %	1	3	310,44	0
<b>HER11</b>							
	4,27	28,26	99.25 %	1	3	310,44	0
<b>HER12</b>							
	2,41	56,33	89.57 %	2	4	310,40	0
<b>HER13</b>							

	3,61	48,49	92.27 %	2	4	311,55	0
<b>HER14</b>							
	3,56	65,56	86.38 %	2	5	368,48	0
<b>HER15</b>							
	3,94	34,48	97.11 %	0	4	324,42	0
<b>HER16</b>							
	3,04	54,70	90.13 %	1	5	354,45	0
<b>HER17</b>							
	3,28	39,26	95.46 %	2	3	294,40	0
<b>HER18</b>							
	3,89	28,26	99.25 %	1	3	308,43	0
<b>HER19</b>							
	3,86	37,50	96.07 %	1	4	340,47	0
<b>HER20</b>							
	4,29	43,71	93.92 %	0	5	368,48	0
<b>HER21</b>							

	2,82	51,78	91.13 %	2	4	312,41	0
<b>HER22</b>							
	3,89	62,41	87.47 %	1	5	354,45	0
<b>HER23</b>							
	4,44	20,70	101.86 %	0	3	326,48	0
<b>HER24</b>							
	3,45	40,93	94.88 %	1	4	342,48	0
<b>HER25</b>							
	3,23	71,19	84.44 %	2	5	338,41	0
<b>HER26</b>							
	3,81	65,56	86.38 %	2	5	382,50	0
<b>HER27</b>							
	2,82	43,71	93.92 %	0	5	366,46	0
<b>HER28</b>							
	3,38	37,63	96.02 %	1	4	338,45	0
<b>HER29</b>							
	4,11	43,71	93.92 %	0	5	366,46	0
<b>HER30</b>							

	3,23	37,63	96.02 %	1	4	340,47	0
<b>HER31</b>							
	3,88	31,56	98.11 %	1	3	310,44	0
<b>HER32</b>							
	3,84	34,48	97.11 %	0	4	336,44	0
<b>HER33</b>							
	3,42	72,84	83.87 %	1	5	354,40	0
<b>HER34</b>							
	3,55	65,75	86.32 %	0	5	338,36	0
<b>HER35</b>							
	4,89	35,54	96.74 %	0	3	326,44	0
<b>HER36</b>							

Legendas: LogP: coeficiente de partição; TPSA: área da superfície polar topológica; ABS %: percentagem de absorção oral; nDLH: doadores de ligação hidrogênio; nALH: aceptores de ligação hidrogênio; P.M: peso molecular; nV: número de violações da regra de lipinski.

Fonte: Próprio autor, 2021.

Vale salientar que é de suma importância que os candidatos apresentem valores de logP nem tão altos e nem tão baixos, pois substâncias muito hidrofílicas tendem a permanecer em solução aquosa, se dissolvendo rapidamente e não conseguindo atravessar as membranas biológicas, ricas em lipídios. Ao contrário de substâncias muito hidrofóbicas, que se dissolvem nas membranas e nelas se acumulam, acarretando em um efeito cumulativo e gerando toxicidade, por exemplo.

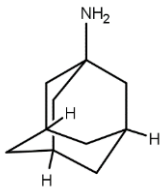
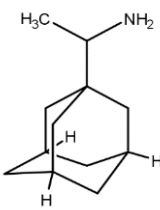
O peso molecular e a área da superfície polar topológica (TPSA) influenciam também na permeabilidade dos candidatos frente a membrana biológica, pois se os mesmos tiverem um

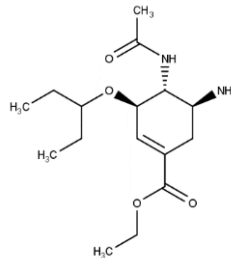
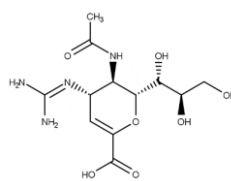
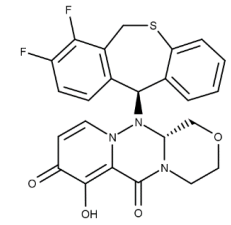
peso molecular e TPSA altos, irá ocasionar em uma redução da sua permeabilidade pela membrana, já que pressupõe que será uma molécula grande e pesada, tornando difícil sua passagem pela membrana. Já o peso molecular e TPSA baixos, aumentam a probabilidade de permeabilidade nas membranas biológicas (ULLAH et al., 2019). Além disso, a literatura também indica que candidatos com áreas de superfície polar topológica (TPSA) menores ou igual a 140 Å (ångström), apresentam melhor biodisponibilidade oral (RODRIGUES, 2019).

Diante de tais considerações, foi constatado que nenhuma das moléculas apresentaram qualquer violação da regra de Lipinski. Tanto o valor do número de doadores de ligação de hidrogênio (nDLH) quanto o de aceptores (nALH), se enquadraram na regra, além disso, o peso molecular ficou dentro dos 500 g/mol, variando de 284,40 a 396,49 g/mol, e o logP de todos os compostos ficou abaixo de 5, confirmando que tais moléculas apresentarão uma boa biodisponibilidade se administrados oralmente. Nesta perspectiva, o valor para a TPSA dos candidatos variou entre 20,70 a 80,03 Å e a partir desses valores, foi possível calcular o percentual de absorção oral teórico (%ABS), com auxílio da equação:  $\%ABS = 109 - (0.345 \times TPSA)$ .

O percentual de absorção oral para todas as moléculas variou entre 80,70 à 101,86%, constatando que em teoria, eles apresentariam uma ótima absorção pela via oral. Adicionalmente, a caráter de comparação e discutir melhor os resultados das violações da regra de Lipinski, fez-se necessário analisar também as violações dos principais fármacos comercializados para o tratamento do influenza A, servindo como padrão para a discussão dos resultados dessa pesquisa, assim temos os seguintes apresentados na tabela (6):

Tabela 6. Resultado dos parâmetros da regra de Lipinski e percentual de absorção dos fármacos controles.

<b>Fármaco</b>	<b>LogP</b>	<b>TPSA</b>	<b>ABS %</b>	<b>nDLH</b>	<b>nALH</b>	<b>P.M</b>	<b>nV</b>
 <b>Amantadina</b>	2,65	26,02	100,02%	2	1	151,25	0
 <b>Rimantadina</b>	2,46	26,02	100,02%	2	1	179,31	0

	0,85	90,66	77,72 %	3	6	312,41	0
<b>Oseltamivir</b>							
	-3,64	200,72	39,75 %	9	11	332,31	2
<b>Zanamivir</b>							
	2,44	75,01	83,12%	1	7	483,50	0
<b>Baloxavir</b>							

Legendas: LogP: coeficiente de partição; TPSA: área da superfície polar topológica; ABS %: percentagem de absorção oral; nDLH: doadores de ligação hidrogênio; nALH: aceptores de ligação hidrogênio; P.M: peso molecular; nV: número de violações da regra de lipinski.

Fonte: Próprio autor, 2021.

De acordo com os dados da tabela 6, é possível observar que 1 dos 5 fármacos utilizados como controles, apresentou 2 violações da regra de Lipinski, especificamente o Zanamivir, apresentando o número de doadores e aceptores de ligação de hidrogênio acima do permitido pela regra, além disso, tal fármaco demonstrou um percentual de absorção de apenas 39,75% por via oral. Diante de tais resultados, é possível demonstrar que todas as 36 moléculas naturais preditas como ativas apresentarão teoricamente um ótimo perfil de biodisponibilidade e absorção, se ingeridos oralmente, além de servir como estímulo para as próximas etapas da triagem.

### 5.3 Toxicidade

Durante a pesquisa e desenvolvimento de novos candidatos a fármacos, inúmeros candidatos são descartados devido a problemas relacionados principalmente com a toxicidade.

A toxicidade é um parâmetro de caráter vitalício no processo de desenvolvimento e planejamento de candidatos a fármacos, sendo um dos parâmetros da análise ADMET que junto com a eficiência da ação terapêutica é possível estimar doses e estudar a viabilidade de uso do candidato como fármaco, além de possíveis combinações de acordo com as necessidades.

Com isto, dentre as análises feitas utilizando o *Osisris Data Warrior*, foram selecionados 4 tipos de predição de riscos:

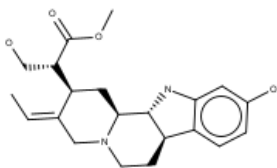
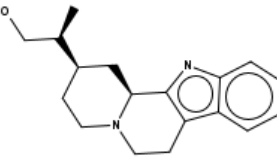
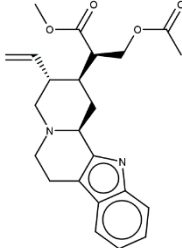
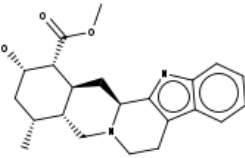
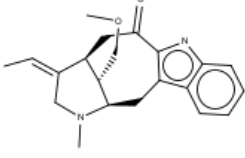
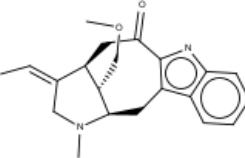
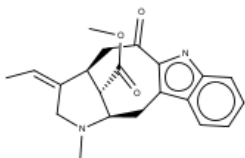
1. Mutagenicidade - capacidade que a substância tem de provocar mutações no DNA do organismo no qual está inserido;
2. Carcinogenicidade - capacidade que a substância tem em provocar o surgimento de tumores;
3. Efeito no sistema reprodutor – capacidade que a substância tem em interferir no sistema de reprodução humana;
4. Irritabilidade do tecido – capacidade que a substância tem em ocasionar quaisquer tipo de irritabilidade na pele ou tecido dos órgãos humano.

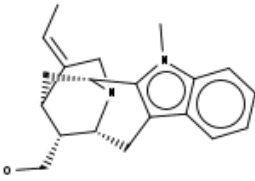
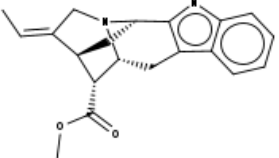
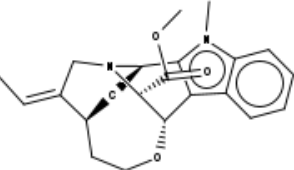
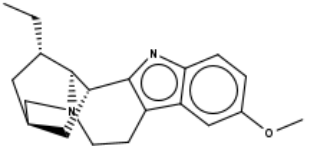
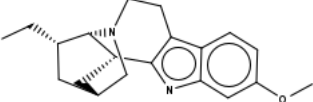
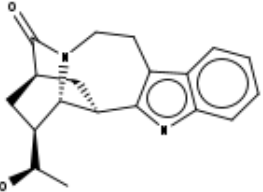
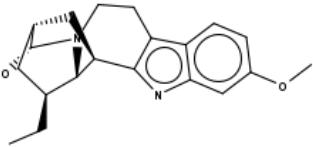
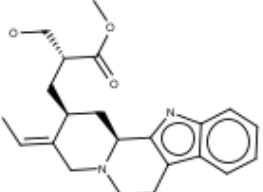
Vale salientar que a predição dos riscos no sistema reprodutor e de irritabilidade não são frequentemente analisados em pesquisas envolvendo o planejamento racional de fármacos, sem a aplicabilidade *in vitro* e dermatológica. Faz-se necessário predizer se as moléculas candidatas a fármaco poderão agir no sistema reprodutor, principalmente causando a impotência, disfunção de hormônios reprodutivos ou até mesmo infertilidade.

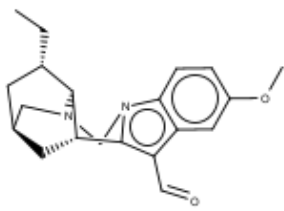
Além disso, a irritabilidade dos tecidos ou até mesmo da pele, pode provocar danos sérios a saúde do paciente que está ingerindo um medicamento, podendo ocasionar danos a mucosa dos mais diversos órgãos e com isso, acarretar em subseqüentes problemas, incluindo o óbito, dependendo da ocasião fisiológica do paciente.

Diante das observações dos parâmetros citotóxicos para esta pesquisa, foi considerado para as próximas etapas da triagem, apenas as moléculas que não apresentaram nenhum tipo de toxicidade frente aos 4 parâmetros analisados e como caráter de exclusão, as moléculas que apresentaram riscos em pelo menos 1 desses parâmetros, mesmo que fosse de natureza baixa, foram desclassificadas para serem possíveis candidatas a fármacos. Dito isto, na tabela (7) estão os dados da análise de toxicidade das 36 moléculas que continuaram na triagem virtual até essa etapa.

Tabela 7. Resultado das análises de toxicidade.

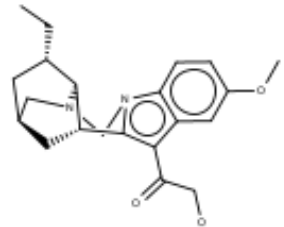
MOLÉCULAS	MUT	CAR	ESR	IRR	TOX
 <p><b>HER01</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
 <p><b>HER02</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
 <p><b>HER03</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
 <p><b>HER04</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
 <p><b>HER05</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
 <p><b>HER06</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
 <p><b>HER07</b></p>	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum

	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER08</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER09</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER10</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER11</b>					
	Nenhum	Nenhum	Alto	Nenhum	Sim
<b>HER12</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER13</b>					
	Nenhum	Nenhum	Alto	Nenhum	Sim
<b>HER14</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER15</b>					



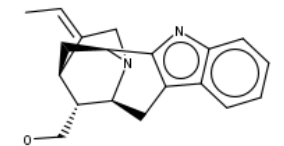
**HER16**

Nenhum    Nenhum    Nenhum    Nenhum    Nenhum



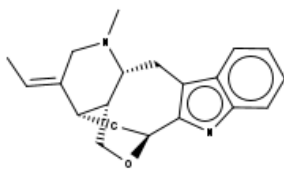
**HER17**

Nenhum    Nenhum    Nenhum    Nenhum    Nenhum



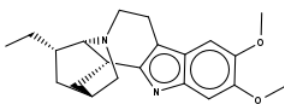
**HER18**

Nenhum    Nenhum    Nenhum    Nenhum    Nenhum



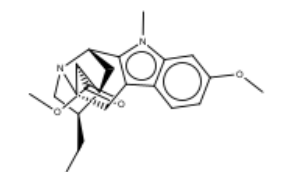
**HER19**

Nenhum    Nenhum    Nenhum    Nenhum    Nenhum



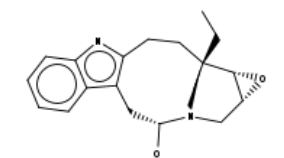
**HER20**

Nenhum    Nenhum    Alto    Nenhum    Sim



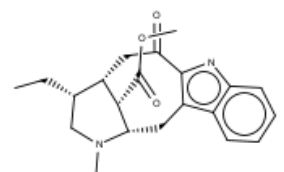
**HER21**

Nenhum    Nenhum    Alto    Nenhum    Sim



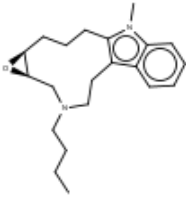
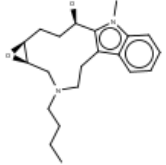
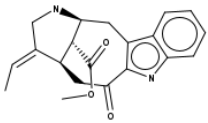
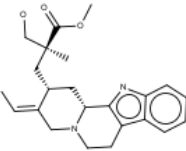
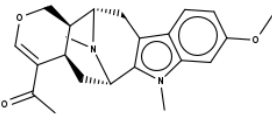
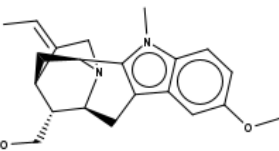
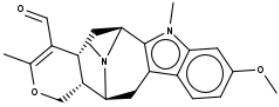
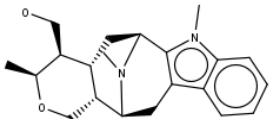
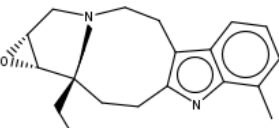
**HER22**

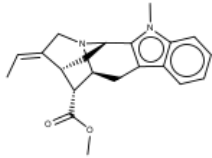
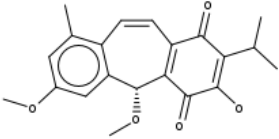
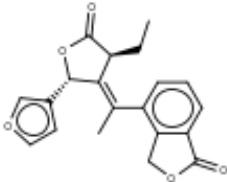
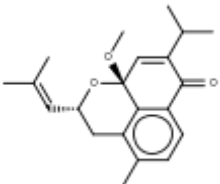
Nenhum    Nenhum    Nenhum    Nenhum    Nenhum



**HER23**

Nenhum    Nenhum    Nenhum    Nenhum    Nenhum

	Nenhum	Baixo	Nenhum	Nenhum	Sim
<b>HER24</b>					
	Nenhum	Baixo	Nenhum	Nenhum	Sim
<b>HER25</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER26</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER27</b>					
	Nenhum	Nenhum	Alto	Baixo	Sim
<b>HER28</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Não
<b>HER29</b>					
	Nenhum	Nenhum	Alto	Nenhum	Sim
<b>HER30</b>					
	Nenhum	Nenhum	Nenhum	Alto	Sim
<b>HER31</b>					
	Nenhum	Nenhum	Nenhum	Alto	Sim
<b>HER32</b>					

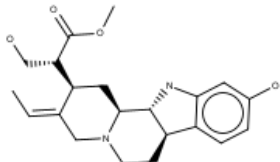
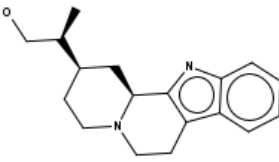
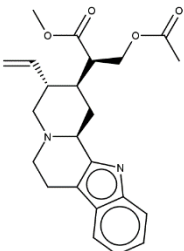
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER33</b>					
	Baixo	Nenhum	Nenhum	Nenhum	Sim
<b>HER34</b>					
	Nenhum	Nenhum	Nenhum	Nenhum	Nenhum
<b>HER35</b>					
	Nenhum	Nenhum	Nenhum	Baixo	Sim
<b>HER36</b>					

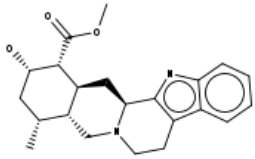
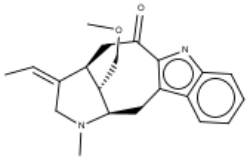
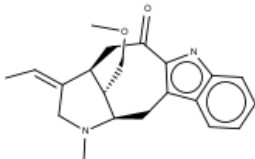
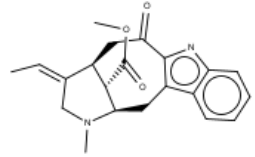
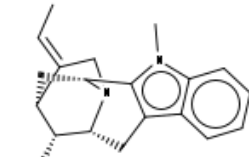
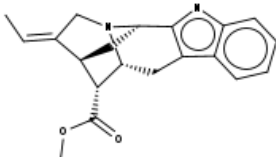
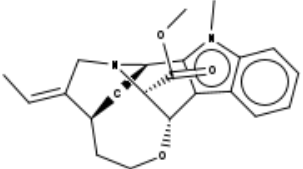
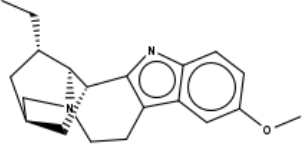
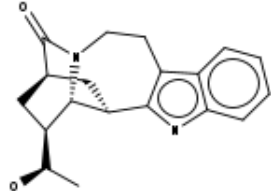
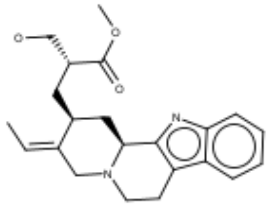
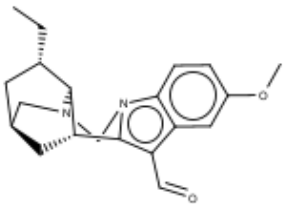
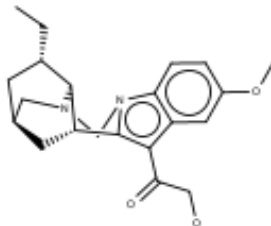
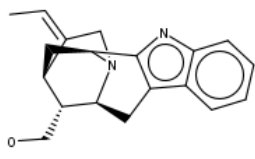
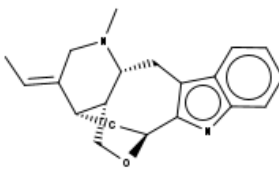
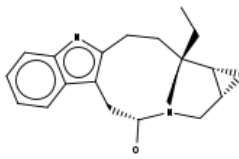
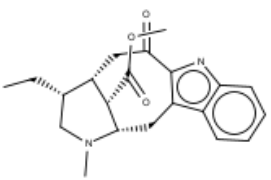
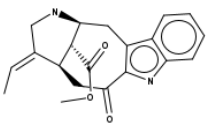
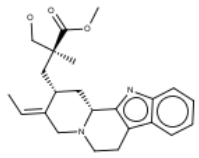
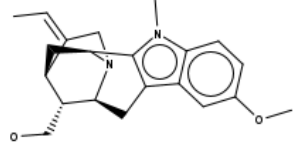
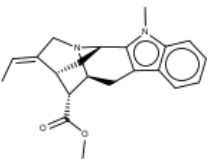
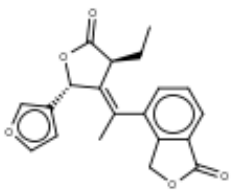
Legendas: MUT: mutagenicidade; CAR: carcinogenicidade; ESR: efeito no sistema reprodutor; IRR: irritabilidade no tecido; TOX: toxicidade geral frente a esses 4 parâmetros.

Fonte: Próprio autor, 2021.

De acordo com os dados da tabela acima, as moléculas HER12, 14, 20, 21, 24, 25, 28, 30, 31, 32, 34 e 36 apresentaram riscos em pelo menos 1 dos parâmetros de toxicidade e com base nisto, foram consideradas inviáveis para as próximas etapas da triagem virtual, tendo em vista que são parâmetros não toleráveis para um candidato a fármaco. Assim, para as próximas etapas, restaram apenas 24 moléculas, uma vez que não apresentaram potenciais mutagênicos, carcinogênicos, que venham a interferir na reprodução ou irritantes de tecido. A tabela (8) apresenta quais são as moléculas consideradas para esta análise:

Tabela 8. Arquitetura molecular em 2D das 24 moléculas restantes.

		
<b>HER01</b>	<b>HER02</b>	<b>HER03</b>

 <b>HER 04</b>	 <b>HER 05</b>	 <b>HER 06</b>
 <b>HER 07</b>	 <b>HER 08</b>	 <b>HER 09</b>
 <b>HER 10</b>	 <b>HER 11</b>	 <b>HER 13</b>
 <b>HER 15</b>	 <b>HER 16</b>	 <b>HER 17</b>
 <b>HER 18</b>	 <b>HER 19</b>	 <b>HER 22</b>
 <b>HER 23</b>	 <b>HER 26</b>	 <b>HER 27</b>
 <b>HER 29</b>	 <b>HER 33</b>	 <b>HER 35</b>

Fonte: Próprio autor, 2021.

#### 5.4 Metabolização no fígado

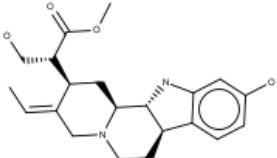
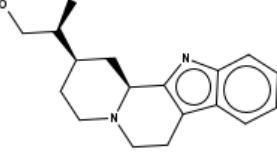
Um dos requisitos essenciais no planejamento de novos fármacos são as análises simulando a metabolização sofrida pelos candidatos no fígado, que é o principal órgão metabolizador, em conjunto com as análises das possíveis toxicidades que tais metabólitos secundários gerados após a interação com o citocromo P450 (CYP450), poderiam ocasionar no sistema fisiológico.

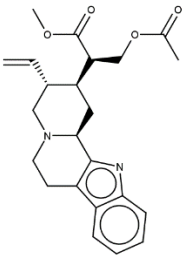
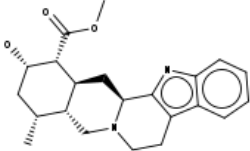
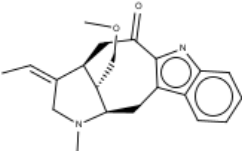
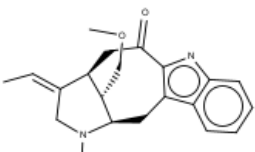
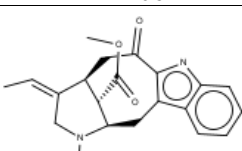
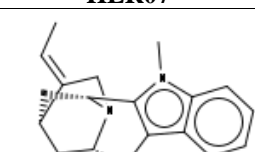
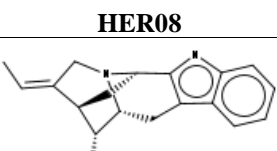
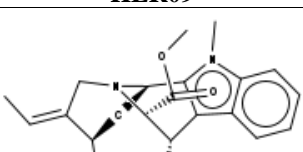
Nesta perspectiva, a próxima etapa da triagem consistiu em analisar quais são os possíveis metabólitos secundários que cada molécula geraria, após sofrer a etapa de metabolização no fígado, especificamente frente ao CYP450, através do *software MetaSite 6.0*. Além disso, após determinar quais são os metabólitos e predizer suas respectivas abundâncias no organismo, foi realizada uma análise de toxicidade de tais metabólitos, através da predição dos mesmos parâmetros analisados anteriormente, a citar: mutagenicidade, carcinogenicidade, efeito no sistema reprodutor e irritabilidade do tecido.

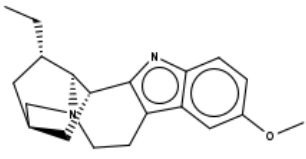
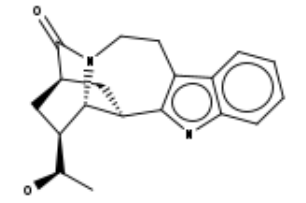
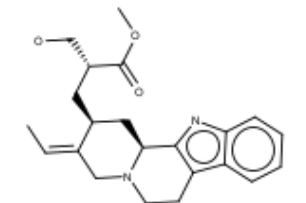
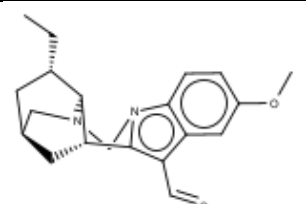
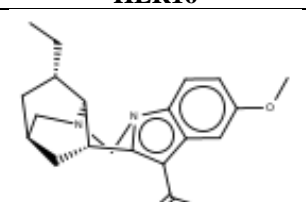
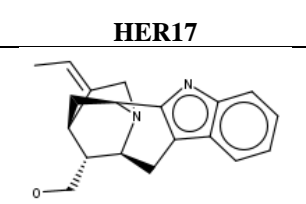
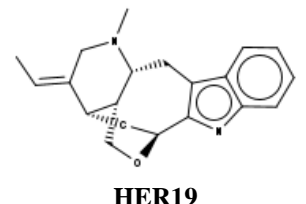
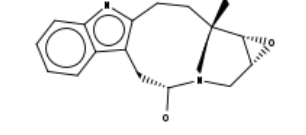
Frente aos metabólitos secundários oriundos das 24 moléculas analisadas, foram considerados apenas as moléculas que não apresentaram toxicidade em nenhum dos 4 parâmetros tóxicos citados nos seus metabólitos com abundância maior ou igual a 50%. Como caráter de exclusão, as moléculas que apresentaram seus metabólitos secundários com abundância maior ou igual a 50% e que estes, apresentarem risco de toxicidade em pelo menos 1 dos parâmetros analisados, foram descartadas.

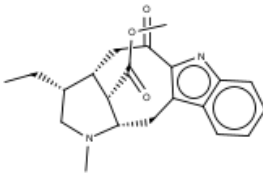
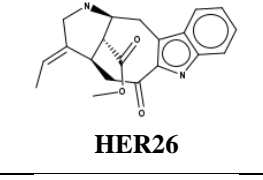
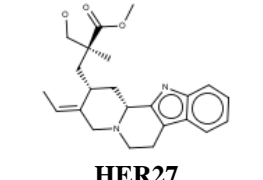
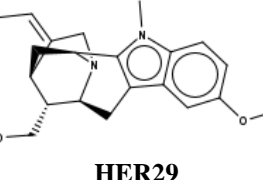
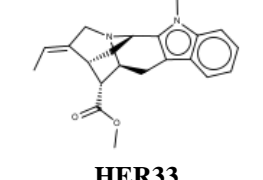
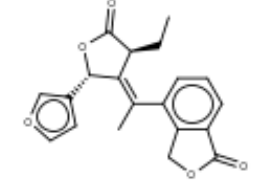
Levando em consideração que o *software* gera centenas de dados e gráficos para cada molécula analisada, optou-se por sumarizar os principais dados de cada um e expor por meio da tabela (9), conforme demonstra a seguir:

Tabela 9. Principais dados das moléculas frente ao metabolismo do CYP450.

Moléculas	nMET	ID	ABD%	TOX
 <p><b>HER01</b></p>	37	MET58 MET61	100% 55,27%	Sim
 <p><b>HER02</b></p>	25	0	0	Não

 <p><b>HER03</b></p>	33	0	0	Não
 <p><b>HER04</b></p>	30	0	0	Não
 <p><b>HER05</b></p>	66	MET173	100%	Sim
 <p><b>HER06</b></p>	5	MET 6	100%	Sim
 <p><b>HER07</b></p>	5	MET 230	100%	Sim
 <p><b>HER08</b></p>	5	MET 259	100%	Sim
 <p><b>HER09</b></p>	44	MET 288	100%	Sim
 <p><b>HER10</b></p>	31	MET 319	100%	Sim

 <p><b>HER11</b></p>	28	0	0	Não
 <p><b>HER13</b></p>	26	MET 359	100%	Sim
 <p><b>HER15</b></p>	34	MET 404	100%	Sim
 <p><b>HER16</b></p>	28	0	0	Não
 <p><b>HER17</b></p>	28	0	0	Não
 <p><b>HER18</b></p>	52	MET 488	100%	Sim
 <p><b>HER19</b></p>	14	MET 517	100%	Sim
 <p><b>HER22</b></p>	23	MET 525 MET 529	100% 100%	Sim

 <p><b>HER23</b></p>	27	MET 549	87%	Sim
 <p><b>HER26</b></p>	60	MET 592	100%	Sim
 <p><b>HER27</b></p>	32	MET 624	100%	Sim
 <p><b>HER29</b></p>	32	MET 655	100%	Sim
 <p><b>HER33</b></p>	5	MET 687	100%	Sim
 <p><b>HER35</b></p>	18	0	0	Não

Legendas: nMET: número de metabólitos formados para cada molécula; ID: id dos metabólitos que apresentaram toxicidade em pelo menos 1 dos quatro parâmetros analisados; ABD%: percentual de abundância dos metabólitos que poderão ocasionar toxicidade; TOX: toxicidade geral apresentada pelos metabólitos, avaliadas como “Sim ou Não”, sendo “Não” a molécula não apresenta nenhum tipo de metabólito abundando que venha a ocasionar toxicidade.

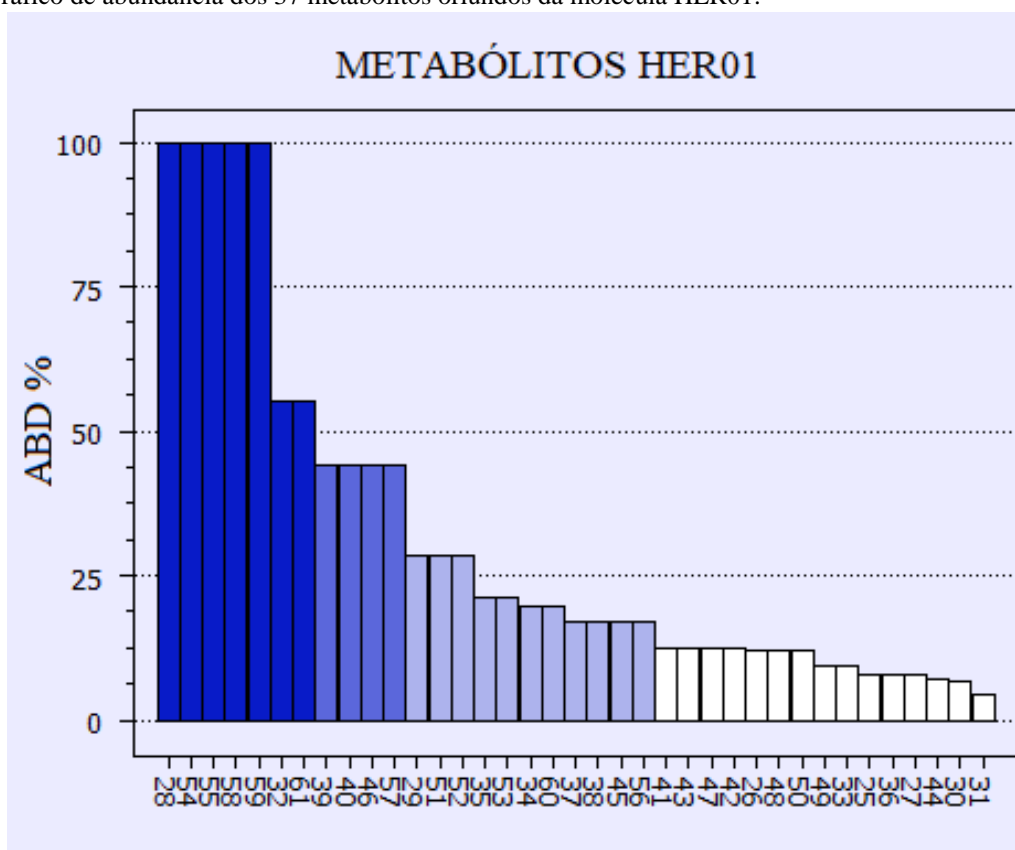
Fonte: Próprio autor, 2021.

A Tabela (9) demonstra o número de metabólitos gerados após a passagem pelo CYP450 para cada molécula, bem como ID (identificação) dos metabólitos abundantes que apresentaram toxicidade em ao menos um dos riscos analisados (mutagenicidade, carcinogenicidade, efeito no sistema reprodutor e irritabilidade na pele), seguida de seu percentual de abundância e por fim, se a molécula gerou metabólitos abundantes com algum tipo de toxicidade, avaliadas em “Sim” caso presente e “Não” caso a molécula não gere nenhum metabólito com percentual de

abundancia maior ou igual a 50% e que este não apresente toxicidade para os parâmetros avaliados.

Neste sentido, a molécula HER01, por exemplo, após sofrer a metabolização oriunda do CYP450, gerou 37 metabólitos secundários, cada qual com seu respectivo percentual de abundância no organismo (ABD%), sendo seus respectivos ID, fornecidos pelo programa, como demonstra a figura (13).

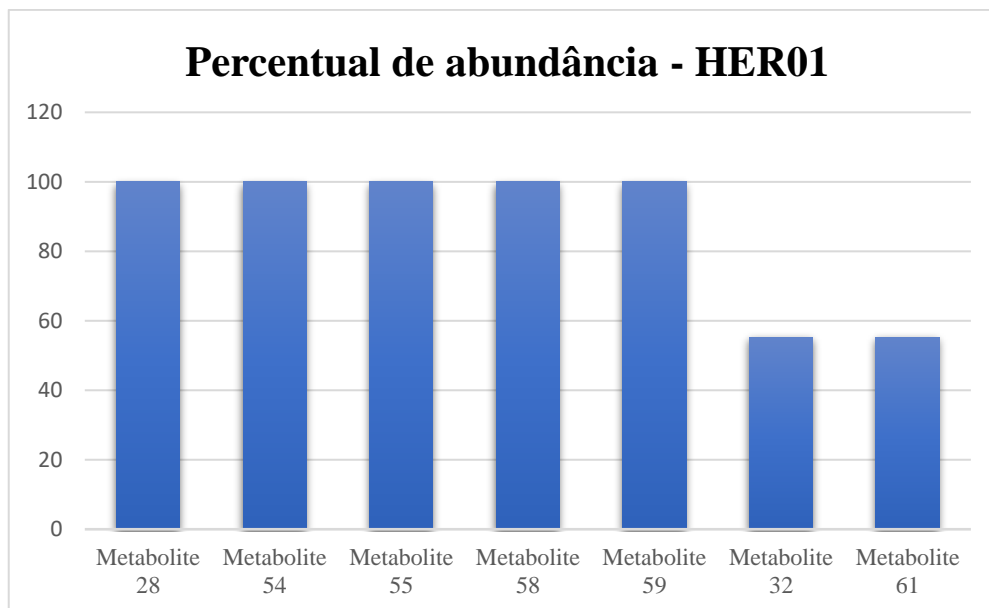
Figura 13. Gráfico de abundância dos 37 metabólitos oriundos da molécula HER01.



Fonte: Próprio autor, 2021.

E diante desses 37 metabólitos gerados, 7 apresentaram abundância maior ou igual a 50%, e destes, os metabólitos MET58 (100% de abundância) e MET61 (55,27% de abundância) apresentaram toxicidade em pelo menos um dos parâmetros analisados, conforme demonstra o gráfico (1).

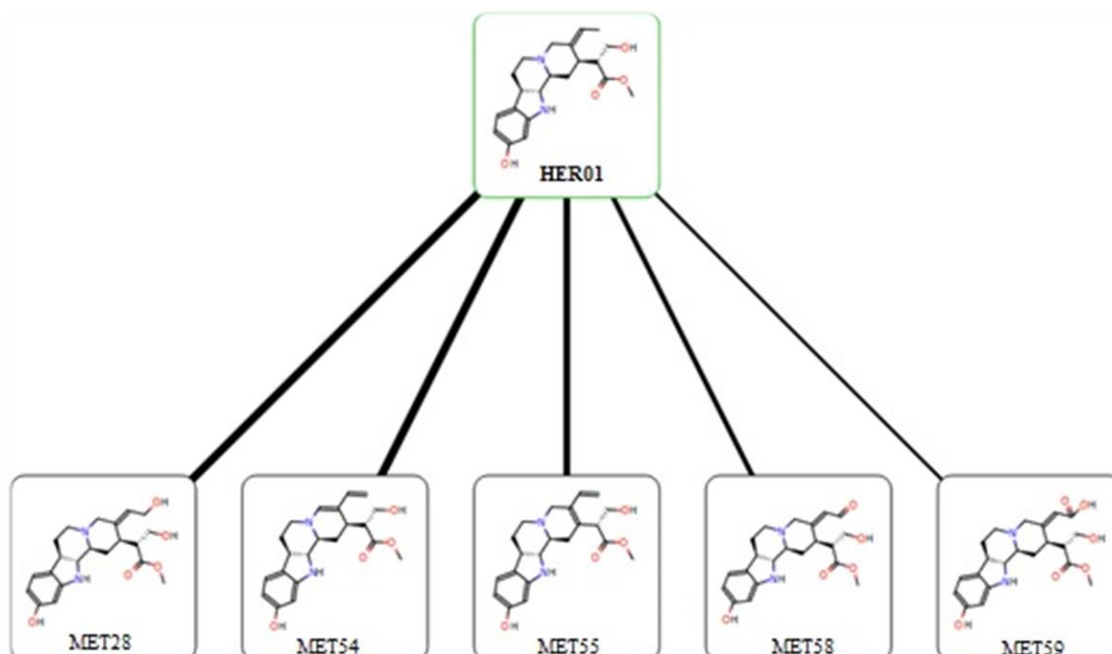
Gráfico 1. Percentual de abundância dos 7 metabólitos da molécula HER01.



Fonte: Próprio autor, 2021.

Para melhor visualização, a figura (14) por exemplo, demonstra a estrutura molecular de alguns metabólitos, oriundos da molécula HER01.

Figura 14. Estruturas em 2D de alguns dos possíveis metabólitos da molécula HERB01.



Fonte: Próprio autor, 2021.

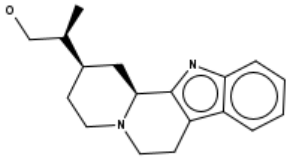
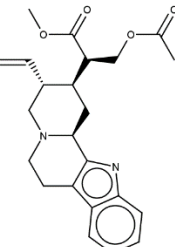
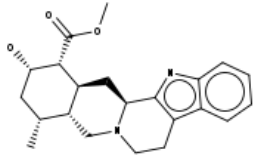
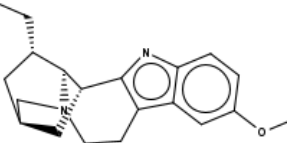
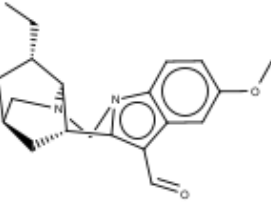
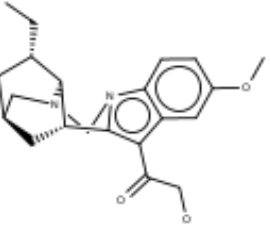
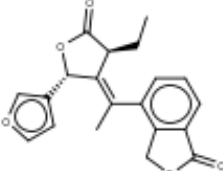
Em suma, observou-se que cerca de 70% das moléculas analisadas, apresentaram metabólitos secundários com algum tipo de toxicidade e que a maioria desses metabólitos apresentaram um percentual de abundância de 100%, ou seja, após a ingestão daquela molécula

e metabolização pelo CYP450, alguns dos metabólitos secundários mais abundantes oriundos de tal processo, será consideravelmente tóxico para o corpo humano.

Sendo assim, tais moléculas foram consideradas inviáveis para serem candidatas a fármaco, tendo em vista que não é viável um paciente ingerir um fármaco com ou sem toxicidade, mas que quando o mesmo é metabolizado pelo corpo, origine metabólitos secundários com elevada toxicidade, podendo acarretar sérios riscos ao paciente.

Frente a essas considerações, as moléculas que não apresentaram nenhum tipo de toxicidade, mediante os parâmetros analisados, em seus metabólitos abundantes (>50%) foram: HER02, HER03, HER04, HER11, HER16, HER17 e HER35, e estas irão seguir adiante no processo da triagem, sendo suas respectivas estruturas demonstrada na tabela (10).

Tabela 10. Arquitetura molecular das candidatas que não apresentaram metabólitos secundários abundantes tóxicos.

 <p style="text-align: center;"><b>HER02</b></p>	 <p style="text-align: center;"><b>HER03</b></p>	 <p style="text-align: center;"><b>HER 04</b></p>
 <p style="text-align: center;"><b>HER11</b></p>	 <p style="text-align: center;"><b>HER16</b></p>	 <p style="text-align: center;"><b>HER17</b></p>
 <p style="text-align: center;"><b>HER35</b></p>		

Fonte: Próprio autor, 2021.

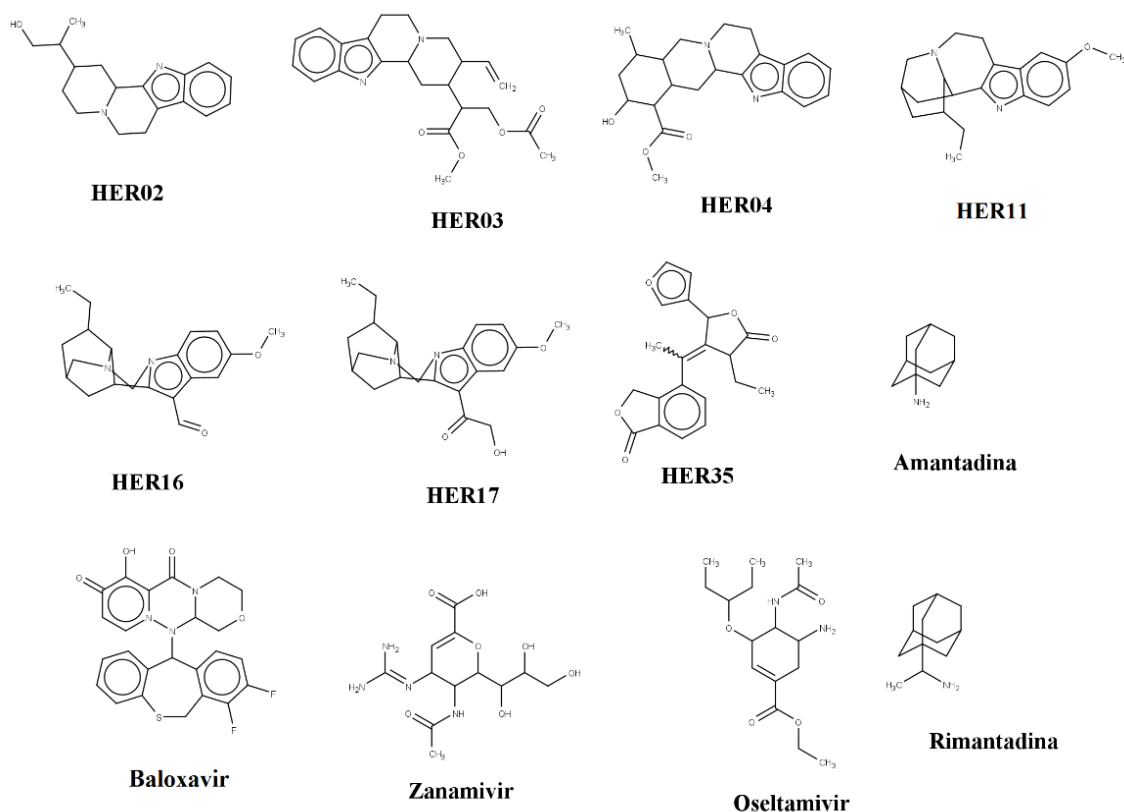
## 5.5 Docking Molecular

O docking molecular foi utilizado objetivando prever a melhor posição e orientação de um ligante em comparação a outra molécula, formando um complexo estável (PINZI &

RASTELLI, 2019). Em outras palavras, o docking molecular consiste na ancoragem para a formação do complexo ligante-receptor, em que quanto melhor for a interação para a formação desse complexo, mais baixa será sua energia total e conseqüentemente maiores as chances de a molécula apresentar atividade contra a patologia estudada.

Os fármacos atualmente utilizados para o tratamento do influenza A, visam o canal iônico formado pela proteína M2 (Amantadina e Rimantadina), a neuraminidase (oseltamivir e zanamivir) e mais recentemente, a polimerase (Baloxavir). Diante disto, vale salientar que ainda não existem fármacos que visem a hemaglutinina do influenza A. Assim, as 7 moléculas restantes e estes 5 fármacos foram submetidos aos estudos de docking e suas estruturas podem ser visualizadas na figura (15).

Figura 15. Estrutura 2D das 7 moléculas restantes + 5 fármacos controles.



Fonte: Próprio autor, 2021.

Deste modo, a ancoragem foi realizada frente as quatro proteínas envolvidas no ciclo de replicação viral do influenza A: hemaglutinina, neuraminidase, canal M2 e polimerase. Adicionalmente, foi feito o redocking molecular para cada proteína, utilizando seu próprio ligante natural. O redocking consiste em comparar a posição de um ligante complexado com a

proteína, com a posição deste mesmo ligante que foi docado no sítio ativo desta mesma proteína (NUNES et al., 2019).

Além disso, o redocking é feito a fim de analisar o valor da RMSD, que é o desvio quadrático médio para o desvio médio dos átomos de uma estrutura molecular, relativamente a uma segunda estrutura Y e quanto maior o valor, maior é a diferença estrutural entre as proteínas ou estruturas que estão a ser alvo de comparação e com isto, o *docking molecular* não apresenta confiabilidade (VELÁZQUEZ-LIBERA et al., 2020).

Conforme descrito na metodologia, sendo os principais dados do docking podem ser visualizados na tabela (11).

Tabela 11. Valores de energia (MolDockScore) das moléculas.

ID	3BKD*	1RUY*	2HTY*	6FS7*
HER02	-90,476	-12,7961	-286,723	-62,699
HER03	-125,268	-34,5391	-482,851	-112,652
HER04	-112,74	-38,4944	-450,388	-84,917
HER11	-83,518	-12,4963	-251,676	-69,915
HER16	-90,429	3,390	-181,101	-67,993
HER17	-105,59	-2,480	-205,339	-82,831
HER35	-107,821	-30,516	-317,499	-91,091
AMA	-88,702	15,843	-	-
RIM	-92,695	-21,828	-	-
OSEL	-	13,842	-396,934	-
ZAN	-	-21,8288	-310,663	-
BAL	-	-38,1105	-	-103,859
LIGANTE	-105,306	-14,114	-417,174	-81,235
RMSD	0,289	0,127	0,190	0,181

Legenda: AMA: amantadina; RIM: rimantadina; OSEL: oseltamivir; ZAN: zanamivir, BAL: baloxavir; \*: \*kcal/mol.

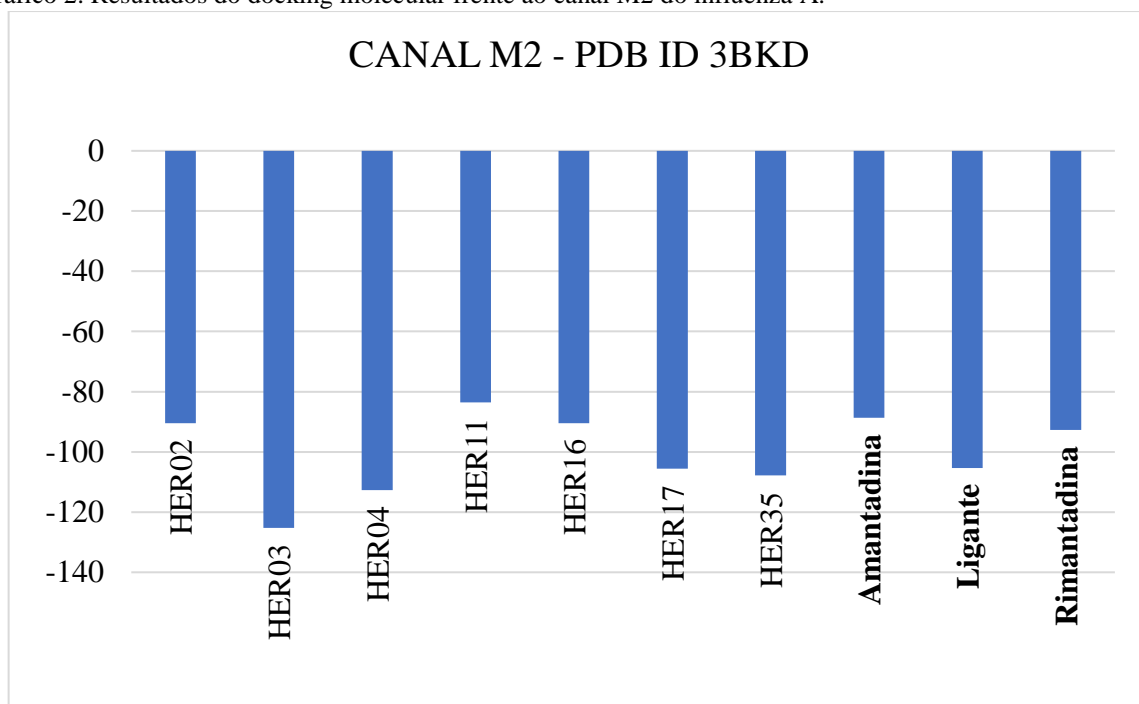
Fonte: Próprio autor, 2021.

De acordo com os dados demonstrados na tabela (11), o RMSD oriundo do redocking das 4 proteínas apresentou ótimos valores, sugeriu que a docagem molecular para todas as proteínas submetidas e moléculas alvo, apresentaram confiabilidade e validação para com os resultados. Neste sentido, a fim de melhorar o entendimento da discussão, os resultados foram divididos em basicamente 4 partes, cada qual tendo como base as proteínas submetidas a docagem molecular.

### 5.5.1 Docagem molecular do canal M2

O canal M2 (PDB - 3BKD) que possui os fármacos amantadina e rimantadina como inibidores de referência, foi analisado frente as 7 moléculas candidatas, bem como o seu ligante natural (já fornecido complexado juntamente com a proteína), conforme demonstra o gráfico (2).

Gráfico 2. Resultados do docking molecular frente ao canal M2 do influenza A.



Fonte: Próprio autor, 2021.

Neste sentido, frente ao canal M2, todas as moléculas apresentaram ótimos scores de interação comparando-se aos fármacos referência, com valores de energia variando de -125,268 à -90,476 kcal/mol, ao passo que a amantadina e rimantadina apresentaram energias de interação de -88,702 e -92,695 kcal/mol, respectivamente, não esquecendo do seu ligante natural (-105,306). Tais resultados são de suma importância, tendo em vista que, com exceção da HER11, todas as moléculas obtiveram uma afinidade melhor que os próprios fármacos utilizados para tal proteína.

Adicionalmente, a tabela (12) demonstra quais foram os resíduos de aminoácidos que as moléculas e os fármacos atuantes contra o M2, apresentaram.

Tabela 12. Interações das moléculas com os resíduos de aminoácidos do M2.

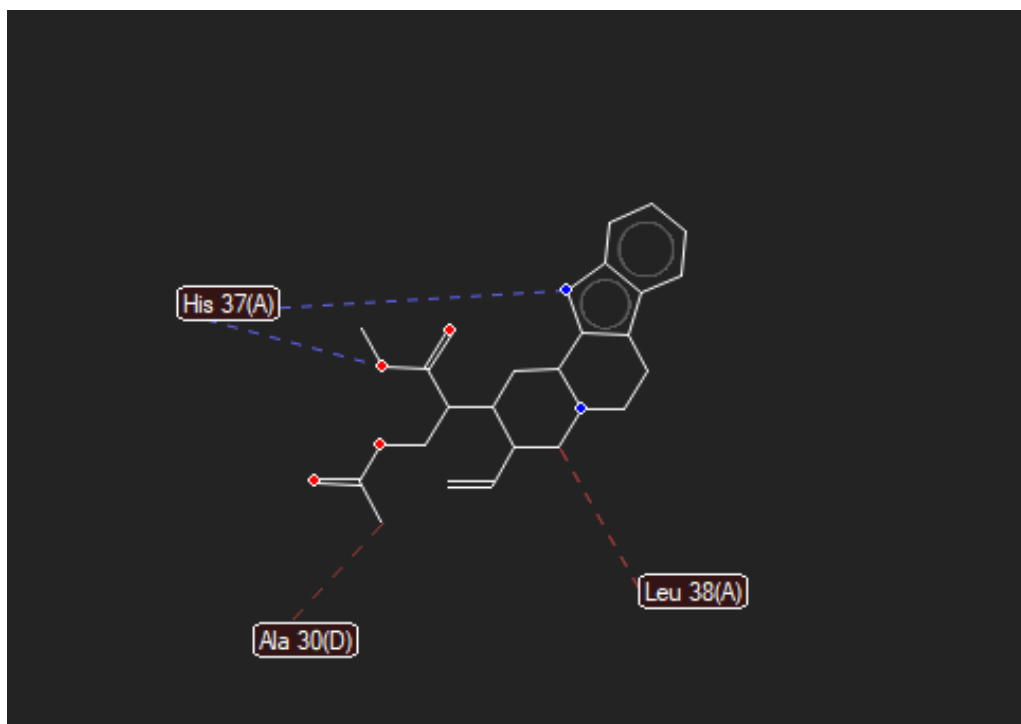
MOLÉCULA	RESÍDUOS DE AA
HER02	HIS 37
HER03	HIS 37; ALA 30; LEU 38

HER04	HIS 37; LEU 38; GLY 34 (A); GLY 34 (D)
HER11	LEU 38
HER16	LEU 38; ARG 45
HER17	LEU 38; ARG 45; HIS 37; TRP 41
HER35	HIS 37 (A); HIS 37 (D); GLY 34
AMA	HIS 37; TRP 41
RIM	HIS 37; TRP 41
LIGANTE	HIS 37

Legenda: HIS: histidina; ALA: alanina; LEU: leucina; GLY: glicina; ARG: arginina; TRP: triptofano.  
 Fonte: Próprio autor, 2021.

Frente aos dados das tabelas 11 e 12, é notável que a molécula HER03 foi a que apresentou melhor score de interação (-122,575 kcal/mol), interagindo especificamente, com os resíduos de aminoácidos His 37, este de suma importância, tendo em vista que tanto o ligante natural quanto os fármacos utilizados como controle, apresentarão a mesma interação neste resíduo de aminoácido, o que pode justificar tal score de interação para a HER03. Além disso, tal molécula interagiu com os resíduos ALA 30 e LEU 38, com interações de hidrogênio (em azul) e estéricas (em vermelho) conforme demonstra a figura (16).

Figura 16. Interações da HER03 com os resíduos de aminoácidos da M2.



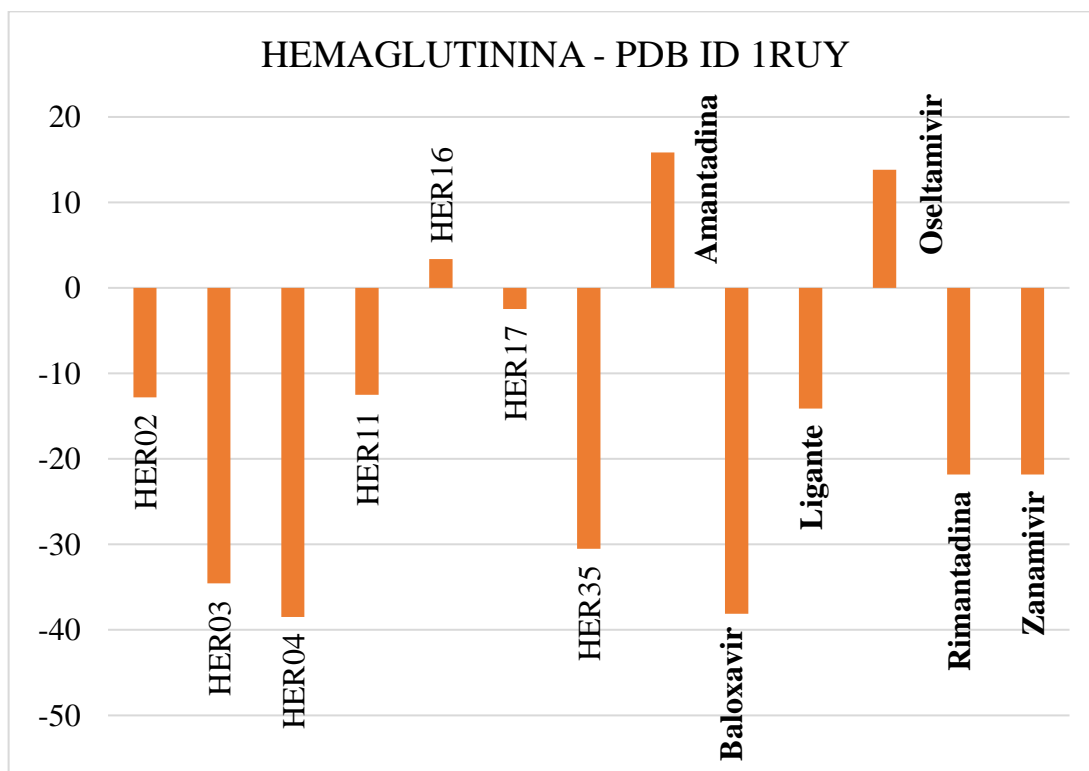
Fonte: Próprio autor, 2021.

Tais resultados são encorajadores, tendo em vista que, de acordo com Duncan et al. (2020), todas as cepas de vírus influenza A atualmente são resistentes aos “adamantanos”, principalmente devido à prevalência generalizada de uma variante M2 que codifica uma mutação de serina para asparagina 31 (S31N). Sendo assim, a molécula HER03 pode ser uma opção contra a resistência desta vírus.

### 5.5.2 Docagem molecular da hemaglutinina

Diante das análises frente a hemaglutinina (PDB - 1RUY), apesar de não existir fármacos comerciais para esta proteína, foi utilizado as energias de interação dos 5 fármacos para fins de comparação, bem como a do seu ligante natural. Nesta perspectiva, foi observado que os fármacos apresentaram um score que variou de 15,843 à -38,110 kcal/mol, lembrando que quanto mais positivo for a energia de interação, menor é aquela interação, sendo assim, dois fármacos praticamente não apresentaram interações para tal enzima, a citar: Amantadina (15,843 kcal/mol) e Oseltamivir (13,842 kcal/mol), conforme demonstra o gráfico (3).

Gráfico 3. Resultados do docking molecular frente a hemaglutinina do influenza A.



Fonte: Próprio autor, 2021.

Já frente aos resultados das moléculas, apresentaram valores de 3,390 à -38,494 kcal/mol. Basicamente a molécula HER16 não demonstrou afinidade com tal enzima, contrariando o valor demonstrado pela molécula HER04, que apresentou score de interação maior que todos os fármacos submetidos a análise, com -38,494 kcal/mol. A molécula HER03 e HER04 apresentam os maiores scores, até mesmo ao se comparar com o próprio ligante da proteína.

Tal fato, também pode ser explicado pelas interações que ocorrem os resíduos de aminoácidos da proteína, conforme demonstra a tabela (13).

Tabela 13. Interações das moléculas com os resíduos de aminoácidos da hemaglutinina.

MOLÉCULA	RESÍDUOS DE AA
HER02	ARG 224
HER03	ARG 224
HER04	ARG 224; ALA 137
HER11	ARG 224
HER16	ARG 224
HER17	ARG 224
HER35	ARG 224
AMA	-
RIM	-
OSEL	ARG 224
ZAN	ARG224
BAL	ARG224
LIGANTE	ARG 224

Legenda: ARG: arginina; ALA: alanina.

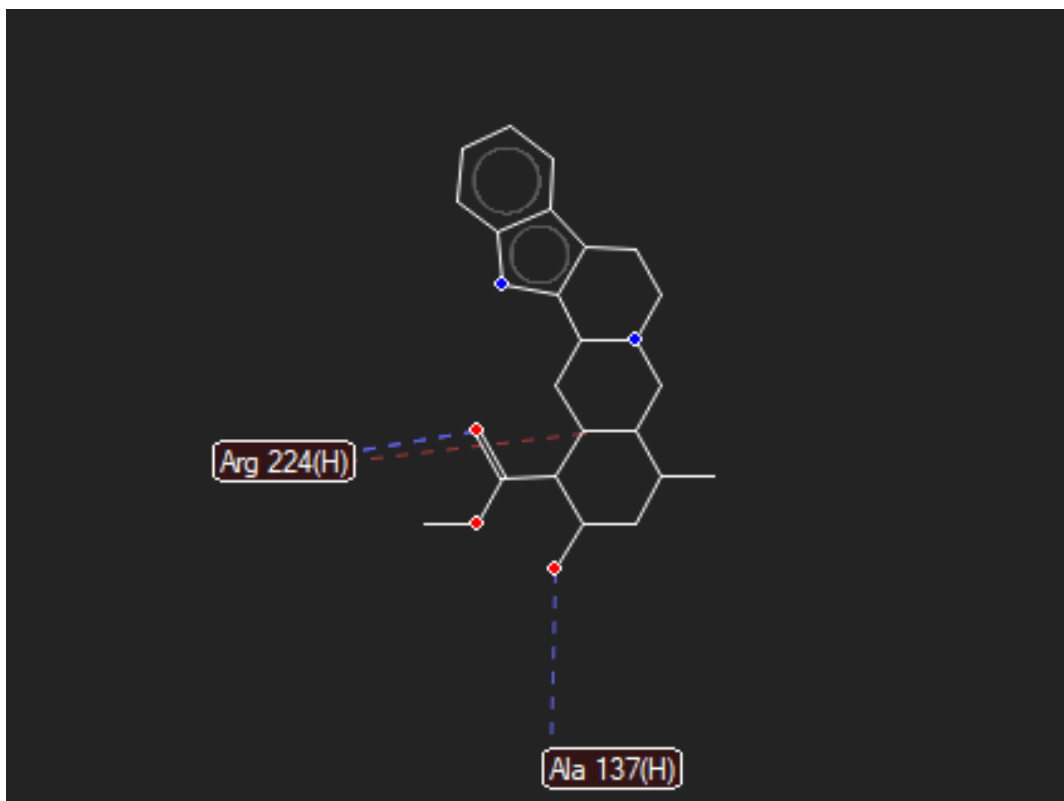
Fonte: Próprio autor, 2021.

Os resultados da tabela acima, demonstraram que o ligante da proteína apresentou interações apenas com o resíduo de aminoácido arginina (ARG 224) e tal interação, foi preservada nas demais moléculas e fármacos, com exceção da amantadina e rimantadina, que não interagiu com nenhum dos aminoácidos presentes na proteína.

Deste modo, vale destacar que as duas moléculas que apresentaram melhor score frente a tal proteína, foram a HER03 e HER04, sendo esta última com melhor pontuação e tal fato pode ser explicado pelo fato de que tal molécula foi a única que interagiu, além do resíduo ARG

224, com o resíduo ALA 137, apresentando interações de hidrogênio (azul) e estéricas (em vermelho), conforme demonstra a figura (17).

Figura 17. Interações da HER04 com os resíduos de aminoácidos da HA.

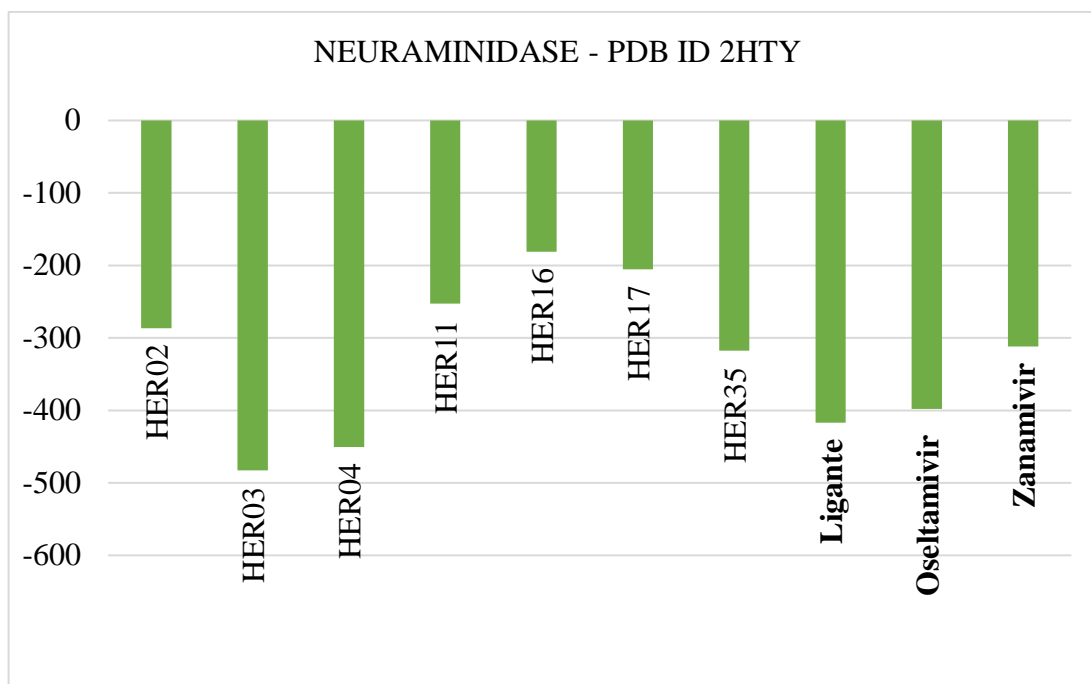


Fonte: Próprio autor, 2021.

### 5.5.3 Docagem molecular da neuraminidase

Em relação a neuraminidase (PDB - 2HTY), os fármacos utilizados como controle foram o oseltamivir e zanamivir, que apresentaram energias de interação de -396,934 e -310,663 kcal/mol, ao passo que o ligante natural de tal proteína apresentou um valor de 417,174 kcal/mol. Tais valores foram superados pelas moléculas HER03 e HER04, tendo em vista que ambas apresentaram score de interação de -482,851 e -450,388kcal/mol. Já em relação ao zanamivir, apenas uma molécula, além das citadas anteriormente, apresentou melhor score, a HER35, com -317,499kcal/mol, conforme mostra o gráfico (4).

Gráfico 4. Resultados do docking molecular frente a neuraminidase do influenza A.



Fonte: Próprio autor, 2021.

Já em relação as interações com os resíduos de aminoácidos, os dados estão sumarizados na tabela (14).

Tabela 14. Interações das moléculas com os resíduos de aminoácidos da neuraminidase.

MOLÉCULA	RESÍDUOS DE AA
HER02	ASN 146; ILE 437; THR 148
HER03	ASN 146; VAL 149; THR 148
HER04	ASN 146; ILE 437; THR 148
HER11	ASN 146; ILE 437; ILE 468
HER16	ASN 146; ILE 468
HER17	ASN 146; ILE 437; THR 148
HER35	ASN 146; ILE 468
OSEL	ASN 146
ZAN	ASN 146; ILE 437; THR 148
LIGANTE	ASN 146

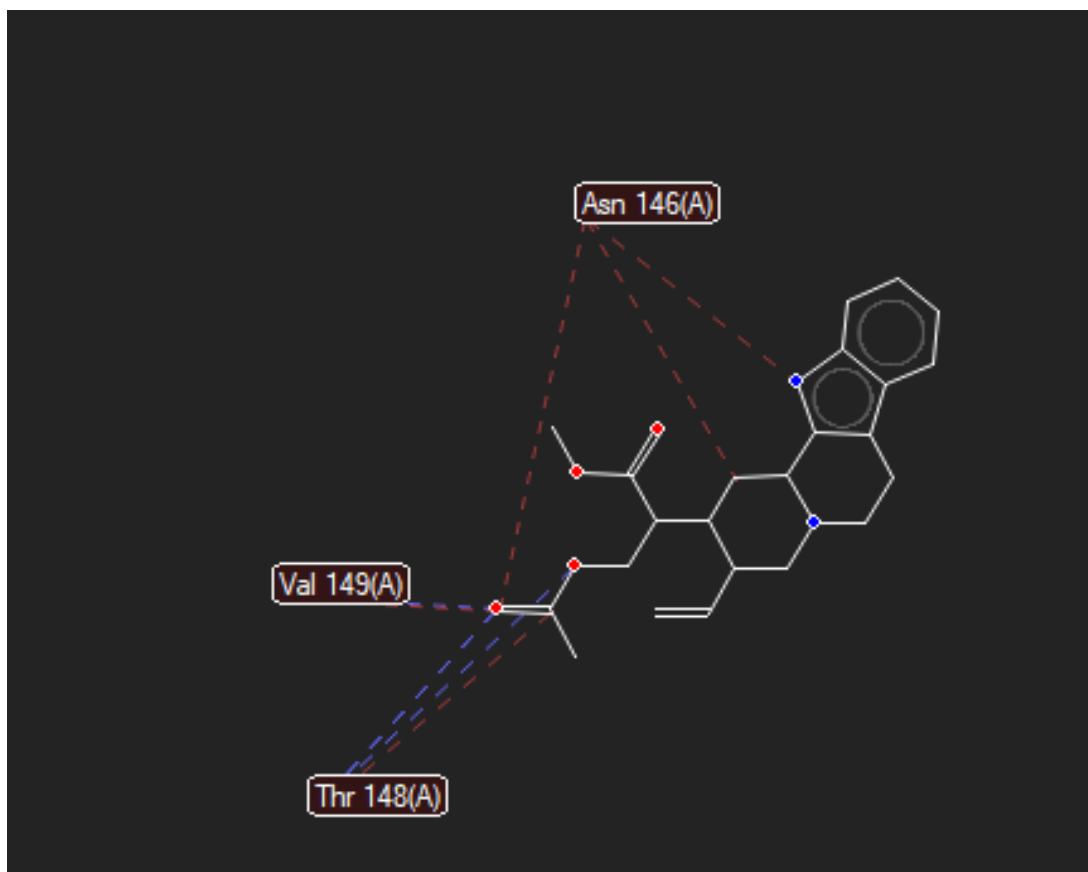
Legenda: ASN: asparagina; ILE: isoleucina; VAL: valina; THR: treonina.

Fonte: Próprio autor, 2021.

Diante dos resultados sumarizados na tabela acima, é possível observar que o ligante e o oseltamivir (controle) interagiram apenas com o resíduo de aminoácido ASN 146, e que tal interação foi preservada também para todas as moléculas e para o zanamivir (controle), ao qual apresentaram em adição outras interações com diferentes aminoácidos. A molécula HER03,

que apresentou maior “afinidade” para esta proteína, demonstrou interações com os resíduos ASN 146, VAL 149 e THR 148, conforme demonstra a figura (18).

Figura 18. Interações da HER03 com os resíduos de aminoácidos da NA.



Fonte: Próprio autor, 2021.

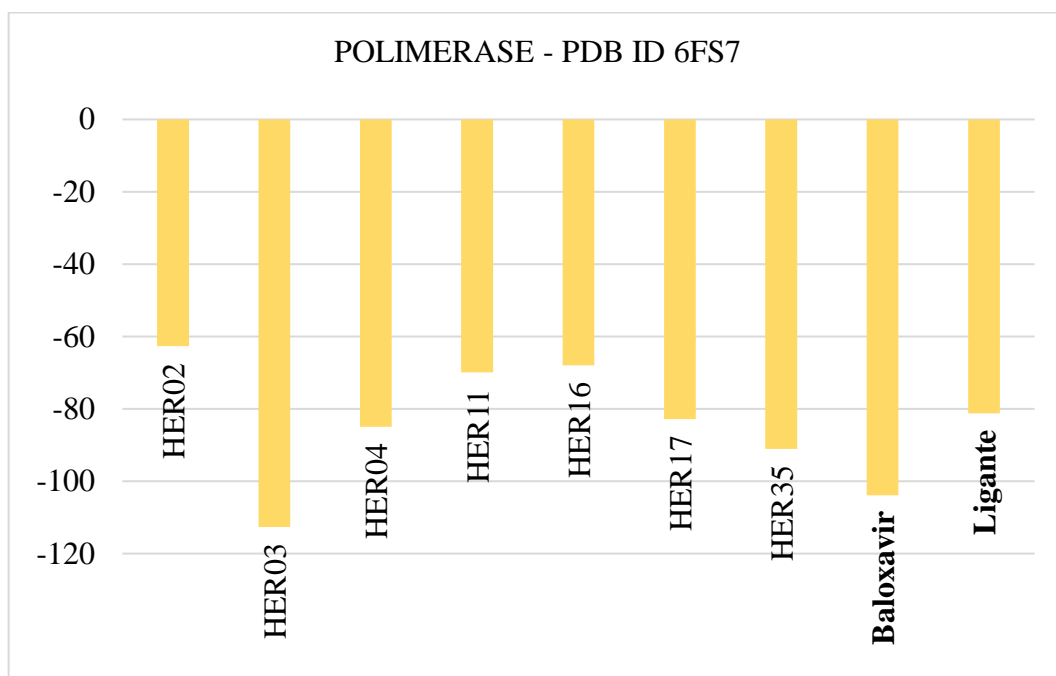
A HER03 conforme demonstra a figura acima, apresentou diversas interações de hidrogênio (em azul), estéricas (em vermelho) e eletrostáticas (em roxo), para com os resíduos de aminoácidos da neuraminidase, interações estas, de suma importância para justificar tal afinidade, uma vez que, foi a única que interagiu com a VAL 149, possivelmente esta interação foi mais forte que as outras.

#### 5.5.4 Docagem molecular da polimerase

E para a última proteína analisada, a polimerase (PDB – 6FS7), que possui como fármaco atuante, apenas o Baloxavir e tal fármaco apresentou um score de -103,859 kcal/mol, ao passo que o ligante apresentou um valor de -81,235 kcal/mol. Frente a tais resultados, o score

de interação das moléculas variou entre -62,699 à -112,652 kcal/mol, sendo este último valor oriundo da molécula HER03, conforme demonstra o gráfico (5).

Gráfico 5. Resultados do docking molecular frente a polimerase do influenza A.



Fonte: Próprio autor, 2021.

Em relação as interações com os resíduos de aminoácidos da polimerase, os dados estão sumarizados na tabela (15).

Tabela 15. Interações das moléculas com os resíduos de aminoácidos da polimerase.

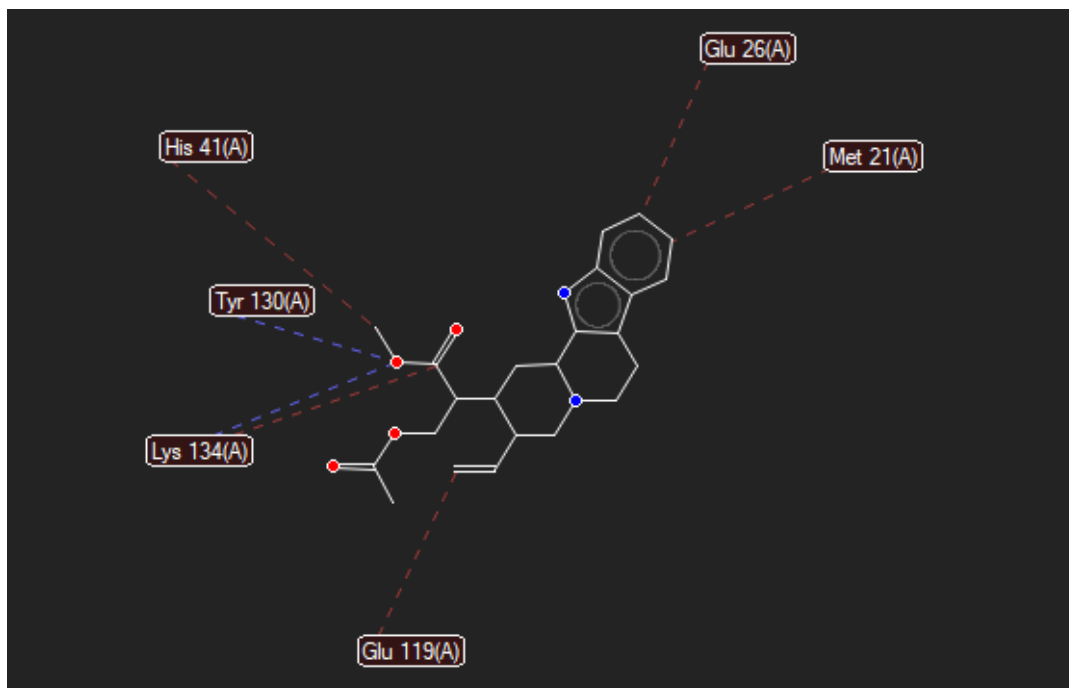
MOLÉCULA	RESÍDUOS DE AA
HER02	THR 38; ALA 37
HER03	GLU 119; GLU 26; LYS 134; TYR 130; HIS 41; MET 21
HER04	GLU 119; GLU 26; LYS 134; TYR 130; MET 21; LYS 34, HIS 41
HER11	LYS 134; LYS 104; THR 38
HER16	GLU 119; GLU 26; TYR 24; ILE 120; LYS 134; MET 21
HER17	GLU 119; GLU 198; GLU 26; LYS 134; LYS 34; THR 38; ALA 37; ILE 120; MET 21; TYR 24;
HER35	TYR 24; MET 21
BAL	GLU 119; GLU 198; GLU 80; ASP 108; LYS 134; ALA 37; HIS 41; THR 38
LIGANTE	GLU 119; GLU 80; ASP 108; LYS 134; ILE 120; HIS 41

Legenda: GLU: glutamato; ILE: isoleucina; THR: treonina; LYS: lisina; ASP: aspartato; ALA: alanina; HIS: histidina; MET: metionina.

Fonte: Próprio autor, 2021.

Frente a tabela acima, o ligante para tal proteína apresentou interações com os resíduos de aminoácido GLU 119, GLU 80, ASP 108; LYS 134; ILE 120; HIS 41, que foi preservado na maioria das moléculas analisadas e para o fármaco controle, o baloxavir. A molécula HER03 que novamente apresentou o melhor score de interação, sendo tais interações feitas nos resíduos GLU 119; GLU 26; LYS 134; TYR 130; HIS 41; MET 21, conforme demonstra a figura (19).

Figura 19. Interações da HER03 com os resíduos de aminoácidos da PA.



Fonte: Próprio autor, 2021.

Tal figura demonstra que as interações feitas pela molécula HER03 com os resíduos de aminoácidos da polimerase foram principalmente por interações de hidrogênio (azul) e estéricas (vermelho).

Com tais resultados, no geral todas as moléculas apresentaram ótimas interações frente as 4 principais proteínas do influenza A, com destaque, para a molécula HER03, um alcalóide que apresentou ótimos scores de energias de interação e para todas as 4 proteínas, com exceção da hemaglutinina, que basicamente ficou em segundo lugar no ranking dos scores mais altos, demonstrando que tal molécula apresenta promissores e potenciais resultados para ser uma candidata a fármaco anti influenza A.

## 6. CONCLUSÃO

A princípio, frente aos resultados estatísticos dos três modelos de predição de atividade biológica para o influenza A, foi possível comprovar que tais modelos apresentaram confiabilidade em prever quais foram as moléculas ativas e as inativas, para então dar continuidade ao processo da triagem virtual híbrida. Deste modo, as 36 moléculas resultantes das análises por consenso dos três modelos de predição foram então submetidas a análises da biodisponibilidade oral e o percentual da taxa de absorção, que demonstraram que todas as 36 moléculas naturais preditas como ativas apresentarão teoricamente um ótimo perfil de biodisponibilidade e absorção, se ingeridos oralmente, servindo de estímulo para as próximas etapas da triagem.

Deste modo, após as análises do risco de toxicidade, para as próximas etapas de triagem, restaram apenas 24 moléculas em potencial, uma vez que não apresentaram potenciais mutagênicos, carcinogênicos, que venham a interferir na reprodução ou irritantes de tecido. Neste sentido, tais moléculas foram submetidas a análises do metabolismo hepático, especificamente no CYP 450, em que as moléculas que não apresentaram nenhum tipo de toxicidade, mediante os parâmetros analisados, em seus metabólitos secundários abundantes (>50%) foram: HER02, HER03, HER04, HER11, HER16, HER17 e HER35, e apenas estas, seguiram adiante no processo da triagem, com a docagem molecular. Assim, na docagem molecular, no geral todas as seis moléculas apresentaram afinidade frente as quatro principais proteínas envolvidas no ciclo de replicação do influenza A.

Em suma, diante dos inúmeros resultados provenientes da triagem virtual híbrida de 986 moléculas naturais, foi demonstrado que destas, seis moléculas apresentaram potenciais para serem candidatas a fármacos frente ao influenza A, com destaque para a molécula HER03, que apresentou promissores e potenciais resultados para ser uma candidata a fármaco anti influenza A atuante em múltiplos alvos, possuindo a capacidade de bloquear mais de um alvo de tal vírus, obtendo melhor eficácia e perfis de resistência.

## REFERÊNCIAS

ABDELJOUAD, F. Z.; BRAHAMI, M.; MATTA, N. A hybrid approach for heart disease diagnosis and prediction using machine learning techniques. **ICOST**, Springer, 2020.

ABED, Y.; SAIM-MAMOUN, A.; BOIVIN, G. Fitness of influenza A and B viruses with reduced susceptibility to baloxavir: A mini-review. **Reviews in Medical Virology**, v. 31, n. 3, p. 1–11, 2021.

RAO, A. S. J.; SHETTY, N. Structure-based Assessment of Homologous Analogues of Natural products: A computational approach to predict the therapeutic effects of natural products. **Research Square**, 2021.

AHLQUIST, P. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. **Nature Reviews Microbiology**, v. 4, n. 5, p. 371–382, 2006.

ALMEIDA, R. O. (2018). **Predição de rotas metabólicas de enzimas utilizando aprendizado de máquina**. Tese (doutorado em biotecnologia) - Universidade Estadual Paulista, Instituto de Biociências de Botucatu, Botucatu-SP.

ALVES, M. S.; MEDEIROS, M. A. A.; PEREIRA, C. T.; SIMÃO, K. L. A.; SIMÃO, B. L. A.; FILHO, A. A. O. Avaliação da atividade antineoplásica e antiviral do monoterpene Ascaridol presente em plantas da caatinga: estudo *in silico*. **Revista Brasileira de Gestão Ambiental**, v. 13, n. 3, p. 23-26, 2019.

ALVES, V. M.; BRAGA, R. C.; MURATOV, E. N.; ANDRADE, C. H. Cheminformatics: An introduction. **Quimica Nova**, v. 41, n. 2, p. 202–212, 2018.

AMIN, S. A.; GHOSH, K.; GAYEN, S.; JHA, T. Chemical-informatics approach to COVID-19 drug discovery: Monte Carlo based QSAR, virtual screening and molecular docking study of some in-house molecules as papain-like protease (PLpro) inhibitors. **Journal of Biomolecular Structure and Dynamics**, v. 39, n. 13, p. 4764–4773, 2021.

ARIS-BROSOU, S.; PARENT, L.; IBEH, N. Viral long-term evolutionary strategies favor stability over proliferation. **Viruses**, v. 11, n. 8, p. 1–21, 2019.

ARRAIS, E. L.; SILVA, A. F. Pandemias na História o que há de semelhante e de novo na Covid-19 - *Sanar Medicina*. 14, 2019.

ASADI, S., GAALOUL BEN HNIA, N., BARRE, R. S., WEXLER, A. S., RISTENPART, W. D., & BOUVIER, N. M. Influenza A virus is transmissible via aerosolized fomites. **Nature Communications**, v. 11, n. 1, p. 1–9, 2020.

BAI, L.; ZHAO, Y.; DONG, J.; LIANG, S.; GUO, M.; LIU, X.; WANG, X.; HUANG, Z.; SUN, X.; ZHANG, Z.; DONG, L.; LIU, Q.; ZHENG, Y.; NIU, D.; XIANG, M.; SONG, K.; YE, J.; ZHENG, W.; TANG, Z.; XU, K. Coinfection with influenza A virus enhances SARS-CoV-2 infectivity. **Cell Research**, v. 31, n.4, p. 395–403, 2021.

BAÑOS, J. E.; GUARDIOLA, E. Historiofarmacoetimología de la morfina. **Actualidad en farmacología y terapéutica**, v. 16, n.1, p. 53-59, 2018.

BARACUHY, J. G. V.; FURTADO, D. A.; FRANCISCO, P. R. M.; LIMA, J. L. S.; PEREIRA, J. P. G. Plantas medicinais de uso comum no nordeste do Brasil. 2ª edição. Campina Grande – PB: EDUFPG, 2016.

BARROS, R. P. C. **Triagem virtual de metabólitos secundários com potencial atividade antimicrobiana do gênero *solanum* estudo fitoquímico de *solanum capsicoides***. 2017. 215f. Dissertação (Mestrado em Produtos Naturais e Sintéticos Bioativos) - Universidade Federal da Paraíba, João Pessoa – PB, 2017.

BASSO, E.; LOCATELLI, A. Plantas medicinais no ensino de ciências à luz de um “estado da arte”. **Revista de Produtos Educacionais e Pesquisas em Ensino**, v. 4, n. 2, p. 183-209, 2020.

BAUMEISTER, E. G.; PONTORIERO, A. V. Influenza Viruses, Biology, Epidemiology, and Control. **Springer Cham**, p. 255–65, 2017.

BEDFORD, J.; FARRAR, J.; IHEKWEAZU, C.; KANG, G.; KOOPMANS, M.; NKENGASONG, J. A new twenty-first century science for effective epidemic response. **Nature**, v. 575, n. 7781, p. 130–136, 2019.

BEZERRAa, L. C. C.; QUEIROZ, E. W. A.; FREIRE, J. E. C. Predição físico-química, modelagem e análise do mecanismo de interação da quitinase mo-chi1[moringaoleifera, lam.],

com poli- $\beta$ -(1-4)-n-acetil-d-glucosamine: uma abordagem *in silico*. **Revista Desafios**, v. 5, n. 1, p. 1-16, 2018.

BEZERRA<sup>b</sup>, J. J. L. Plantas potencialmente medicinais do gênero solanum l. (solanaceae): uma revisão. Anais III CONBRACIS... Campina Grande: Realize Editora, 2018. Disponível em: <<https://editorarealize.com.br/artigo/visualizar/41286>>, acesso em 10 de dezembro de 2021.

BISSO-ANDRADE, A. Resistencia a los antimicrobianos Resistance to antimicrobials. **Rev Society Peru Medicinal Interna**, v. 31, n. 2, p. 50–59, 2018.

BORGES, L. P.; AMORIM, V. A. Metabólitos secundários de plantas. **Revista Agrotecnologia**, v.11, n.1, p.54-67, 2020.

BORTOLUZZI, M. M.; SCHMITT, V.; M, C. E. Efeito fitoterápico de plantas medicinais sobre a ansiedade: uma breve revisão. **Research, Society and Development**, v. 9, n.1, p. 1-13, 2019.

BRASIL. Agência Nacional de Vigilância Sanitária (ANVISA). Lista DCB Plantas Medicinais atualizada em dezembro 2017. 2017. Brasília. Disponível em <[http://portal.anvisa.gov.br/resultado-debusca?p\\_p\\_id=101&p\\_p\\_lifecycle=0&p\\_p\\_state=maximized&p\\_p\\_mode=view&p\\_p\\_col\\_id=column1&p\\_p\\_col\\_count=1&\\_101\\_struts\\_action=%2Fasset\\_publisher%2Fview\\_content&\\_101\\_assetEntryId=3881905&\\_101\\_type=document](http://portal.anvisa.gov.br/resultado-debusca?p_p_id=101&p_p_lifecycle=0&p_p_state=maximized&p_p_mode=view&p_p_col_id=column1&p_p_col_count=1&_101_struts_action=%2Fasset_publisher%2Fview_content&_101_assetEntryId=3881905&_101_type=document)>, acesso em 07 de outubro de 2021.

BROWN, F. K. Chemoinformatics: what is it and how does it impact drug discovery. **Annual Reports in Medicinal Chemistry**, v. 33, p.375–384, 1998.

CARACCIOLO, G.; FAROKHZAD, O. C.; MAHMOUDI, M. Biological identity of nanoparticles in vivo: clinical implications of the protein corona. **Trends in Biotechnology**, v. 35, n. 3, p. 257–264, 2017.

CECE-ESENCAN, E. N.; FONTAINE, F.; PLASENCIA, G.; TEPPNER, M.; BRINK, A.; PÄHLER, A.; ZAMORA, I. Software-aided cytochrome P450 reaction phenotyping and kinetic analysis in early drug discovery. **Rapid Communications in Mass Spectrometry**, v. 30, n. 2, p. 301–310, 2016.

Centers for Disease Control and Prevention - CDC. Section 11: Epidemic Disease Occurrence. In U.S. Department of Health & Human Services (Ed.), *Principles of Epidemiology in Public Health Practice - An Introduction to Applied Epidemiology and Biostatistics*, p. 1–6, 2020.

CHEN, Y.; JIANG, Z.; LEI, Z.; PING, J.; SU, J. Effect of rifaximin on gut-lung axis in mice infected with influenza A virus. **Comparative Immunology, Microbiology and Infectious Diseases**, v. 75, 2021.

CHENG, Y.; MA, J.; WANG, H.; WANG, X.; HU, Z.; LI, H.; ZHANG, H.; LIU, X. Co-infection of influenza A virus and SARS-CoV-2: A retrospective cohort study. **Journal of Medical Virology**, v. 93, n. 5, p. 2947–2954, 2021.

CHOWAŃSKI, S.; ADAMSKI, Z.; MARCINIAK, P.; ROSIŃSKI, G.; BÜYÜKGÜZEL, E.; BÜYÜKGÜZEL, K.; FALABELLA, P.; SCRANO, L.; VENTRELLA, E.; LELARIO, F.; BUFO, S. A. A review of bioinsecticidal activity of Solanaceae alkaloids. **Toxins**, v. 8, n. 3, p. 60, 2016.

CLEMENTI, N.; GHOSH, S.; DE SANTIS, M.; CASTELLI, M.; CRISCUOLO, E.; ZANONI, I.; CLEMENTI, M.; MANCINI, N. Viral respiratory pathogens and lung injury. **Clinical Microbiology Reviews**, v.34, n. 3, p. 1–45, 2021.

COLEY, C. W.; BARZILAY, R.; GREEN, W. H.; JAAKKOLA, T. S.; JENSEN, K. F. Convolutional Embedding of Attributed Molecular Graphs for Physical Property Prediction. **Journal of Chemical Information and Modeling**, v. 57, n. 8, p. 1757-1772, 2017.

COMAS-GARCIA, M. Packaging of genomic RNA in positive-sense single-stranded RNA viruses: a complex story. **Viruses**, v. 11, n. 3, p. 253, 2019.

COSTA, L. M. C. D. A.; MERCHAN-HAMANN, E. Pandemias de influenza e a estrutura sanitária brasileira: breve histórico e caracterização dos cenários. **Revista Pan-Amazônica de Saúde**, v. 7, n. 1, p. 11–25, 2016.

CUI, Y.; ZHANG, Y.; ZHOU, K.; SUN, J.; ZHOU, Z. H. Conservative transcription in three steps visualized in a double-stranded RNA virus. **Nature Structural & Molecular Biology**, v. 26, n. 11, p. 1023–1034, 2019.

DAR, K. B.; BHAT, A. H.; AMIN, S.; HAMID, R.; ANEES, S.; ANJUM, S.; RESHI, B. A.; ZARGAR, M. A.; MASOOD, A.; GANIE, S. A. Modern Computational Strategies for Designing Drugs to Curb Human Diseases: A Prospect. **Current Topics in Medicinal Chemistry**, v. 18, n. 31, p. 2702–2719, 2019.

DAS, S.; ALAM, M. M.; ZHANG, R.; HISANO, S.; SUZUKI, N. Proof of Concept of the Yadokari Nature: a Capsidless Replicase-Encoding but Replication-Dependent Positive-Sense Single-Stranded RNA Virus Hosted by an Unrelated Double-Stranded RNA Virus. **Journal of Virology**, v. 95, n.17, 2021.

DAVIDSON, S. Treating influenza infection, from now and into the future. **Frontiers in Immunology**, v. 9, p. 1–14, 2018.

DIETZ, C.; RUEDEN, C. T.; HELFRICH, S.; DOBSON, E. T. A.; HORN, M.; EGLINGER, J.; EVANS, E. L.; MCLEAN, D. T.; NOVITSKAYA, T.; RICKE, W. A.; SHERER, N. M.; ZIJLSTRA, A.; BERTHOLD, M. R.; ELICEIRI, K. W. Integration of the ImageJ Ecosystem in KNIME Analytics Platform. **Frontiers in Computer Science**, v. 2, p. 1–17, 2020.

DOGRU, S.; SUBASI, A. Traffic accident detection using random forest classifier. **15th Learning and Technology Conference**, p. 40-45, 2018.

DOU, D.; REVOL, R.; ÖSTBYE, H.; WANG, H.; DANIELS, R.; DANIELS R. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. **Frontiers in Immunology**, v. 9, n. 1581, p.1–17. 2018.

DUNCAN, M. C.; ONGUÉNÉ, P. A.; KIHARA, I.; NEBANGWA, D. N.; NAIDU, M. E.; WILLIAMS, D. E.; BALGI, A. D.; ANDRAE-MAROBELA, K.; ROBERGE, M.; ANDERSEN, R. J.; NIIKURA, M.; NTIE-KANG, F.; TIETJEN, I. Virtual screening identifies chebulagic acid as an inhibitor of the M2(S31N) viral ion channel and influenza a virus. **Molecules**, v. 25, n. 12, 2020.

FALCÓN-CANO, G.; CABRERA-PÉREZ, M. Á.; MOLINA, C. (2020). ADME prediction with KNIME: In silico aqueous solubility consensus model based on supervised recursive random forest approaches. **ADMET and DMPK**, v. 8, n. 3, p. 251–273, 2020.

FARIA, N. R.; AZEVEDO, S. S. R.; KRAEMER, M. U. G.; SOUZA, R.; CUNHA, M. S.; HILL, S. C.; THÉZÉ, J.; BONSALE, M. B.; BOWDEN, T. A.; RISSANEN, I.; ROCCO, I. M.;

NOGUEIRA, J. S.; MAEDA, A. Y.; DA SILVA VASAMI, F. G.; DE LIMA MACEDO, F. L.; SUZUKI, A.; RODRIGUES, S. G.; CRUZ, A. C. R.; NUNES, B. T.; VASCONCELOS, P. F. C. Zika virus in the Americas: Early epidemiological and genetic findings. **Science**, v. 352, n. 6283, p. 345–349, 2016.

FASHOTO, S. G.; METFULA, A. S.; MATSEBULA, B. B.; FASHOTO, B. Y. Multi-Target Regression Prediction on Cervical Cancer for. **Asian Journal of Information Technology**, v. 17, n. 2, p. 160–166, 2018.

FERNANDA, G.; ANDRADE, V. R. M. The Influenza virus : narrative literature review. **Revista Interdisciplinar em Ciências da Saúde e Biológicas**, v. 3, n. 2, p. 74–82, 2019.

FIELDING, B. C.; FILHO, C. S. M. B.; ISMAIL, N. S. M.; SOUSA, D. P. Alkaloids: Therapeutic Potential against Human Coronaviruses. **Molecules**, v. 25, p. 5496, 2020.

FILLBRUNN, A.; DIETZ, C.; PFEUFFER, J.; RAHN, R.; LANDRUM, G. A.; BERTHOLD, M. R. KNIME for reproducible cross-domain analysis of life science data. **Journal of Biotechnology**, v. 261, p 149–156, 2017.

FLERLAGE, T.; BOYD, D. F.; MELIOPOULOS, V.; THOMAS, P. G. Influenza virus and SARS-CoV-2 : pathogenesis and host responses in the respiratory tract. **Nature Reviews Microbiology**, v. 19, p. 425–441, 2021.

FORTI, M. **Técnicas de machine learning aplicadas na recuperação de crédito do mercado brasileiro**. 2018. 74f. Dissertação (Mestre em economia), Escola de economia de São Paulo, São Paulo - SP.

FOWLER, S.; MORCOS, P. N.; CLEARY, Y.; MARTIN-FACKLAM, M.; PARROTT, N.; GERTZ, M.; YU, L. Progress in Prediction and Interpretation of Clinically Relevant Metabolic Drug-Drug Interactions: a Minireview Illustrating Recent Developments and Current Opportunities. **Current Pharmacology Reports**, v. 3, n. 1, p. 36-49, 2017.

FUKUHARA, T. Endornaviruses: persistent dsRNA viruses with symbiotic properties in diverse eukaryotes. **Virus Genes**, v. 55, n. 2, p. 165–173, 2019.

GAMBLIN, S. J., HAIRE, L. F., RUSSELL, R. J., STEVENS, D. J., XIAO, B., HA, Y., VASISHT, N., STEINHAEUER, D. A., DANIELS, R.S., ELLIOT, A., WILEYAND, D.

C., SKEHEL, J. J. The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin. **Science**, v. 303, n. 5665, p. 1838-1842, 2004.

GARCÍA, J. P.; FERREIRA, J. C.; PATINO, C. M. Análise ROC: uma aliada na pandemia. **Journal Brasileiro de Pneumologia.**, v. 47, n. 2, 2021.

GIRE, S. K.; GOBA, A.; ANDERSEN, K. G.; SEALFON, R. S. G.; PARK, D. J.; KANNEH, L.; JALLOH, S.; MOMOH, M.; FULLAH, M.; DUDAS, G.; WOHL, S.; MOSES, L. M.; YOZWIAK, N. L.; WINNICKI, S.; MATRANGA, C. B.; MALBOEUF, C. M.; QU, J.; GLADDEN, A. D.; SCHAFFNER, S. F.; LANDER, E. S. During the 2014 Outbreak. **Science**, v. 345, n. 6202, p. 1369–1372, 2014.

GOMES, J. N. S. **Estudo *in silico* de derivados acridínicos com potencial atividade antitumoral.** 2019. 108 f. Trabalho de Conclusão de Curso (Graduação em Farmácia) - Centro de Ciências Biológicas e da Saúde, Universidade Estadual da Paraíba, Campina Grande-PB, 2019.

GOIS, M. A. F.; LUCAS, F. C. A.; COSTA, J. C. M.; MOURA, P. H. B.; LOBATO, G. J. M. Etnobotânica de espécies vegetais medicinais no tratamento de transtornos do sistema gastrointestinal. **Revista Brasileira de Plantas Medicinai**s, v. 18, n. 2, p. 547-557, 2016.

GONZÁLEZ-CHAVARRO, C. F.; GUTIERREZ, M. C.; PULIDO-BLANCO, V. C.; RUIZ, X. M. C. Amaryllidaceae: fuente potencial de alcaloides. **Ciencia y Agricultura**, v. 17, n. 3, p. 78-94, 2020.

GUERRA, T. M. **Estudos de docking molecular de derivados da tiazolidina como potenciais inibidores da enzima cruzaina de trypanosoma cruzi.** 2019. 58 f. Trabalho de Conclusão de Curso (Licenciatura em Química) – Unidade Acadêmica de Serra Talhada, Universidade Federal Rural de Pernambuco, Serra Talhada-PE, 2019.

GUREVICH, K. G.; URAKOV, A. L.; ROZIT, G. A.; KLEN, E.; SAMORODOV, A. V.; KHALIULLIN, F. A. Synthesis and Antiplatelet and Anticoagulant Activity of Thietane-Containing 2-(5-Bromo-2,4-Dihydro-3-Oxo-1,2,4-Triazolyl-4)Acetate Salts. **Pharmaceutical Chemistry Journal**, v. 55, n. 5, p. 417–422, 2021.

GURUNG, A. B.; ALI, M. A.; LEE, J.; FARAH, M. A.; AL-ANAZI, K. M. An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. **BioMedical Research International**, 2021.

HANSMANN, S.; MIYAJI, Y.; DRESSMAN, J. An in silico approach to determine challenges in the bioavailability of ciprofloxacin, a poorly soluble weak base with borderline solubility and permeability characteristics. **European Journal of Pharmaceutics and Biopharmaceutics**, v.122, p. 186–196, 2018.

HARSHBARGER, W. D.; DEMING, D.; LOCKBAUM, G. J.; ATTATIPPAHOLKUN, N.; KAMKAEW, M.; HOU, S.; SOMASUNDARAN, M.; WANG, J. P.; FINBERG, R. W.; ZHU, Q. K.; SCHIFFER, C. A.; MARASCO, W. A. (2021). Unique structural solution from a VH 3-30 antibody targeting the hemagglutinin stem of influenza A viruses. *Nature Communications*, 12(1), 1–13. <https://doi.org/10.1038/s41467-020-20879-6>

HASHEMI, S. A.; SAFAMANESH, S.; GHASEMZADEH-MOGHADDAM, H.; GHAFOURI, M.; AZIMIAN, A. High prevalence of SARS-CoV-2 and influenza A virus (H1N1) coinfection in dead patients in Northeastern Iran. **Journal of Medical Virology**, v. 93, n. 2, 2021.

HAYLES, N. K. Novel corona: Posthuman virus. **Crit Inq**, v. 47, n.2, p. 68–72, 2021.

HISANO, S.; ZHANG, R.; FARUK, M. I.; KONDO, H.; SUZUKI, N. A neo-virus lifestyle exhibited by a (+)ssRNA virus hosted in an unrelated dsRNA virus: Taxonomic and evolutionary considerations. **Virus Research**, v. 244, p. 75–83. 2018.

HOO, Z. H.; CANDLISH, J.; TEARE, D. What is an ROC curve? *Emergency Medicine Journal*, v. 34, n. 6, p. 357–359, 2017.

HUTCHINSON, E. C. Influenza Virus. **Trends Microbiol**, v. 26, n. 9, p. 809–810, 2018.

ISON, M. G. Antiviral Treatments. **Clinics in Chest Medicine**, v. 38, n. 1, p. 139–153, 2017.

JAMIN, M.; YABUKARSKI, F. Nonsegmented negative-sense rna viruses—structural data bring new insights into nucleocapsid assembly. **Advances in Virus Research**, v. 97, p. 143–185, 2017.

JASSBI, A. R.; ZARE, S.; ASADOLLAHI, M.; SCHUMAN, M. C. Ecological roles and biological activities of specialized metabolites from the genus *Nicotiana*. **Chemical reviews**, v. 117, n. 19, p. 12227-12280, 2017.

KALRA, S.; KUMAR, A.; JARHYAN, P.; UNNIKRISHNAN, A. G. Endemic or epidemic? Measuring the endemicity index of diabetes. **I. Journ. Endoc. Met.**, v. 19, n. 1, p. 5–7, 2015.

KLAJMAN, C. A Gripe sob a ótica da História Ecológica: um estudo comparativo entre as pandemias de 1918 e 2009. **História Revista**, v. 20, n. 3, p. 118, 2016.

KOONIN, L. M.; PATEL, A. Timely antiviral administration during an influenza pandemic: Key components. **American Journal of Public Health**, v. 108, p. 215–S220, 2018.

KRISCHUNS, T; LUKARSKA, M.; NAFFAKH, N.; CUSACK, S. Influenza Virus RNA-Dependent RNA Polymerase and the Host Transcriptional Apparatus. **Annual Review Biochemistry**, v. 90, p. 321–348, 2021.

LAM, P. C.H.; ABAGYAN, R.; TOTROV, M. Ligand-biased ensemble receptor docking (LigBEnD): a hybrid ligand/receptor structure-based approach. **Journal of Computer-Aided Molecular Design**, v. 32, n. 1, p. 187–198, 2018.

LIMA, V. C. O.; MACHADO, R. J. A.; MONTEIRO, N. K. V.; LYRA, I. L.; CAMILLO, C. S.; SERQUIZ, A. C.; OLIVEIRA, A. S.; RUFINO, F. P. S.; MACIEL, B. L. L.; UCHÔA, A. F.; SANTOS, E. A.; MORAIS, A. H. A. Gastroprotective and antielastase effects of protein inhibitors from *Erythrina velutina* seeds in an experimental ulcer model. **Biochemistry and Cell Biology**, v. 95, n.2, 2017.

LINHARES, J.; MIRANDA, I.; SILVA, W.; XAVIER, D.; XAVIER, J.; VELOSO, W. Introdução à triagem virtual. **Revista Brasileira de Bioinformática**, v.1, p. 1-18, 2021.

LINS, A. J. C. C. Aplicação de aprendizagem de máquina no diagnóstico de declínio cognitivo e demência de Alzheimer baseado em testes cognitivos e marcadores genéticos. **Semantic Scholar**, v. 1, 2018.

LIRA, E. L. S.; SOUSA, L. A. G.; LINS, S. R. O. Levantamento sobre plantas medicinais utilizadas em distúrbios do sistema digestivo no Município de Bezerros-PE. **Brazilian Journal of Development**, v. 6, n. 12, p. 1-12, 2020.

LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, v. 23, n.3, p. 4-25, 1997.

LIU, X.; XU, F.; REN, L.; ZHAO, F.; HUANG, Y.; WEI, L.; WANG, Y.; WANG, C.; FAN, Z.; MEI, S.; SONG, J.; ZHAO, Z.; CEN, S.; LIANG, C.; WANG, J.; & GUO, F. Inhibits influenza A virus infection by targeting viral M2 protein for ubiquitination-dependent degradation in lysosomes. **Nature Communications**, v. 12, n. 1. 2021.

LOO, J. S. E.; EMTAGE, A. L.; NG, K. W.; YONG, A. S. J.; DOUGHTY, S. W. Assessing GPCR homology models constructed from templates of various transmembrane sequence identities: Binding mode prediction and docking enrichment. **Journal of Molecular Graphics and Modelling**, v. 80, p. 38–47, 2018.

LUISE, C.; ROBAA, D. Application of Virtual Screening Approaches for the Identification of Small Molecule Inhibitors of the Methyllysine Reader Protein Spindlin1. **Rational Drug Design**, p. 347–370, 2018.

LUO, M.; TERRELL, J. R.; MCMANUS, S. A. Nucleocapsid Structure of Negative Strand RNA Virus. **Viruses**, v. 12, n. 8, p. 835, 2020.

MAIDA, Y.; YASUKAWA, M.; FURUUCHI, M.; LASSMANN, T.; POSSEMATO, R.; OKAMOTO, N.; KASIM, V.; HAYASHIZAKI, Y.; HAHN, W. C.; MASUTOMI, K. An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. **Nature**, v. 461, n. 7261, p. 230–235, 2009.

MACHADO, R. J. A.; MONTEIRO, N.; MIGLIOLO, L.; SILVA, O. N. Characterization and Pharmacological Properties of a Novel Multifunctional Kunitz Inhibitor from Erythrina velutina Seeds. **Plos One**, v. 8, n. 5, 2013.

MARTINS, E. J. V. **Data Mining em aplicações de Desenho racional de Fármacos**. 2017. 65f. Dissertação (Mestrado em Engenharia Informática e Computação), Universidade do Porto.

MASCARENHAS, R. M. G.; LIMA, C. A.; RODRIGUES, R. P.; KITAGAWA, R. R.; FARAONI, A. S.; OLIVEIRA, T. B.; KITAGAWA, R. R.; FARAONI, A. S.; OLIVEIRA, T. B.V. Avaliação de toxicidade geral em constituintes alimentares utilizando ferramentas in silico. **Revista Virtual de Química**, v. 11, n. 2, p. 543-553, 2019.

MASSARO, A.; MARITATI, V.; GALIANO, A.; BIRARDI, V.; PELLICANI, L. ESB Platform Integrating Knime Data Mining Tool Oriented on Industry 4.0 Based on Artificial Neural Network Predictive Maintenance. **International Journal of Artificial Intelligence & Applications**, v. 9, n. 3, p. 01–17, 2018.

MATROSOVICH, M. N.; MATROSOVICH, T. Y.; GRAY, T.; ROBERTS, N. A.; KLENK, H.D. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. **Proceedings of the National Academy of Sciences**, v. 101, n. 13, p. 4620–4624, 2004.

MEDEIROS, H. I. R.; CAIANA, R. R. A.; FERNANDES, N. D.; MEDEIROS JÚNIOR, F. C.; SCOTTI, M. T.; SCOTTI, L.; FREITAS, J. C. R. de. Rational planning of a drug candidate: in silico studies, synthesis and structural elucidation. **Research, Society and Development**, v. 9, n. 11, p. 1-27, 2020. DOI: 10.33448/rsd-v9i11.10605.

METASITE, versão 6.0. Molecular Discovery Ltd. London. Disponível em: <https://www.moldiscovery.com/software/metasite/>, acessado em setembro de 2021.

MILLER, A. D. Retroviral Vectors in Gene Therapy. **E LS**, 2001.

MOLINSPIRATION. Molinspiration Molecule Viewer: Chem informatics on the Web. 2018. Disponível em: < <http://www.molinspiration.com>>. Acesso em setembro de 2021.

MOREIRA, R. F. **Determinação do perfil farmacocinético de medicamentos contendo fármacos de ação central**. 2014. 161f. Tese (Doutorado em Ciências), Faculdade de Ciências Médicas, Campinas - SP.

NAMASIVAYAM, V.; SILBERMANN, K.; WIESE, M.; PAHNKE, J.; & STEFAN, S. M. COPA: Computer-Aided Pattern Analysis to Predict Multitarget ABC Transporter Inhibitors. **Journal of Medicinal Chemistry**, v. 64, n. 6, p. 3350–3366, 2021.

NEVEU, E.; POPOV, P.; HOFFMANN, A.; MIGLIOSI, A.; BESSERON, X.; DANOY, G.; BOUVRY, P.; GRUDININ, S. RapidRMSD: Rapid determination of RMSDs corresponding to motions of flexible molecules. **Bioinformatics**, v. 1, n. 9, 2018.

NEWMAN, D. J.; CRAGG, G. M. Natural Products as Sources of New Drugs *from 1981 to 2014*. **Journal of natural products**, v. 79, n. 3, p. 629-661, 2016.

NIE, C.; STADTMÜLLER, M.; PARSHAD, B.; WALLERT, M.; AHMADI, V.; KERKHOFF, Y.; BHATIA, S.; BLOCK, S.; CHENG, C.; WOLFF, T.; & HAAG, R. Heteromultivalent topology-matched nanostructures as potent and broad-spectrum influenza A virus inhibitors. **Science Advances**, v. 7, n. 1, p. 1–10, 2021.

NOGUEIRA, T.; PONCE, R. O vírus da Gripe. **Revista de Ciência Elementar**, v. 9, n. 2, p. 1–5, 2021.

NUNES, R. R.; DA FONSECA, A. L.; PINTO, A. C. D. S.; MAIA, E. H. B.; DA SILVA, A. M.; VAROTTI, F. D. P.; & TARANTO, A. G. Brazilian malaria molecular targets (BraMMT): Selected receptors for virtual high-throughput screening experiments. **Memorias Do Instituto Oswaldo Cruz**, v. 114, n. 2, p. 1–10, 2019.

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

ODOARDI, S.; MESTRIA, S.; BIOSA, G.; ARFÈ, R.; TIRRI, M.; MARTI, M.; STRANO ROSSI, S. Metabolism study and toxicological determination of mephtetramine in biological samples by liquid chromatography coupled with high-resolution mass spectrometry. **Drug Testing and Analysis**, v. 13, n. 8, p. 1516–1526, 2021.

OLIVEIRA, V. F.; OLIVEIRA, H. M. B. F.; MEDEIROS, C. I. S.; FILHO, A. A. O.; REGO, T. G. Análise farmacológica e toxicológica in silico do flavonoide 5- hidroxí-4', 7- dimetoxiflavona. **Journal of Medicine and Health Promotion**, v. 3, n. 1, p. 913-921, 2018.

OMOTO, S.; SPERANZINI, V.; HASHIMOTO, T.; NOSHI, T.; YAMAGUCHI, H.; KAWAI, M.; KAWAGUCHI, K.; UEHARA, T.; SHISHIDO, T.; NAITO, A.; & CUSACK, S. Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil. **Scientific Reports**, v. 8, n. 1, p. 1–15. 2018.

PADHORN, D.; HALL, D. R.; MIRZAEI, H.; MAMONOV, A. B.; MOGHADASI, M.; ALEKSEENKO, A.; BEGLOV, D.; & KOZAKOV, D. Protein–ligand docking using FFT based sampling: D3R case study. **Journal of Computer-Aided Molecular Design**, v. 32, n. 1, p. 225–230, 2018.

PAGADALA, N. S.; SYED, K.; & TUSZYNSKI, J. Software for molecular docking: a review. **Biophysical Reviews**, v. 9, n. 2, p. 91–102, 2017.

- PANDEY, P.; LYNN, A. M.; BANDYOPADHYAY, P. Identification of Inhibitors against  $\alpha$ -Isopropylmalate Synthase of Mycobacterium Tuberculosis Using Docking-MM/PBSA Hybrid Approach. *Bioinformation* 2017, 13 (5), 144–148. <https://doi.org/10.6026/97320630013144>.
- POUTARAUD, A.; MICHELOT-ANTALIK, A.; PLANTUREUX, S. Grasslands: a source of secondary metabolites for livestock health. **Journal of agricultural and food chemistry**, v. 65, n. 31, p. 6535-6553, 2017.
- REN, J.-X.; CHENG, Z.; HUANG, Y.-X.; ZHAO, J.-F.; GUO, P.; ZOU, Z.-M.; XIE, Y. Identification of Novel Dual-Specificity Phosphatase 26 Inhibitors by a Hybrid Virtual Screening Approach Based on pharmacophore and molecular docking. **Biomed Pharmacother**, v. 89, p. 376–385. 2017.
- RODRIGUES, F. T. S.; SOUSA, C. N. S.; XIMENES, N. C.; ALMEIDA, A. B.; CABRAL, L. M.; PATROCÍNIO, C. F. V.; SILVA, A. H.; LEAL, L. K. A. M.; JÚNIOR, J. E. R. H.; MACEDO, D.; VASCONCELOS, S. M. M. Effects of standard ethanolic extract from *Erythrina velutina* in acute cerebral ischemia in mice. **Biomedicine & Pharmacotherapy**, v.96, n.1, 2017.
- ROSALES, P. F. **Alcaloides indólicos de *Tabernaemontana catharinensis*: fracionamento, isolamento, síntese e estudos de atividade *in vitro* e *in silico***. 2020. 144 f. Tese (Doutorado em biotecnologia) – Universidade de Caxias do Sul, Caxias do Sul-RS.
- RUIZ-TORRES, V.; ENCINAR, A. J.; HERRANZ-LÓPEZ, M.; PÉREZ-SÁNCHEZ, A.; GALIANO, V.; BARRAJÓN-CATALÁN, E.; MICOL, V. An Updated Review on Marine Anticancer Compounds: The Use of Virtual Screening for the Discovery of Small-Molecule Cancer Drugs. *Molecules* . 2017. <https://doi.org/10.3390/molecules22071037>.
- SAMPAIO, V. S.; VIEIRA, I. M. F.; JUNIOR, E. A. L.; LOIOLA, M. I. B. Flora do Ceará, Brasil: *Solanum* (Solanaceae). **Rodriguésia**, v. 70, n. 1, p. 1-27, 2019.
- SANTOS, R. N.; FERREIRA, L. G.; ANDRICOPULO, A. D. Practices in Molecular Docking and Structure-Based Virtual Screening BT. **Computational Drug Discovery and Design**, p. 31–50, 2018.
- SCHUELER, P. O que é uma Pandemia? Fundação Oswaldo Cruz - FioCruz. **Bio-manguinhas**, v. 1, p. 1–1, 2020.

SHANG, J.; DAI, X.; LI, Y.; PISTOLOZZI, M.; WANG, L. HYBRIDSIM-VS: A Web Server for Large-Scale Ligand-Based Virtual Screening Using Hybrid Similarity Recognition Techniques. *Bioinformatics* **2017**, *33* (21), 3480–3481.

SILVAa, D. A. A. **Mobilização de reservas e perfil preliminar de metabólitos secundários durante e após a germinação em *Erythrina velutina***. 2020. 54f. Dissertação (Mestrado em Ciências Florestais) - Escola Agrícola de Jundiá, Universidade Federal do Rio Grande do Norte, Natal-RN.

SILVAb, N. G. M. **Estudo do metabolismo do 7-nitroindazol: uma abordagem *in silico***. 2020. 54 f. Trabalho de Conclusão de Curso (Graduação em Farmácia) – Universidade Federal de Santa Catarina, graduação em farmácia, Florianópolis – SC.

SILVAc, R. S.; VELOSO, C. L.; SOUSA, A. C. P.; SILVA, R. L. V.; OLIVEIRA, J.; NASCIMENTO, D. L.; MARQUES, F. J. Utilidades do Mulungu (*Erythrina velutina* Willd.). **Brazilian Journal of Development**, v. 6, n. 3, 2020.

PETKOVIC, D.; ALTMAN, R.; WONG, M.; VIGIL, A. Improving the explainability of Random Forest classifier – User centered approach. **Pacific Symposium on Biocomputing**, v. 23, p. 204–215, 2018.

PFLUG, A.; LUKARSKA, M.; RESA-INFANTE, P.; REICH, S.; & CUSACK, S. Structural insights into RNA synthesis by the influenza virus transcription-replication machine. **Virus Research**, v. 234, p. 103–117, 2017.

PICCIRILLO, E.; AMARAL, A. T. Virtual screening of bioactive compounds: concepts and applications. **Química Nova**, v. 41, n. 6, p. 662–677, 2018.

RADCHENKO, T.; BRINK, A.; SIEGRIST, Y.; KOCHANSKY, C.; BATEMAN, A.; FONTAINE, F.; MORETTONI, L.; & ZAMORA, I. Software-aided approach to investigate peptide structure and metabolic susceptibility of amide bonds in peptide drugs based on high resolution mass spectrometry. **PloS One**, v. 12, n. 11, 2017.

RIBEIRO, T. P. P. Estudo teórico (modelagem molecular e QSAR) de compostos quinolínicos com atividade herbicida. **Physics**, 2017.

ROBLES, M. B. G.; LÉON, M. P. P.; REYES, A. R.; PARA, M. G. DE LA. Seguridad del tratamiento con mexicanos con influenza y enfermedad tipo influenza. **Enfermedades Infecciosas y Microbiología**, v. 40, n. 3, p. 81–86, 2020.

RODRIGUES, G. DOS S.; AVELINO, J. A.; SIQUEIRA, A. L. N.; RAMOS, L. F. P.; & SANTOS, G. B. DOS. O uso de softwares livres em aula prática sobre filtros moleculares de biodisponibilidade oral de fármacos. **Química Nova**, v. 00, p. 1–9. 2021.

RODRIGUES, L.M.S. **Estudos de Docking molecular de derivados tiazóis com a BSA e como potenciais intercaladores de DNA e inibidores da topoisomerase II**. 2019. 61 f. Monografia (Licenciatura em Química) – Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Serra Talhada, Serra Talhada-PE, 2019.

ROSENBERG, C. E. What Is an Epidemic?. **The MIT Press on Behalf of American Academy of Arts & Sciences**, v. 118, n. 2, p. 1–17, 2016.

RUSSELL, R. J.; HAIRE, L. F.; STEVENS, D. J.; COLLINS, P. J.; LIN, Y. P.; BLACKBURN, G. M.; HAY, A. J.; GAMBLIN, S. J.; SKEHEL, J. J. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. **Nature**, v. 443, p. 45-49, 2006.

RUISÁNCHEZ, I.; JIMÉNEZ-CARVELO, A. M.; & CALLAO, M. P. ROC curves for the optimization of one-class model parameters. A case study: Authenticating extra virgin olive oil from a Catalan protected designation of origin. **Talanta**, v. 222, 2021.

SALMAN, M. M.; AL-OBAIDI, Z.; KITCHEN, P.; LORETO, A.; BILL, R. M.; WADE-MARTINS, R. Advances in Applying Computer-Aided Drug Design for Neurodegenerative Diseases. **International Journal of Molecular Sciences**, v. 22, n.9, p. 4688, 2021.

SÁNCHEZ-ESPINOZA, A.; CONONA-NÚÑEZ, T. J.; ROJO-CONTRERAS, E. W.; AGUILAR-MADRIGAL, F.; REYES-VÉLEZ, M. N. Influenza A (H1N1): Un nuevo desafío , que llegó para quedarse .... **Salud Jalisco**, v. 1, p. 164–172, 2020.

SCOTTI, M. T.; HERRERA-ACEVEDO, C.; OLIVEIRA, T.B.; COSTA, R. P. O.; SANTOS, S.Y. K. O.; RODRIGUES, R. P.; SCOTTI, L.; DA-COSTA, F. B. Sistemax, an Online Web-Based Cheminformatics Tool for Data Management of Secondary Metabolites. **Molecules**, v. 23, p. 103, 2018.

SENHORAS, E. M. Coronavirus and the role of pandemics in human history. **Boletim de Conjuntura**, v. 1, p. 31–34, 2020.

SILVAd, R. S. As aplicações da informática no ensino de química como alternativa para o ensino e aprendizagem. **Revista Amor Mundi**, v. 1, n. 2, p. 43–56, 2020.

STOUFFER, A. L.; ACHARYA, R.; SALOM, D.; LEVINE, A. S.; DI COSTANZO, L.; SOTO, C. S.; TERESHKO, V.; NANDA, V.; STAYROOK, S.; & DEGRADO, W. F. Structural basis for the function and inhibition of an influenza virus proton channel. **Nature**, v. 451, n. 7178, p. 596–599, 2008.

SYDOW, D.; WICHMANN, M.; RODRÍGUEZ-GUERRA, J.; GOLDMANN, D.; LANDRUM, G.; & VOLKAMER, A. TeachOpenCADD-KNIME: A Teaching Platform for Computer-Aided Drug Design Using KNIME Workflows. **Journal of Chemical Information and Modeling**, p. 8–11, 2019.

THÉBAUD, G.; CHADŒUF, J.; MORELLI, M. J.; MCCAULEY, J. W.; & HAYDON, D. T. The relationship between mutation frequency and replication strategy in positive-sense single-stranded RNA viruses. *Proceedings of the Royal Society B: Biological Sciences*, v. 277, n. 1682, p. 809–817, 2010.

TIAN, L.; QIANG, T.; LIANG, C.; REN, X.; JIA, M.; ZHANG, J.; LI, J.; WAN, M.; YUWEN, X.; LI, H.; CAO, W.; & LIU, H. RNA-dependent RNA polymerase (RdRp) inhibitors: The current landscape and repurposing for the COVID-19 pandemic. **European Journal of Medicinal Chemistry**, v. 213, p. 113201, 2021.

TOROPOVA, M. A. Drug metabolism as an object of computational analysis by the Monte Carlo method. **Current Drug Metabolism**, v. 18, n. 12, p. 1123-1131, 2017.

TRIGUERO, C. R.; MAGADÁN, C. R.; & PÉREZ, V. S. Aplicaciones quimioinformáticas en el descubrimiento de fármacos. **Dianas**, v. 5, n. 2, p. 1–10, 2016.

ULLAH, A.; PROTTOY, N. I.; ARAF, Y.; HOSSAIN, S.; SARKAR, B.; SAHA, A. Molecular Docking and pharmacological property analysis of phytochemicals from clitoria ternatea as potent inhibitors of cell cycle checkpoint proteins in the cyclin/cdk pathway in cancer cells. **Scientific Research Publishing**, v. 9, n. 3, p. 1-14, 2019.

VAZ, T. A. (2017). **Modelo de dados para treinamento de inteligência artificial na pesquisa em saúde: um estudo prático sobre infecções hospitalares**. 2017. 103f. Dissertação (Mestrado em pesquisa clínica), Universidade Federal do Rio Grande do Sul - Hospital de Clínicas de Porto Alegre, Porto Alegre - RS.

VAVOUGIOS, G. D.; ZAROGIANNIS, S. G.; KROGFELT, K. A.; GOURGOULIANIS, K.; MITSIKOSTAS, D. D.; HADJIGEORGIOU, G. Novel candidate genes of the PARK7 interactome as mediators of apoptosis and acetylation in multiple sclerosis: an *in silico* analysis. **Multiple Sclerosis and Related Disorders**, v. 19, n. 1, p. 8-14, 2018.

VELÁZQUEZ-LIBERA, J. L.; DURÁN-VERDUGO, F.; VALDÉS-JIMÉNEZ, A.; VALDÉS-JIMÉNEZ, A.; NÚÑEZ-VIVANCO, G.; & CABALLERO, J. LigRMSD: A web server for automatic structure matching and RMSD calculations among identical and similar compounds in protein-ligand docking. **Bioinformatics**, v. 36, n. 9, p. 2912–2914, 2020.

VELTHUIS, A. J. W. Common and unique features of viral RNA-dependent polymerases. *Cellular and Molecular Life Sciences*, v. 71, n. 22, p. 4403–4420, 2014.

VENKATARAMAN, S.; PRASAD, B. V. L. S.; & SELVARAJAN, R. RNA dependent RNA polymerases: insights from structure, function and evolution. **Viruses**, v. 10, n. 2, p. 76, 2018.

WANG, D.; ZHU, W.; YANG, L.; & SHU, Y. The epidemiology, virology, and pathogenicity of human infections with avian influenza viruses. **Cold Spring Harbor Perspectives in Medicine**, v. 11, n. 4, 2021.

WANG, Z.; LORENZI, J. C. C.; MUECKSCH, F.; FINKIN, S.; VIANT, C.; GAEBLER, C.; CIPOLA, M.; HOFFMANN, HH.; OLIVEIRA, T. Y.; OREN, D. A.; RAMOS, V.; NOGUEIRA, L.; MICHAILIDIS, E.; ROBBIANI, D. F.; GAZUMYAN, A.; RICE, C. M.; HATZIOANNOU, T.; BIENIASZ, P. D.; CASKEY, M.; NUSSENZWEIG, M. C. Enhanced SARS-CoV-2 neutralization by dimeric IgA. **Science Translational Medicine**, v. 13, n. 577, p. 1555, 2020.

WORLD HEALTH ORGANIZATION (WHO). Coronavirus disease (COVID-19): situation report, 162. 2020. Disponível em: <<https://apps.who.int/iris/handle/10665/332970>>, acesso em 28 de outubro de 2021.

XU, T. Y.; WANG P.; TIAN J. S.; QING, S. L.; WANG, S. N.; HUANG, Y. H.; XU, J. Y.; SU, D. F.; LIU, J. G.; MIAO, C. Y. Pharmacological characterization of MT-1207, a novel multitarget antihypertensive agent. *Acta Pharmacol Sin.* 2021 Jun;42(6):885-897. doi: 10.1038/s41401-021-00636-1.

YU, W.; & MACKERELL, A. D. Computer-Aided Drug Design Methods. **In *Biotechnology and Bioengineering***, v. 7, n. 1, p. 85–106, 2017.

ZWEIG, M. H.; CAMPBELL, G. Receiver-operating characteristic (ROC) plots. ***Clinical Chemistry***, v. 29, p. 561-577, 1993.

ZENG, B.; ZHANG, H.; XU, X.; WU, Z.; & XIONG, C. A correlation analysis of short-term imaging manifestations and long-term function using ROC curve after tibial fracture surgery. ***American Journal of Translation Research***, v. 13, n. 6, p. 6724-6730, 2021.

ZHANG, Z.; HE, G.; FILIPOWICZ, N. A.; RANDALL, G.; BELOV, G. A.; KOPEK, B. G.; WANG, X. Host lipids in positive-strand RNA virus genome replication. ***Frontiers in Microbiology***, v. 10, p. 286, 2019.

ZHENG, W.; TAO, Y. J. Structure and assembly of the influenza A virus ribonucleoprotein complex. ***FEBS Letters***, v. 587, n. 8, p. 1206–1214, 2013.

ZHU, L.; LING, J.; ZHU, Z.; TIAN, T.; SONG, Y.; & YANG, C. Selection and applications of functional nucleic acids for infectious disease detection and prevention. ***Analytical and Bioanalytical Chemistry***, v. 413, p. 4563–4579, 2021.