

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE MICROBIOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA**

Mateus Sá Magalhães Serafim

**PLANEJAMENTO DE MOLÉCULAS BIOATIVAS
CONTRA PROTEASES DE SARS-CoV-2:
estudos in silico e potencial biológico**

Belo Horizonte

2023

Mateus Sá Magalhães Serafim

**PLANEJAMENTO DE MOLÉCULAS BIOATIVAS
CONTRA PROTEASES DE SARS-CoV-2:
estudos *in silico* e potencial biológico**

Tese apresentada ao Programa de Pós-Graduação em Microbiologia (PPGM) da Universidade Federal de Minas Gerais (UFMG), como requisito parcial à obtenção do título de Doutor em Microbiologia.

Orientador: Prof. Jônatas Santos Abrahão (ICB/UFMG)

Coorientadores: Prof. Bruno Eduardo Fernandes Mota (FAFAR/UFMG)

Prof. Vinícius Gonçalves Maltarollo (FAFAR/UFMG)

Colaboradores: Dr. Thales Kronenberger (Universidade de Tübingen, Alemanha)

Dr. Anthony J. O'Donoghue (Universidade de San Diego, Califórnia, Estados Unidos)

Dr. Conor R. Caffrey (Universidade de San Diego, Califórnia, Estados Unidos)

Belo Horizonte

2023

043

Serafim, Mateus Sá Magalhães.

Planejamento de moléculas bioativas contra proteases de SARS-COV-2: estudos in silico e potencial biológico [manuscrito] / Mateus Sá Magalhães Serafim. – 2023.

136 f. : il. ; 29,5 cm.

Orientador: Prof. Jônatas Santos Abrahão. Coorientadores: Prof. Bruno Eduardo Fernandes Mota; Prof. Vinícius Gonçalves Maltarollo. Colaboradores: Dr. Thales Kronenberger. Dr. Anthony J. O'Donoghue. Dr. Conor R. Caffrey.

Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Microbiologia.

1. Microbiologia. 2. Antivirais. 3. Betacoronavirus. 4. Inibidores de Proteases. I. Abrahão, Jônatas Santos. II. Mota, Bruno Eduardo Fernandes. III. Maltarollo, Vinícius Gonçalves. IV. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. V. Título.

CDU: 579



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PÓS-GRADUAÇÃO EM MICROBIOLOGIA

ATA DE DEFESA DE TESE

ATA DA DEFESA DE TESE DE **MATEUS SÁ MAGALHÃES SERAFIM**

Nº REGISTRO: **2020715974**

Às 13:30 horas do dia **17 de fevereiro de 2023**, reuniu-se, por via remota, a Comissão Examinadora composta pelos Drs. Jordana Graziella Alves Coelho dos Reis (Departamento de Microbiologia/ICB /UFMG), Débora Maria Abrantes Costa (UFMG), Prof. Dr. Glaucio Monteiro Ferreira (Universidade de São Paulo), Jair Lage de Siqueira Neto (University of California San Diego, EUA) e o Prof. Dr. Jônatas Santos Abrahão (Orientador), para julgar o trabalho final **“Planejamento de moléculas bioativas contra proteases de SARS-COV-2: estudos in silico e potencial biológico”**, do aluno **Mateus Sá Magalhães Serafim**, requisito final para a obtenção do Grau de **DOUTOR EM CIÊNCIAS BIOLÓGICAS: MICROBIOLOGIA**. Abrindo a sessão, o Presidente da Comissão, Prof. Dr. Jônatas Santos Abrahão, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para a apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. O candidato foi considerado **APROVADO**. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ata, que será assinada por todos os membros participantes da Comissão Examinadora. O candidato tem 60 (sessenta) dias, a partir desta data, para entregar a versão final da tese ao Programa de Pós-graduação em Microbiologia da UFMG e requerer seu diploma.

Belo Horizonte, 17 de fevereiro de 2023

Membros da Banca:

Profa. Dra. Jordana Graziella Alves Coelho dos Reis

Profa. Dra. Débora Maria Abrantes Costa

Prof. Dr. Glaucio Monteiro Ferreira

Prof. Dr. Jair Lage de Siqueira Neto

De acordo:

Prof. Dr. Jônatas Santos Abrahão

(Orientador)

Profa. Daniele da Glória de Souza
(Coordenadora do Programa de Pós-graduação
em Microbiologia)



Documento assinado eletronicamente por **Glauco Monteiro Ferreira, Usuário Externo**, em 23/02/2023, às 10:11, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Débora Maria Abrantes Costa, Professora do Magistério Superior**, em 23/02/2023, às 10:39, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jonatas Santos Abrahao, Professor do Magistério Superior**, em 23/02/2023, às 12:13, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Daniele da Glória de Souza, Coordenador(a) de curso de pós-graduação**, em 23/02/2023, às 13:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jordana Graziela Alves Coelho dos Reis, Membro de comitê**, em 27/02/2023, às 21:35, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jair Lage de Siqueira Neto, Usuário Externo**, em 28/02/2023, às 16:44, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2054131** e o código CRC **1F8417C4**.

Dedico esta tese à dona Rosa, que me criou e sonhava em ter um neto doutor. Eu te amo, vó.

Dedico em memória ao meu primo “Rick”, meu tio “Liminha”, minha vó “Estela” e meu fiel escudeiro, “Ralf”. Descansem em paz.

Agradecimentos

Agradeço às instituições que me formaram, UFMG e UCSD;

Agradeço às agências de fomento, CAPES, CNPq e FAPEMIG;

Agradeço aos orientadores que me guiaram, Bruno, Jônatas e Vinícius;

Agradeço aos supervisores que me aceitaram, Anthony e Conor;

Agradeço aos colaboradores, colegas e profissionais que me auxiliaram;

Agradeço aos meus amigos e à minha segunda família que me suportaram;

Agradeço à minha irmã, Vitória “pirralha”, meu anjo da guarda;

Agradeço à minha mãe, Regina, meu exemplo de resiliência e fonte de amor;

Agradeço ao meu pai, Fernando, meu exemplo de integridade e fonte de respeito;

Agradeço à minha vó, Rosa, meu exemplo de vida e fonte de inspiração;

Agradeço à minha noiva e futura esposa, minha luz e eterno amor, Ellen;

Agradeço a você, leitor, pela paciência e tempo dispostos por aqui.

“A ciência mostra os fatos, independente da sua opinião”.

RESUMO

Aproximadamente 664,8 milhões de infecções por SARS-CoV-2 foram notificados no mundo até 25 de janeiro de 2023, com mais de 6,72 milhões de mortes notificadas de em três anos. Embora medidas não farmacológicas, desenvolvimento e distribuição de vacinas, utilização de fármacos antivirais nirmatrelvir e rendesivir, bem como de anticorpos monoclonais tenham sido empregados contra COVID-19, baixas coberturas vacinais e o acúmulo de mutações de SARS-CoV-2 dificultam o controle da pandemia. Ademais, a possibilidade de seleção de linhagens resistentes às opções terapêuticas se apresenta como um agravante do cenário. Nesse sentido, se faz necessária a contínua obtenção de novas moléculas bioativas contra SARS-CoV-2. Portanto, o presente trabalho objetiva, empregando estratégias *in silico* e *in vitro*, a obtenção de moléculas bioativas com potencial atividade inibitória contra as proteases, M^{pro} e PL^{pro}, de SARS-CoV-2. Inicialmente, um modelo de HQSAR ($q^2 = 0,964$ e $r^2 = 0,975$) foi selecionado para predição de ligantes com potencial atividade inibitória contra M^{pro}. Foram identificados 24 potenciais inibidores que seguiram para avaliação de inibição da atividade enzimática, sendo todos inativos (menor que 50% de inibição) a 10 μ M. Posteriormente, foram selecionadas estruturas cristalográficas das proteases para predições de sítios de ligação e feitas análises de acoplamento molecular (*redocking*) para construção de modelos farmacofóricos baseados nas M^{pro} e PL^{pro} de SARS-CoV-2 e outros coronavírus. Estes foram utilizados em triagens virtuais para seleção de potenciais inibidores em quimiotecas (BraCoLi, DrugBank, GMR, NuBBE e ZINC15) totalizando 6,8 milhões de moléculas. A triagem virtual dos alvos foi conduzida e 46 *hits* de duas quimiotecas disponíveis (DrugBank e GMR) foram obtidos. Dentre as moléculas triadas (14 à pronta disposição), nenhuma apresentou inibição da atividade enzimática (maior que 50%) contra M^{pro} e PL^{pro} até 10 μ M. Posteriormente, uma série de 127 análogos estruturais dos 46 *hits* virtuais indisponíveis foi avaliada nos ensaios enzimáticos, e cinco dentre as 127 moléculas tiveram suas concentrações inibitória (IC₅₀) determinadas para M^{pro} de SARS-CoV-2 (0.086 ± 0.0015 a 16.42 ± 0.095 μ M), sendo dois ativos contra as M^{pro} de SARS-CoV e MERS-CoV (0.11 ± 0.0175 a 3.786 ± 0.64 μ M) e todos inativos para PL^{pro}. Adicionalmente, extratos e isolados triterpenos planejados contra SARS-CoV-2 foram avaliados quanto a sua citotoxicidade e atividade antiviral pelo ensaio de MTT, determinando as concentrações citotóxica (CC₅₀) e efetiva de 50% (CE₅₀) contra um coronavírus murino (MHV-3). Quatro substâncias apresentaram atividade antiviral, com índices de seletividade (IS) entre 1,08 e 34,36. Tomadas em conjunto, as estratégias computacionais empregadas foram capazes de obter inibidores preditos *in silico* e determinados *in vitro*, demonstrando o potencial e sucesso no planejamento de moléculas bioativas contra SARS-CoV-2.

Palavras-chave: Atividade antiviral. Inibição enzimática. M^{pro}. PL^{pro}. SARS-CoV-2.

ABSTRACT

As of January 25, 2023, approximately 664.8 million SARS-CoV-2 infections have been reported worldwide, with over 6.72 million deaths notified over the last three years. Whilst non-pharmacological interventions, vaccines' development and distribution, direct antiviral drugs approval (e.g., nirmatrelvir and remdesivir), as well as monoclonal antibodies have been employed against COVID-19, low vaccine coverage and new SARS-CoV-2 mutations negatively impact the control of the pandemic scenario. In addition, the possibility of resistant strains can limit therapeutic options. In this sense, it is still necessary to design and develop bioactive molecules against SARS-CoV-2. Therefore, the present work aims to obtain new inhibitors against SARS-CoV-2 proteases, M^{pro} and PL^{pro}, employing *in silico* and *in vitro* approaches. Initially, an HQSAR model ($q^2 = 0.964$ and $r^2 = 0.975$) was selected to predict ligands with potential inhibitory activity against M^{pro}. Here, 24 potential inhibitors obtained followed to enzymatic inhibition activity determination, all of which were inactive (less than 50% inhibition) at 10 μ M. Subsequently, crystallographic structures of the proteases were selected for binding sites predictions and molecular docking analysis (redocking) to build pharmacophoric models based on SARS-CoV-2 and other coronaviruses M^{pro} and PL^{pro}. These were used in virtual screening campaigns to select potential inhibitors in chemical libraries (BraCoLi, DrugBank, GMR, NuBBE and ZINC15), from a total of 6.8 million compounds. Virtual screening was conducted to both targets obtaining 46 compounds from two available libraries (DrugBank and GMR). Among the screened compounds (14 available in house), none showed inhibitory activity (over 50% inhibition) against M^{pro} and PL^{pro} up to 10 μ M. A series of 127 synthetic structural analogues of the 46 unavailable virtual hits were evaluated in enzymatic assays, and five out of 127 compounds had their 50% inhibitory concentrations (IC₅₀) determined for SARS-CoV-2 M^{pro} (0.086 ± 0.0015 to 16.42 ± 0.095 μ M). Two compounds were also active against SARS-CoV and MERS-CoV M^{pro} (0.11 ± 0.0175 to 3.786 ± 0.64 μ M), and all 127 were inactive against PL^{pro}. In addition, plant extracts and triterpene isolates designed against SARS-CoV-2 were evaluated for their cytotoxicity and antiviral activity by MTT assay, determining 50% cytotoxic concentration (CC₅₀) and 50% effective concentration (EC₅₀) against a murine coronavirus (MHV-3). Four compounds showed antiviral activity, with selectivity indexes (SI) between 1.08 and 34.36. Taken together, the computational strategies employed were able to obtain inhibitors predicted *in silico* and determined *in vitro*, demonstrating the potential and success in the planning of bioactive molecules against SARS-CoV-2.

Keywords: Antiviral activity. Enzymatic inhibition. M^{pro}. PL^{pro}. SARS-CoV-2.

LISTA DE FIGURAS

1.	Partículas virais de SARS-CoV observadas por microscopia eletrônica de transmissão (MET)	16
2.	Representação de partícula viral dos coronavírus	17
3.	Modelo de organização de proteínas traduzidas do genoma dos coronavírus	18
4.	Estruturas tridimensionais das enzimas de SARS-CoV-2	21
5.	Ciclo de multiplicação dos coronavírus	23
6.	Estrutura química do nirmatrelvir	29
7.	Etapas do mecanismo de ação da protease principal (M ^{pro}) de SARS-CoV-2	34
8.	Fluxograma de atividades das etapas executadas no projeto de doutorado	40
9.	Estrutura geral das moléculas utilizadas na construção do conjunto de dados de HQSAR	51
10.	Mapas de contribuição química das moléculas com maiores valores de pIC ₅₀ (23 e 21m)	55
11.	Seleção de descritores moleculares empregando aleatorização de N modelos	57
12.	Potenciais sítios de ligação e grupamentos de moléculas-sonda preditos na PL ^{pro} de SARS-CoV-2	63
13.	Potenciais sítios de ligação e grupamentos de moléculas-sonda preditos na M ^{pro} de SARS-CoV-2	63
14.	Potenciais sítios de ligação preditos nas M ^{pro} de diferentes coronavírus	64
15.	Predições de sítios de ligação com FTSite em um monômero da M ^{pro} de SARS-CoV-2	65
16.	Predições de potenciais sítios de ligação com PrankWeb nas M ^{pro} de diferentes coronavírus	66
17.	Predições de potenciais sítios de ligação com PrankWeb nas PL ^{pro} de diferentes coronavírus	67
18.	Fração do alinhamento de sequências de aminoácidos das M ^{pro} de diferentes coronavírus	70
19.	Alinhamento 3D de estruturas das M ^{pro} de diferentes coronavírus	72
20.	Redocking de GC376 na cadeia B da M ^{pro} de SARS-CoV-2, utilizando programa GOLD	74
21.	Redocking de VIR251 na PL ^{pro} de SARS-CoV-2, utilizando programa GOLD	77
22.	Simulações de DM do ligante co-cristalizado (GC376) da M ^{pro} de SARS-CoV-2 (PDB ID: 7D1M)	78
23.	Simulações de DM do ligante co-cristalizado (boceprevir) da M ^{pro} de SARS-CoV-2 (PDB ID: 7C6S)	79
24.	Simulações de DM do ligante co-cristalizado (GC376) da M ^{pro} de SARS-CoV-2 (PDB ID: 7C6U)	80
25.	Modelo farmacofórico construído na M ^{pro} de SARS-CoV-2 (PDB ID: 7D1M)	82
26.	Modelo farmacofórico construído na PL ^{pro} de SARS-CoV-2 (PDB ID: 6WX4)	83
27.	Hits das triagens virtuais contra as proteases de SARS-CoV-2	85
28.	Conjunto de moléculas selecionadas contra as proteases de SARS-CoV-2	87
29.	Curvas de dose-resposta dos inibidores contra M ^{pro} de SARS-CoV-2	93
30.	Curvas de dose-resposta de WRR-180 e WRR-181 contra as M ^{pro} de SARS-CoV e MERS-CoV	94
31.	Simulações de DM de WRR-180 e WRR-181 na M ^{pro} de SARS-CoV-2 (PDB ID: 7D1M)	96

LISTA DE TABELAS

Tabela 1	Funções das proteínas não estruturais (nsp) dos coronavírus	22
Tabela 2	Parâmetros definidos para os modelos de triagem virtual de ZINC15 (Molport)	44
Tabela 3	Relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) do conjunto de dados em triplicata de inibidores da M ^{pro} de SARS-CoV-2	53
Tabela 4	Variação do tamanho (quatro átomos) dos fragmentos do modelo HQSAR selecionado	54
Tabela 5	Características dos descritores moleculares utilizados no software PaDEL	56
Tabela 6	Predições de toxicidade e farmacocinética das 24 moléculas com admetSAR 2.0	58
Tabela 7	Atividade inibitória das moléculas triadas por HQSAR contra M ^{pro} de SARS-CoV-2	59
Tabela 8	Características das estruturas cristalográficas das proteases de SARS-CoV-2 selecionadas no banco de dados Protein Data Bank (PDB, disponível em https://www.rcsb.org/)	61
Tabela 9	Estruturas cristalográficas de enzimas de demais coronavírus selecionadas no banco de dados Protein Data Bank (PDB, disponível em https://www.rcsb.org/)	62
Tabela 10	Potenciais resíduos de interação preditos com PrankWeb nas diferentes M ^{pro}	67
Tabela 11	Potenciais resíduos de interação preditos com PrankWeb na PL ^{pro} de SARS-CoV-2	69
Tabela 12	Resíduos de interação com ligantes co-cristalizados nas M ^{pro} de SARS-CoV-2 preditos em análises com programa GOLD	75
Tabela 13	Resíduos de interação com ligantes co-cristalizados nas M ^{pro} de outros coronavírus preditos em análises com programa GOLD	76
Tabela 14	Moléculas selecionadas após o protocolo de triagem virtual contra as proteases	86
Tabela 15	Atividade inibitória (μM) contra as M ^{pro} de SARS-CoV-2, SARS-CoV e MERS-CoV	91
Tabela 16	Avaliação biológica de extratos e triterpenos contra MHV-3 em células L929	98

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

Å	Angstrom
ACE2	<i>Angiotensin converting enzyme 2</i> (enzima conversora de angiotensina 2)
°C	Grau <i>Celsius</i>
CC ₅₀	Concentração citotóxica 50%
CDC	<i>Centers for Disease Control and Prevention</i> (Centro de Controle e Prevenção de Doenças)
cDNA	DNA complementar
CE ₅₀	Concentração efetiva 50%
CMC	Carboximetilcelulose
CO ₂	Dióxido de carbono
DMSO	Dimetilsulfóxido
DNA	<i>Deoxyribonucleic acid</i> (ácido desoxirribonucleico)
EDTA	Ácido etilendiaminotetracético
ERGIC	Compartimento intermediário do retículo endoplasmático-Golgi
FDA	<i>Food and Drug Administration</i>
GMR	Gerwick, McKerrow and Roush
h	Hora(s)
HCoV	<i>Human coronavirus</i> (coronavírus humano)
HQSAR	<i>Hologram Quantitative Structure-Activity Relationship</i> (relação estrutura-atividade quantitativa por holograma)
IS	Índice de seletividade
LBDD	<i>Ligand-Based Drug Design</i> (planejamento de fármacos baseado na estrutura do ligante)
MEM	<i>Minimum essential medium</i> (meio essencial mínimo)
MERS-CoV	Coronavírus da síndrome respiratória do Oriente Médio
ML	<i>Machine learning</i> (aprendizado de máquina)
mg	Miligrama(s)
mL	Mililitro(s)
M ^{pro}	<i>Main protease</i> (protease principal)
MTT	Brometo de 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazólio
µL	Microlitro(s)
µM	Micromolar
µs	Microsegundo(s)
nm	Nanômetro

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS (*continuação*)

nM	Nanomolar
ns	Nanossegundo(s)
OMS	Organização Mundial de Saúde
PCR	Proteína C Reativa
PL ^{pro}	<i>Papain-like protease</i> (protease do tipo papaína)
RNA	<i>Ribonucleic acid</i> (ácido ribonucleico)
RNA _m	RNA mensageiro
RT	<i>Reverse transcriptase</i> (transcriptase reversa)
RT-qPCR	<i>Reverse Transcription Quantitative Polymerase Chain Reaction</i> (transcrição reversa seguida da reação em cadeia da polimerase quantitativa)
SARS	<i>Severe Acute Respiratory Syndrome</i> (síndrome respiratória aguda grave)
SARS-CoV	Coronavírus relacionado à síndrome respiratória aguda grave
SARS-CoV-2	Coronavírus 2 relacionado à síndrome respiratória aguda grave
SBDD	<i>Structure-Based Drug Design</i> (planejamento de fármacos baseado na estrutura do alvo)
SRAG	Síndrome respiratória aguda grave
U	Unidade Internacional
UCSD	<i>University of California San Diego</i>
UCSF	<i>University of California San Francisco</i>
UFP	Unidade formadora de placa
UNESP	Universidade Estadual Paulista
UFMG	Universidade Federal de Minas Gerais
v/v	Volume por volume
WHO	<i>World Health Organization</i> (Organização Mundial de Saúde)
WRR	William R. Roush
%	Porcentagem
®	Marca registrada

SUMÁRIO

1 INTRODUÇÃO	16
1.1 Família <i>Coronaviridae</i>	16
1.1.1 Estrutura da partícula e genoma dos coronavírus	17
1.1.2 Proteínas estruturais	19
1.1.3 Proteínas não estruturais	20
1.1.4 Ciclo de multiplicação	23
1.2 Doença do coronavírus 2019 (COVID-19): transmissão, sinais e sintomas clínicos	25
1.2.1 Diagnóstico laboratorial de SARS-CoV-2	27
1.2.2 Profilaxias contra COVID-19: medidas não farmacológicas e vacinas	28
1.2.3 Fármacos e tratamento da COVID-19	30
1.3 Estratégias no planejamento, descoberta e desenvolvimento de moléculas bioativas com atividade antiviral	32
1.3.1 Inibidores enzimáticos contra SARS-CoV-2	34
1.3.2 Triagem virtual de moléculas bioativas	36
1.3.3 Simulações de dinâmica molecular (DM)	37
1.3.4 Técnicas complementares: predições de sítios de ligação em alvos moleculares	38
1.3.5 Técnicas complementares: predições de propriedades farmacocinéticas, farmacodinâmicas e de toxicidade de ligantes	38
2 JUSTIFICATIVA	40
3 OBJETIVOS	41
3.1 Objetivo Geral	41
3.2 Objetivos Específicos	41
4 MATERIAL E MÉTODOS	42
4.1 Análises <i>in silico</i>	43
4.1.1 Construção dos conjuntos de dados	43
4.1.2 Relação quantitativa entre a estrutura química e a atividade biológica por holograma (HQSAR) e predição de atividade biológica	43
4.1.3 Seleção e preparo das estruturas das proteases de SARS-CoV-2	44
4.1.4 Predição dos sítios de ligação das enzimas	44
4.1.5 Acoplamento molecular (<i>docking</i>) de ligantes e potenciais moléculas bioativas	45
4.1.6 Construção de modelos farmacofóricos e triagem virtual de bancos de dados	46
4.1.7 Simulações de dinâmica molecular (DM)	47
4.1.8 Predições de propriedades farmacocinéticas, farmacodinâmicas e de toxicidade	48
4.2 Ensaios <i>in vitro</i>	48
4.2.1 Ensaios de inibição da atividade enzimática da M ^{pro} de SARS-CoV-2	48
4.2.2 Ensaios de inibição da atividade enzimática da M ^{pro} de SARS-CoV e MERS-CoV	49
4.2.3 Ensaios de inibição da atividade enzimática da PL ^{pro} de SARS-CoV-2	49
4.2.4 Vírus e linhagens celulares	49
4.2.5 Propagação viral	50
4.2.6 Titulação viral	51
4.2.7 Avaliação da concentração citotóxica 50% (CC ₅₀)	51
4.2.8 Avaliação da concentração efetiva 50% (CE ₅₀) e índices de seletividade (IS)	52

5 RESULTADOS E DISCUSSÃO	53
5.1 HQSAR de inibidores da M ^{pro} de SARS-CoV-2: triagem de moléculas bioativas e avaliação da inibição da atividade enzimática.....	53
5.2 Estratégia de planejamento de moléculas bioativas baseado na estrutura de alvos.....	63
5.2.1 Seleção de estruturas de enzimas no PDB	63
5.2.2 Mapeamento das proteases de SARS-CoV-2 e de outros coronavírus.....	64
5.2.3 Análises de interações entre ligantes e alvos moleculares de SARS-CoV-2 e diferentes coronavírus.....	74
5.2.4 Simulações de dinâmica molecular: M ^{pro} e PL ^{pro} de SARS-CoV-2	79
5.2.5 Construção de modelos farmacofóricos da M ^{pro} e PL ^{pro} de SARS-CoV-2.....	83
5.2.6 Construção dos conjuntos de dados e triagens virtuais de moléculas bioativas.....	85
5.3 Ensaio biológico	90
5.3.1 Avaliação da inibição da atividade enzimática.....	90
5.3.2 Avaliação da atividade antiviral de moléculas bioativas contra MHV-3	99
6 CONCLUSÃO	102
7 REFERÊNCIAS BIBLIOGRÁFICAS	104
APÊNDICE A – Atividade inibitória (%) da série de moléculas análogos dos <i>hits</i> (WRR) contra M^{pro} de SARS-CoV-2.....	126
APÊNDICE B – Moléculas (BraCoLi) selecionadas no modelo de HQSAR.....	129
APÊNDICE C – Alinhamento de sequências das proteases de SARS-CoV-2	132
APÊNDICE D – Percurso, atividades e produções acadêmicas.....	134

1. INTRODUÇÃO

Os coronavírus são microrganismos que infectam peixes, aves e mamíferos, e apresentam importância ambiental, veterinária e humana de saúde pública. Estes vírus possuem tropismo principalmente por células epiteliais, geralmente associados a infecções gastrointestinais e respiratórias que podem ser agudas ou se tornarem crônicas, e a transmissão ocorre por vias aéreas e oral-fecal, além de fômites. Geralmente as infecções são assintomáticas ou apresentam sintomas leves, mas algumas espécies de coronavírus causam manifestações graves, como hepatites, acometimentos neurológicos, insuficiência e falência renal, além de síndromes respiratórias agudas graves (do inglês, *Severe Acute Respiratory Syndrome*; SARS), podendo levar à morte (ICTV, 2021; LEFKOWITZ et al., 2018).

Os primeiros coronavírus foram identificados em aves na década de 1930, enquanto as primeiras espécies que infectam seres humanos tiveram relatos na década de 1960. Alguns coronavírus humanos, como Betacoronavirus 1 e os Human coronavirus (HCoV) 229E, NL63 e HKU1, causam principalmente resfriados, que podem evoluir em manifestações graves do trato respiratório inferior, principalmente em crianças e idosos. Além destes HCoV, outros três denotam relevância clínica: o coronavírus da síndrome respiratória do Oriente Médio (MERS-CoV) e os coronavírus relacionados à síndrome respiratória aguda grave, SARS-CoV e SARS-CoV-2 (AMANAT; KRAMMER, 2020; GORBALENYA et al., 2020; ICTV, 2021).

1.1 Família *Coronaviridae*

Segundo o Comitê Internacional de Taxonomia de Vírus (do inglês, *International Committee on Taxonomy of Viruses*; ICTV) (LEFKOWITZ et al., 2018), a família *Coronaviridae* se enquadra no reino *Riboviria* e ordem *Nidovirales*. A subfamília *Orthocoronavirinae* compreende os gêneros *Alphacoronavirus* e *Betacoronavirus* (usualmente associados a infecções em mamíferos), e *Deltacoronavirus* e *Gammacoronavirus* (comumente associados a infecções em aves). Estes são vírus envelopados com partículas de formato esférico e cerca de 120 a 160 nm de diâmetro, com um genoma de RNA fita simples e polaridade positiva (ssRNA+), linear e não segmentado, contendo aproximadamente 26.000 a 32.000 bases nitrogenadas (GORBALENYA et al., 2020; ICTV, 2021).

Atualmente 52 espécies são descritas por causar infecções em aves e mamíferos, seis delas afetando especificamente seres humanos (MERS-CoV, SARS-CoV, Betacoronavirus 1 e os HCoV-229E, NL63 e HKU1). Os coronavírus foram assim nomeados de acordo com sua

aparência similar à de uma coroa ao redor da partícula viral (do latim, “*corona*”), em razão das glicoproteínas de superfície desses vírus (**Figura 1**), quando observados sob técnicas de microscopia eletrônica (CUI; LI; SHI, 2019; GORBALENYA et al., 2020; ICTV, 2021).

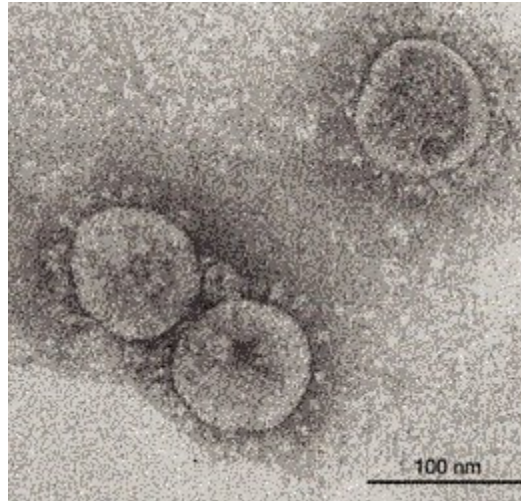


Figura 1. Partículas virais de SARS-CoV observadas por microscopia eletrônica de transmissão (MET). As partículas virais, cultivadas em células Vero (linhagem de epitélio renal de macaco verde africano), apresentam glicoproteínas espículas (do inglês, *spike*) em formato de clavas, resultando em uma aparência da partícula similar a uma coroa. Fonte: (STADLER et al., 2003) (*adaptado*).

Baseado em sequências genômicas depositadas em bancos de dados públicos, os coronavírus que infectam seres humanos têm origem de infecções em outros mamíferos, sendo SARS-CoV-2 (LAU et al., 2020; WACHARAPLUESADEE et al., 2021), SARS-CoV, MERS-CoV, HCoV-NL63 e HCoV-229E considerados originários de morcegos, enquanto HCoV-OC43 (Betacoronavirus 1) e HCoV-HKU1 possivelmente originados de roedores (CUI; LI; SHI, 2019). Nesse âmbito, animais domésticos, além de poderem sofrer doenças causadas por coronavírus transmitidos por morcegos, podem ter papéis importantes como hospedeiros intermediários durante um *spillover*, ou seja, uma transmissão de vírus de hospedeiros naturais para seres humanos (CUI; LI; SHI, 2019).

1.1.1 Estrutura da partícula viral e genoma dos coronavírus

O envelope da partícula madura dos coronavírus apresenta as proteínas de membrana (M) e envelope (E) associadas à sua membrana, enquanto a proteína *spike* (S) é exposta na face externa da partícula. O genoma de RNA interage com a proteína do nucleocapsídeo (N), formando o nucleocapsídeo com simetria helicoidal (**Figura 2**).

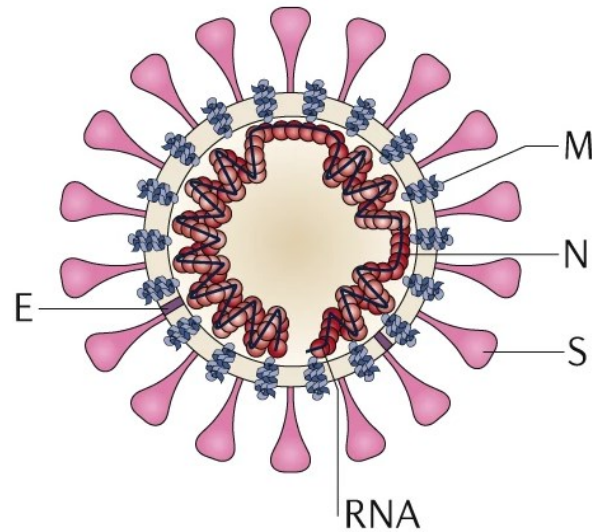


Figura 2. Representação de partícula viral dos coronavírus. Estrutura de uma partícula viral madura dos coronavírus. Corte transversal, indicando por linhas a proteína de envelope E (violeta), a proteína de membrana M (azul), a proteína do nucleocapsídeo N (vermelho) e a proteína S (rosa), envolvendo o genoma viral de RNA (preto). Fonte: (CUI; LI; SHI, 2019) (*adaptado*).

O genoma dos coronavírus (ssRNA⁺), um dos maiores dentre os vírus de RNA (entre 26.000 e 32.000 bases nitrogenadas), contém uma estrutura *cap* na porção 5' e uma cauda poli-A (poliadenilato; repetições de adenina) na porção 3', permitindo que ele atue como RNA mensageiro (RNAm) para a tradução das poliproteínas ou seja parcialmente traduzido na formação do complexo de replicação. A extremidade 5' do genoma contém uma sequência líder e uma região não traduzida (do inglês, *untranslated region*; UTR) que contém estruturas secundárias (elemento *cis-acting*) para a replicação e transcrição de RNA (FEHR; PERLMAN, 2015).

Na organização do genoma, aproximadamente dois terços do genoma de RNA incluem genes das 16 proteínas não estruturais (do inglês, *non structural protein*; nsp), que são traduzidas em duas poliproteínas, pp1a e pp1ab, a partir da expressão de suas respectivas janelas abertas de leitura (do inglês, *open reading frame*; ORF), ORF1a e ORF1b (FEHR; PERLMAN, 2015; V'KOVSKI et al., 2021). Cerca de um terço restante do genoma, a partir de outras ORFs, codifica as proteínas acessórias, intercalados entre os genes de proteínas estruturais (E, M, N e S) na extremidade 3', a exemplo do genoma de SARS-CoV-2 (**Figura 3**). Estas proteínas acessórias, embora não essenciais para replicação viral, já demonstraram importância na patogênese viral, como modulação e evasão do sistema imune (V'KOVSKI et al., 2021; REDONDO et al., 2021).

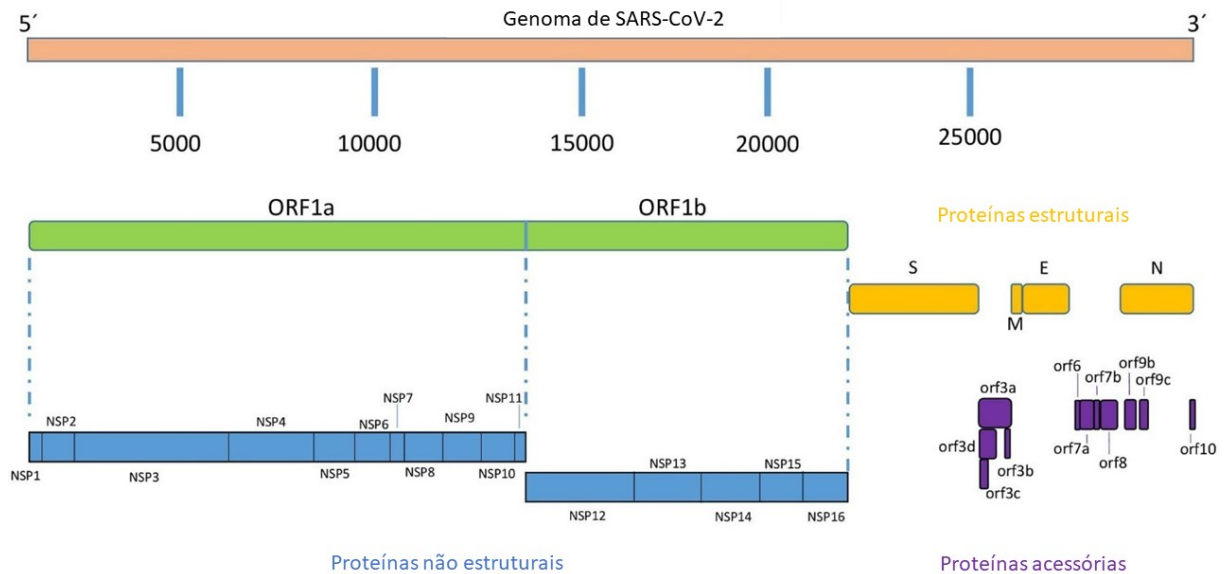


Figura 3. Modelo de organização de proteínas traduzidas do genoma dos coronavírus. Esquema representativo do RNA polaridade positiva (salmão) de SARS-CoV-2 com a tradução das 16 proteínas não estruturais (azul), a partir das respectivas ORF1a e ORF1b (verde), além de quatro proteínas estruturais (amarelo) e 11 acessórias (roxo), orf3a-d, orf6, orf7a-b, orf8, orf9b-c e orf10. Logo abaixo, as poliproteínas pp1a e pp1ab (azul) representam as 16 proteínas não estruturais (nsp1-nsp16). Linhas pretas verticais indicam a região de clivagem das poliproteínas pelas proteases virais (nsp3 e nsp5). Fonte: (REDONDO et al., 2021) (*adaptado*).

1.1.2 Proteínas estruturais

Dentre as quatro proteínas estruturais, a glicoproteína S (~150 kDa) se apresenta como uma proteína de fusão composta por homotrímeros, compondo uma estrutura trimérica distinta na superfície do vírus, responsável por mediar a interação com o receptor celular do hospedeiro, como a enzima conversora de angiotensina 2 (do inglês, *angiotensin converting enzyme 2*; ACE2) na infecção por SARS-CoV-2 (BOSCH et al., 2003; HUANG et al., 2020b; YAN et al., 2020). Este processo ocorre por interação do domínio de ligação do receptor (do inglês, *receptor binding domain*; RBD), via motivo de ligação do receptor (do inglês, *receptor binding motif*; RBM).

Esta estrutura se apresenta em uma das duas subunidades polipeptídicas (S1 e S2) da proteína S, que é clivada, por exemplo, por uma protease do tipo furina da célula hospedeira, bem como por serinoprotease transmembrana II humana (TMPRSS2) e/ou catépsina L, separando as duas subunidades (HOFFMANN et al., 2020; YAN et al., 2020). Além disso, também se mostra como alvo molecular de interesse para o desenvolvimento de testes de diagnóstico (BAKER et al., 2020) e candidatos vacinais (SALVATORI et al., 2020), além de potenciais inibidores candidatos a fármacos (HUANG et al., 2020b).

A proteína E (8-12 kDa), por sua vez, é transmembranar, sendo encontrada em pequenas quantidades na partícula viral e atuando como canal iônico. Já foi observado que a remoção do gene da proteína E em SARS-CoV recombinantes, por exemplo, é capaz de atenuar o vírus, reduzindo a letalidade em modelos murinos. Além de facilitar a montagem e liberação do vírus, esta proteína também possui papel importante na patogênese dos coronavírus, retardando processos de sinalização celular ao interferir no balanço iônico da célula, além de auxiliar na liberação de RNA viral no desnudamento da partícula. (DEDIEGO et al., 2007; NIETO-TORRES et al., 2014; CAO et al., 2021).

Já a proteína M (25-30 kDa) possui três domínios transmembranares, modulando o formato da partícula viral enquanto inserido na membrana do retículo endoplasmático. Esta proteína também pode adotar duas conformações distintas, permitindo promover uma curvatura da membrana e também se ligar ao nucleocapsídeo viral, essencial no processo de montagem da partícula (NAL et al., 2005; NEUMAN et al., 2011; ZHANG et al., 2022).

Finalmente, a proteína N é a única proteína constituindo o nucleocapsídeo viral, composta por dois domínios separados, ambos necessários para se ligar ao genoma. A fosforilação dessa proteína permite uma mudança conformacional necessária à ligação com o RNA, importante para garantia de especificidade na interação entre o RNA viral e o RNA não viral do hospedeiro. Além disso, também é capaz de se ligar à protease do tipo papaína (PL^{pro}), e à proteína M, auxiliando na estabilização do genoma no complexo de replicação e durante sua encapsidação na montagem das partículas virais (CHANG et al., 2006; HURST; KOETZNER; MASTERS, 2013; BAI et al., 2021). Ademais, esta proteína também possui importante papel imunogênico (DOBAÑO et al., 2021), sendo alvo em potencial, por exemplo, no desenvolvimento de testes de diagnóstico (PENG et al., 2021) e de candidatos vacinais (CASTRO et al., 2022).

1.1.3 Proteínas não estruturais

Dentre as 16 nsp contidas nas poliproteínas pp1a e pp1ab dos coronavírus, duas ou três proteases são responsáveis por suas clivagens e essenciais na multiplicação dos coronavírus. A protease principal (do inglês, *main protease*; M^{pro} ou nsp5) ou 3-quimiotripsina (do inglês, *3-chymotrypsin-like protease*; 3CL^{pro}) é uma cisteína protease responsável pelas clivagens de nsp5 a nsp16 (MIELECH et al., 2014; CHEN; LIU; GUO, 2020), sendo alvo do fármaco nirmatrelvir (OWEN et al., 2021). Já a PL^{pro} (nsp3), também uma cisteína protease, é responsável pelas clivagens de nsp1 a nsp4, sendo que a maioria dos coronavírus codifica

duas PL^{pro} na nsp3, exceto MERS-CoV, SARS-CoV e SARS-CoV-2 (CHEN; LIU; GUO, 2020; MIELECH et al., 2014). Além destas, a RNA polimerase dependente de RNA (do inglês, *RNA-dependent RNA polymerase*; RdRp ou nsp12) atua na replicação do RNA genômico viral, ante a etapa de tradução das poliproteínas no complexo de replicação (CHEN; LIU; GUO, 2020), sendo alvo do fármaco antiviral rendesivir (GOTTLIEB et al., 2022).

Dentre estas três enzimas, a M^{pro} (**Figura 4A**) de SARS-CoV-2, um homodímero com duas cadeias principais (A e B) marcada por três domínios (I, II e III), contendo uma díade catalítica de cisteína e histidina (His41 e Cys145), responsável pela hidrólise das ligações peptídicas das poliproteínas (JIN et al., 2020a; RAMOS-GUZMÁN; RUIZ-PERNI; TUÑÓN, 2020). Além de ser conservada entre diversos coronavírus, sua atividade essencial no complexo de replicação, bem como a ausência de homólogos nos seres humanos, a torna um alvo em potencial para o desenvolvimento de inibidores da multiplicação viral (JIN et al., 2020a; CITARELLA et al., 2021).

A PL^{pro} (**Figura 4B**), também uma cisteína protease com tríade catalítica clássica (Cys111, His272, Asp286), é capaz de promover atividade deubiquitinase e regular resposta do tipo interferon, sendo mais um alvo de interesse para obtenção de inibidores da multiplicação viral (SHIN et al., 2020; SERAFIM et al., 2021a). Já a RdRp (**Figura 4C**) atua na replicação de RNA viral, sendo capaz de iniciar a síntese de RNA, com auxílio de demais nsp, como nsp7 e nsp8 (cofatores). Sendo uma enzima conservada entre os diferentes coronavírus, possui resíduos de aminoácidos compartilhados por todas as RdRp, também podendo ser um alvo em potencial para terapias anti-CoV (AHN et al., 2012; KIRCHDOERFER; WARD, 2019; SERAFIM et al., 2021a).

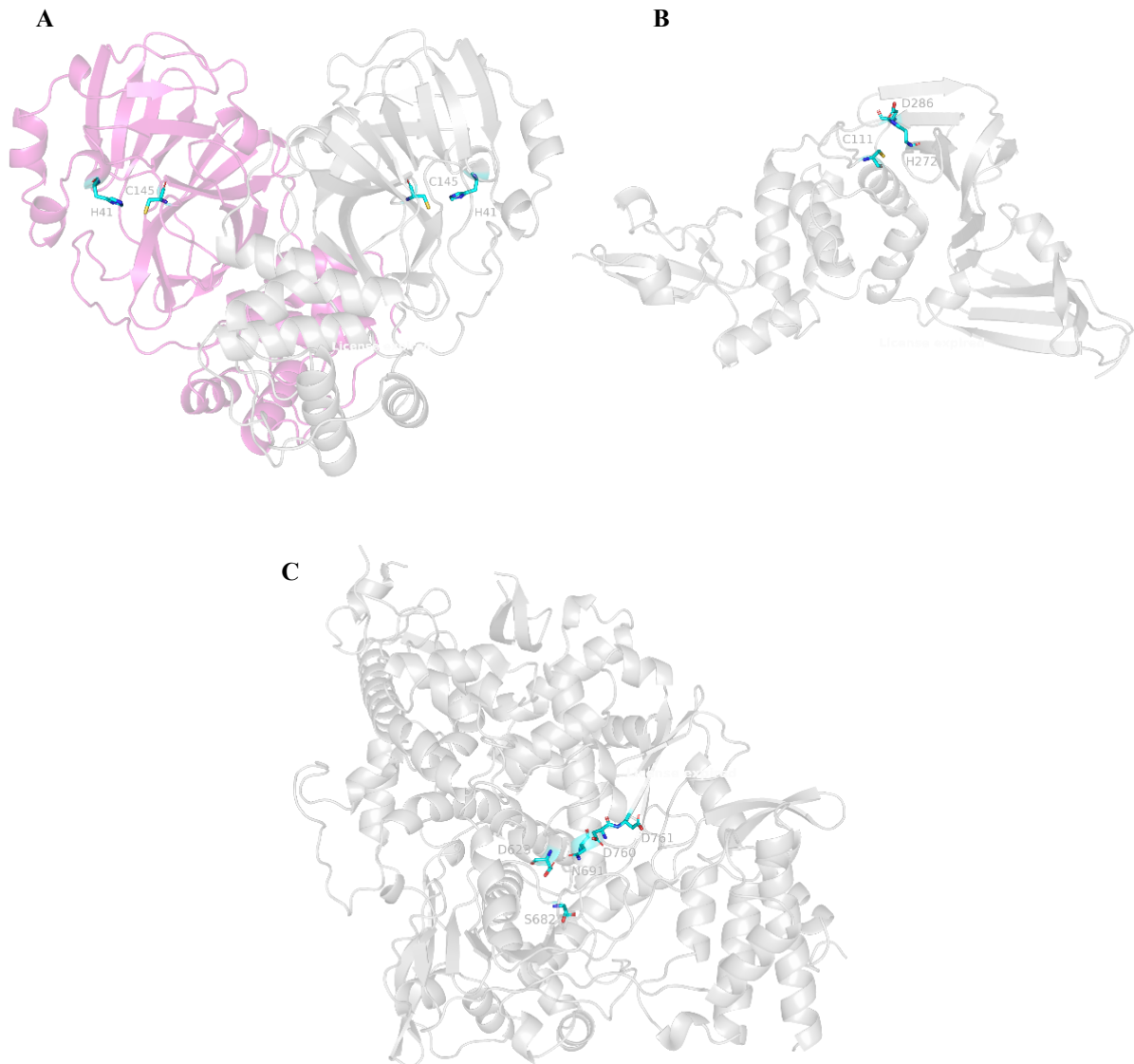


Figura 4. Estruturas tridimensionais das enzimas de SARS-CoV-2. (A) Estrutura representativa em modelo de *cartoon* da M^{pro} de SARS-CoV-2 (PDB ID: 7D1M). Porção da cadeia A (rosa) e cadeia B (cinza) exibindo resíduos da díade catalítica His41 e Cys145 (ciano). (B) Estrutura representativa em modelo de *cartoon* da PL^{pro} (cinza) de SARS-CoV-2 (PDB ID: 6WX4) exibindo resíduos da tríade catalítica Cys111, His272 e Asp286 (ciano). (C) Estrutura representativa em modelo de *cartoon* da RdRp (cinza) de SARS-CoV-2 (PDB ID: 7BV2) exibindo os resíduos catalíticos Asp623, Ser682, Asn691, Asp760 e Asp761 (ciano).

Além dessas três enzimas, outras proteínas foram descritas no ciclo de multiplicação dos coronavírus (**Tabela 1**). A nsp1, por exemplo, é capaz de interferir na resposta imune inata dos hospedeiros, como na inibição de interferon. Já a nsp13 atua como helicase e RNA trifosfatase, enquanto a nsp14 como exoribonuclease e a nsp16 como metiltransferase (CHEN; LIU; GUO, 2020; GADHAVE et al., 2021; XU et al., 2022).

Tabela 1. Funções das proteínas não estruturais (nsp) dos coronavírus

Proteína	Funções principais	Proteína	Funções principais
nsp1	Degradação de RNAm do hospedeiro. Inibição de resposta por interferon.	nsp9	Dimerização proteica. Ligação com RNA.
nsp2	Inibição de resposta por interferon.	nsp10	Cofator de nsp14 e nsp16
nsp3	PL ^{pro} : clivagem da pp1a. Inibição de sinalização celular.	nsp11	Desconhecida (sugerida atuando em dinâmicas conformacionais).
nsp4	Formação de vesículas de dupla membrana (VDM).	nsp12	RNA polimerase dependente de RNA
nsp5	M ^{pro} : clivagem da pp1ab.	nsp13	Helicase e RNA trifosfatase
nsp6	Restrição de autofagossomo. Formação de VDM.	nsp14	Exoribonuclease. Metiltransferase.
nsp7	Cofator da RNA polimerase	nsp15	Endoribonuclease. Evasão de sensores de dsRNA.
nsp8	Cofator da RNA polimerase	nsp16	Metiltransferase. Regulação imune.

Fonte: (CHEN; LIU; GUO, 2020; GADHAVE et al., 2021; XU et al., 2022) (*adaptado*).

1.1.4 Ciclo de multiplicação

Assim como alguns vírus ssRNA+, a formação da progênie dos coronavírus ocorre no citoplasma das células do hospedeiro. Na infecção por SARS-CoV-2 (SHANG et al., 2020; WAN et al., 2020), por exemplo, as partículas virais se adsorvem a receptores na superfície celular por interação da proteína S (RBM) com o receptor ACE2. Alguns coronavírus que infectam porcos (PDEV e TGEV), bem como HCoV-229E, utilizam aminopeptidases como o receptores celulares ao invés de ACE2 (YEAGER et al., 1992; CUI et al., 2020).

A penetração da partícula viral pode ser mediada por duas vias simultâneas (JACKSON et al., 2022): (i) fusão direta com a membrana da célula do hospedeiro, via clivagem proteolítica da proteína S por furina ou TMPRSS2, liberando o genoma de RNA do vírus no citoplasma celular (HOFFMANN et al., 2020; JACKSON et al., 2022); e/ou (ii) via endossomo, com internalização mediada por clatrina, seguida de uma clivagem proteolítica ácido-dependente da proteína S por catepsina L (ZHAO et al., 2021), fundindo a membrana do envelope viral com o endossomo celular, seguido da liberação do genoma viral no citoplasma (FEHR; PERLMAN, 2015; MALONE et al., 2022).

O RNA então pode ser replicado ou diretamente traduzido e, após a tradução, as proteínas estruturais são clivadas pelas duas proteases PL^{pro} e M^{pro}. Este processo se inicia com a auto clivagem (nsp5), seguida da hidrólise das nsp6 a nsp16 e inserção das proteínas na membrana do retículo endoplasmático, seguindo a via secretora do compartimento intermediário do retículo endoplasmático-Golgi (ERGIC), durante a montagem das partículas (DE WIT et al., 2016).

Alguns coronavírus como HCoV-OC43 (Betacoronavirus 1) e HKU1 também codificam uma hemaglutinina esterase, uma proteína multifuncional que atua principalmente na adesão à célula alvo e na liberação da partícula após a exocitose (LANG et al., 2020). Paralelamente, o genoma replicado se associa à proteína N e, por brotamento, migra para o ERGIC finalizando a montagem das partículas virais. As partículas são então transportadas para a superfície celular em vesículas e são liberadas via exocitose (**Figura 5**) (FEHR; PERLMAN, 2015; MALONE et al., 2022).

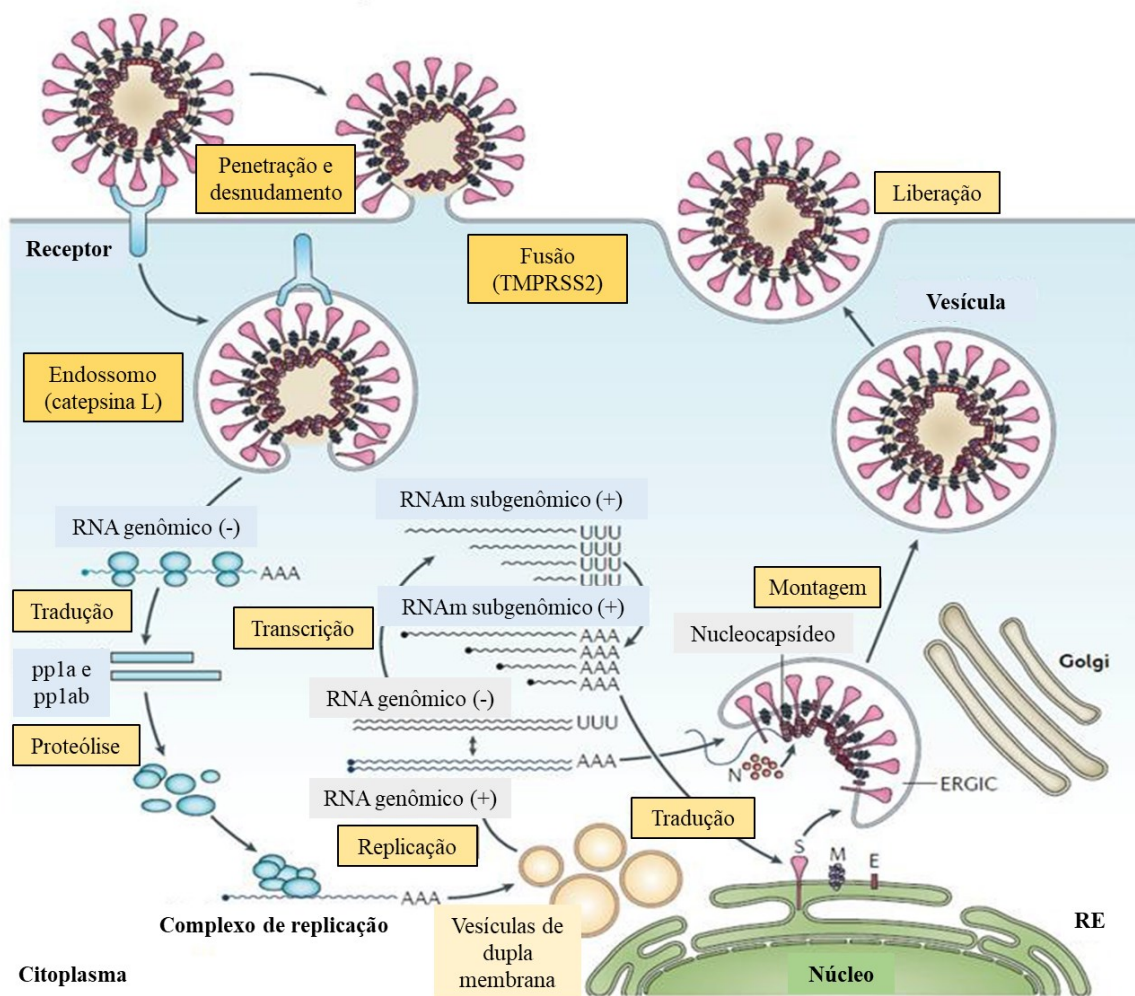


Figura 5. Ciclo de multiplicação dos coronavírus. Partículas virais maduras e infectivas (vírions) podem se ligar via domínio de ligação de receptor (RBD) a um grupo de receptores celulares, por exemplo, ACE2 na infecção por SARS-CoV-2, aderindo à célula do hospedeiro. A internalização das partículas virais pode ocorrer via fusão direta com a membrana celular, após clivagem da proteína S por furina ou TMPRSS2; e/ou via endossomo, com a fusão direta da membrana viral com a membrana celular, após clivagem por catepsina L engatilhada no pH ácido. Ambas acabam promovendo o desnudamento da partícula e liberação do genoma de RNA no citoplasma celular, que pode então ser traduzido. As duas poliproteínas (pp1a e pp1ab) são então processadas por proteases virais e celulares. Proteínas não estruturais do vírus promovem a transcrição e/ou replicação do genoma e, em seguida, o RNA recém-sintetizado se associa à proteína de nucleocapsídeo (N). A montagem das partículas virais ocorre seguindo a via secretora do compartimento intermediário do retículo endoplasmático-Golgi (ERGIC), sendo transportadas em vesículas para a superfície celular e liberadas por exocitose. Fonte: (DE WIT et al., 2016) e (JACKSON et al., 2022) (*adaptado*).

1.2 Doença do coronavírus 2019 (COVID-19): transmissão, sinais e sintomas clínicos

Similar ao primeiro surto de SARS-CoV na província de Guangdong (China) em 2003 (ROSLING; ROSLING, 2003), casos de uma pneumonia viral de etiologia desconhecida foram notificados à Organização Mundial da Saúde (OMS). No dia 30 de dezembro de 2019 a Comissão Municipal de Saúde de Hubei (China) casos similares foram notificados à OMS, que decretaria uma emergência internacional no final de janeiro de 2020 e, ainda no início de março do mesmo ano, a pandemia da então doença do coronavírus 2019 (COVID-19) (CIOTTI et al., 2019; XU et al., 2020).

O agente etiológico, SARS-CoV-2 (anteriormente denominado “*2019 novel coronavirus*” ou 2019-nCoV), viria a ser isolado em linhagens celulares Vero e Huh7 ainda em fevereiro de 2020 e o sequenciamento do genoma ocorreria poucos dias depois das primeiras notificações da nova doença (CIOTTI et al., 2019; ZHANG et al., 2020b). Cerca de três anos após os primeiros relatos, até o momento de redação do presente trabalho (25 de janeiro de 2023), mais de 664,8 milhões de casos e 6,72 milhões de mortes associadas foram notificados no mundo, sendo cerca de 696 mil óbitos somente no Brasil (WHO, 2023).

Usualmente, a transmissão de outros betacoronavírus ocorre por gotículas expiradas, aerossóis, superfícies ou objetos contaminados (ambiental) e via oral-fecal (LI et al., 2005; ONG et al., 2020; OTTER et al., 2016; YU; SUNG, 2004). SARS-CoV-2, além da transmissão ambiental e de indivíduos sintomáticos, apresenta também uma relevante transmissão assintomática, de indivíduos que não apresentaram ou evidenciaram sintomas, dificultando a prevenção da disseminação do vírus e facilitando contágio de demais indivíduos (BAI et al., 2020; FERRETTI et al., 2020; TONG et al., 2020).

Uma vez transmitido, o vírus pode ainda apresentar longos períodos de incubação, de até 14 dias (média de 3-5 dias), podendo ou não evoluir com aparecimento de sintomas da doença (HUANG et al., 2020a). Estão entre os sintomas mais comuns: febre, tosse, fadiga, dores de cabeça, dificuldades para respirar (dispneia), perda total ou parcial de olfato (anosmia) e paladar (ageusia), dores de garganta, dores musculares, diarreia, vômitos, coriza e/ou congestão nasal. Estes podem durar, em média, até 14 dias, podendo se estender por mais de seis semanas em casos graves (CHEN et al., 2020; LONG et al., 2020; TONG et al., 2020).

A doença pode ser dividida de acordo com os sinais e sintomas clínicos e em razão da gravidade e progressão do quadro clínico de cada indivíduo (SIDDIQI; MEHRA, 2020). Na fase inicial, o vírus infecta células epiteliais principalmente do trato respiratório superior, nas regiões da nasofaringe e orofaringe, sendo evidenciados sintomas inespecíficos e/ou similares a outras doenças respiratórias, usualmente resultando em casos leves. O vírus poderia potencialmente seguir uma via alternativa, infectando bulbos e neurônios olfativos (UEHA et al., 2022), o que poderia estar associado às perdas de sensibilidade no olfato e/ou paladar.

Contudo, tais hipóteses foram desconsideradas em seres humanos, sendo as disfunções olfativas e paliativas sugeridas em duas hipóteses viáveis (BUTOWT; BILINSKA; VON BARTHELD, 2023): (i) perda de células suporte no epitélio olfativo, com consequências para função neuronal local; e (ii) danos causados pela resposta imune, afetando a função neuronal de receptores olfativos. Contudo, análises de biópsias sugeriram a mucosa olfativa como possível via para invasão de SARS-CoV-2 em alcançar a região encefálica (MEINHARDT et al., 2021) e acometer o sistema nervoso central (SNC), sugerindo os sinais neurológicos já em uma fase intermediária da doença (LI; XUE; XU, 2020).

A fase intermediária ou pulmonar, por sua vez, apresenta evolução dos sintomas respiratórios, com acometimento pulmonar intenso, em razão da infecção em células do trato respiratório inferior. Nesse momento, a doença pode progredir de casos leves para casos moderados a graves, na qual o estado imune pode não ser capaz de controlar com eficiência o processo inflamatório, comprometendo ainda mais os tecidos acometidos. Uma intensa resposta inflamatória pode levar ao extravasamento celular intenso, aumento da permeabilidade vascular e edema local, agravando, por exemplo, casos pulmonares (SHI et al., 2020; TAY et al., 2020). A progressão da doença nesse estágio também pode evoluir para uma síndrome respiratória aguda grave (SRAG), com o comprometimento da função respiratória, evidenciada por dispneia e hipóxia tecidual, além de baixa saturação de oxigênio.

Esse quadro pode ser acompanhado de acometimento dos tecidos renal, hepático e/ou cardíaco, com uma resposta sistêmica exacerbada, agravando o quadro clínico dos indivíduos (grave a crítico) e podendo ocasionar, por exemplo, em uma síndrome coronariana aguda, miocardites, falência renal, parada cardiorrespiratória, até a morte do indivíduo acometido (LI et al., 2020; TAY et al., 2020).

1.2.1 Diagnóstico laboratorial de SARS-CoV-2

O diagnóstico da infecção por SARS-CoV-2 pode ocorrer pela detecção molecular do material genético do vírus, empregando a técnica de transcrição reversa seguida da reação em cadeia da polimerase quantitativa (RT-qPCR), além de testes imunocromatográficos, imunofluorescentes ou de quimioluminescência para detecção de antígenos virais ou anticorpos (IgA, IgM e/ou IgG) (GUO et al., 2020; HAVERI et al., 2020). A RT-qPCR, “padrão-ouro” no diagnóstico, foi empregada desde o início da pandemia da COVID-19, com diversos protocolos para detecção de genes alvo de SARS-CoV-2 desenvolvidos e reconhecidos pela OMS, a exemplo do protocolo do CDC dos Estados Unidos, com dois (inicialmente três) alvos no gene da proteína N (WHO, 2020). Também é importante ressaltar a variação da sensibilidade do teste em diferentes amostras biológicas, por exemplo, swabs nasofaríngeo e orofaríngeo, escarro, lavado broncoalveolar, biópsia de fibronoscopia, além de fezes, urina e sangue (WANG et al., 2020b).

Nesse sentido, alguns fatores como a carga viral e o momento da coleta são cruciais para uma detecção adequada de RNA viral (WANG et al., 2020). Ainda, uma RT-qPCR negativa não deve ser utilizada como único critério de decisão para se descartar o diagnóstico de uma infecção por SARS-CoV-2. Embora um período ideal de coleta para sintomáticos possa variar usualmente entre os primeiros três a sete dias de sintomas, coletas em dias subsequentes e/ou tipos diferentes de amostras podem ser necessárias para confirmar o diagnóstico. Igualmente, a coleta e transporte adequados da amostra, bem como a qualidade da extração do RNA, também podem influenciar nos resultados (TO et al., 2020; ZHANG et al., 2020b).

Nesse contexto, os ensaios de neutralização e detecção de anticorpos também foram utilizados para auxiliar e/ou complementar o diagnóstico da COVID-19, incluindo casos com RT-qPCR negativos, além de serem indicados para avaliação estado imune de pacientes em uma população (TO et al., 2020), a exemplo da imunização pós-vacinas. Estes ensaios, como ELISA (do inglês, *Enzyme-Linked Immunosorbent Assay*) e os chamados testes rápidos

(ensaios imunocromatográficos ou de *point-of-care*) (LU; STRATTON; TANG, 2020), também permitem a detecção de anticorpos ou antígenos (BRÜMMER et al., 2021; DINNES et al., 2021), e já apresentavam a possibilidade de autotestagem ainda no final de 2021 (CASSUTO et al., 2021; HARMON et al., 2021).

1.2.2 Profilaxias contra COVID-19: medidas não farmacológicas e vacinas

A eficácia de medidas de controle e profilaxia da COVID-19 depende de vários parâmetros sociais, econômicos e epidemiológicos. Modelos e estratégias de quarentena, distanciamento e isolamento social, junto de medidas comunitárias restritivas puderam mitigar a disseminação da doença, como o fechamento ou restrições de serviços, estabelecimentos comerciais e de instituições de ensino, bem como de restrições para frequentar locais públicos e eventos fechados, e o uso de máscaras faciais para proteção individual (AQUINO et al., 2020; KOO et al., 2020; ORABY et al., 2021; TALIC et al., 2021).

O chamado “achatamento de uma curva epidêmica” (SILVA; FIGUEIREDO FILHO; FERNANDES, 2020), ou seja, redução da progressão de infecções e contaminações por uma doença em um surto ou epidemia foi uma importante estratégia contra COVID-19. No entanto, quanto mais restritivas e duradouras as medidas, maiores tendem a ser os impactos socioeconômicos (BARTIK et al., 2020), além do acometimento da própria saúde física e mental dos indivíduos a longo prazo (ORNELL et al., 2020; KOLA et al., 2021).

Naturalmente, esse equilíbrio da implementação e a efetividade de medidas restritivas e de distanciamento em cada país foi dependente de diversos fatores (HOGAN; HOWLETT; MURPHY, 2022), como as normas implementadas por autoridades locais, lideranças político-governamentais, instituições de referência em pesquisa e desenvolvimento, além de questões socioculturais e de saúde pública (HAUG et al., 2020). Não obstante, o quão imediato for a implementação de diferentes medidas (ORABY et al., 2021), mais eficiente e precoce tende a se tornar uma redução do número de casos e mortes, como evidenciado para COVID-19 e em doenças causadas por outros coronavírus no passado (TAN, 2006; LEWNARD; LO, 2020).

Nesse âmbito, uma medida usualmente considerada a primeira linha de defesa utilizada contra o agravamento de infecções virais são as vacinas (JOSEFSBERG; BUCKLAND, 2012; MCCLOSKEY et al., 2014), que no decorrer da pandemia da COVID-19 tiveram o seu desenvolvimento, aprovação (emergencial), distribuição e aplicação em tempo recorde no Brasil e no mundo (MOGHADAS et al., 2021; STODDARD et al., 2021;

EMANUEL; OSTERHOLM; GOUNDER, 2022). O processo de desenvolvimento de uma vacina pode ser longo (> 10 anos) e apresenta desafios de produção e alcance em massa à população, desde uma fase exploratória, ensaios pré-clínicos, fases clínicas (I, II, III e IV), uma revisão e aprovação regulatória, até chegar à produção e controle de qualidade subsequentes (WOUTERS et al., 2021).

Vacinas podem ser baseadas em algumas tecnologias, como vírus vivo atenuado, vírus inativado ou de subunidade, proteínas recombinantes e de vetores virais, além de estratégias mais recentes, como vacinas de DNA e RNAm (JOSEFSBERG; BUCKLAND, 2012; RAUCH et al., 2018). Estas também foram implementadas contra SARS-CoV-2 em colaborações público-privadas direcionados para garantir um acesso global às vacinas (AMANAT; KRAMMER, 2020; WOUTERS et al., 2021), que em poucos meses apresentaram resultados de candidatos vacinais até aprovações para uso emergencial (FOLEGATTI et al., 2020; JACKSON et al., 2020; LOGUNOV et al., 2020; MULLIGAN et al., 2020; XIA et al., 2020), a exemplo da vacina de RNAm desenvolvida pela Pfizer-BioNTech em menos de 11 meses (POLACK et al., 2020).

Vale ressaltar alguns fatores que fundamentam esse êxito em curto período, como resultados de outras vacinas e candidatos vacinais no passado (AMANAT; KRAMMER, 2020; TROGEN; OSHINSKY; CAPLAN, 2020; SU; DU; JIANG, 2021), bem como do conhecimento adquirido no desenvolvimento de candidatos vacinais (recombinantes, atenuadas, inativadas e vetorizadas) para outros coronavírus como SARS-CoV (WEINGARTL et al., 2004; TSENG et al., 2012; PADRON-REGALADO, 2020) e MERS-CoV (AGRAWAL et al., 2016; WANG et al., 2016; HOUSER et al., 2017). Contudo, ainda que os esquemas de imunização em doses múltiplas e de reforço tenham prolongado a modulação de uma resposta imune duradoura contra infecções por SARS-CoV-2 (BARDA et al., 2021), a existência de diversas linhagens e novas subvariantes dificultam a manutenção da cobertura vacinal (RUBIN, 2021; THOMPSON, 2022).

Adicionalmente, ainda que efeitos vacinais adversos graves possam ocorrer, como reações alérgicas (anafilaxias), miocardites, tromboembolismo e a síndrome de Guillain-Barré, estes são raros, na ordem de poucas dezenas por milhão de indivíduos (CDC, 2020), e o benefício individual e coletivo da vacinação sobressai a tais riscos (LAU; GALEA, 2022), que podem ser mais de 14 vezes superiores em indivíduos que desenvolvem a doença (HIPPISEY-COX et al., 2021). Além disso, a distribuição e aquisição de vacinas não é

uniforme entre países de alta e baixa renda per capita (DUAN et al., 2021). Finalmente, o negacionismo, a hesitação e desinformações a respeito das vacinas (ULLAH et al., 2021; GORMAN et al., 2022) retardam o processo de cobertura vacinal com eficiência, facilitando a disseminação viral e uma maior probabilidade de novas linhagens (RUBIN, 2021; OLIVERA MESA et al., 2022). Nesse sentido, estratégias de reposicionamento, descoberta e desenvolvimento de fármacos também são primordiais no combate à COVID-19.

1.2.3 Fármacos e tratamento da COVID-19

Além da primeira linha de defesa contra infecções virais empregando as vacinas, os antivirais se apresentam como um fármaco capaz de inibir a multiplicação e consequente disseminação viral (CDC, 2019). Considerando o contexto da pandemia da COVID-19, estratégias no desenvolvimento e obtenção de novos fármacos antivirais são de interesse, a exemplo da estratégia de reposicionamento de fármacos. Fármacos de reposicionamento são medicamentos que já foram avaliados quanto à sua segurança em ensaios pré-clínicos e clínicos, foram aprovados para uso clínico e disponibilizados comercialmente, e podem ser reposicionados como novas opções terapêuticas para outras doenças ou condições clínicas, reduzindo tempo e custos associados nas etapas de desenvolvimento de um novo medicamento (PUSHPAKOM et al., 2018).

Nesse cenário foram considerados e avaliados diversos fármacos em potencial, como rendesivir, favipiravir, ribavirina, lopinavir e ritonavir, bem como azitromicina e hidroxicloroquina (ALPERN; GERTNER, 2020; ARUN et al., 2020; KALIL, 2020), além de dexametasona como uma opção terapêutica anti-inflamatória (KELLER et al., 2020; MAGRO, 2020). Rendesivir, um fármaco antiviral análogo de nucleosídeo para tratamento de ebola, apresentou atividade *in vitro* contra SARS-CoV-2 ($CE_{50} = 0,77 \mu\text{M}$; $SI > 129,87$) (WANG et al., 2020a), além de ter demonstrado eficácia contra SARS-CoV e MERS-CoV *in vitro* e em modelos animais, como camundongos e primatas não-humanos (WARREN et al., 2016; MULANGU et al., 2019).

O uso controlado de rendesivir nos estudos clínicos da COVID-19 demonstrou redução média de até 5 dias na hospitalização e uma estimativa entre 5 e 87% de redução na mortalidade em diferentes grupos, particularmente de casos mais graves, levando à aprovação deste medicamento ainda em outubro de 2020 (BEIGEL et al., 2020; SHEAHAN et al., 2020; GOTTLIEB et al., 2022). Contudo, poucos foram os potenciais fármacos que alcançaram

algum sucesso as últimas etapas de estudos e fases clínicas contra COVID-19. Molnupiravir, por exemplo, um inibidor de RdRp (via oral) seguiu para fases clínicas I a III (COX; WOLF; PLEMPER, 2021; KHOO et al., 2021) e, posteriormente, recebeu aprovação emergencial para uso no tratamento da doença em dezembro de 2021, especialmente nos casos leves e de pacientes não hospitalizados (LEE; HSIEH; KO, 2021; JAYK BERNAL et al., 2022).

Igualmente, em dezembro de 2021 (OWEN et al., 2021) também foi aprovado o uso emergencial de uma terapia oral combinada de nirmatrelvir (**Figura 6**) e ritonavir. Este medicamento, denominado comercialmente como PAXLOVID™ (PFIZER, 2021), apresentou atividade inibitória contra a M^{pro} de SARS-CoV-2 e demonstrou redução em até 89% (OWEN et al., 2021) de casos leves a moderados nos estudos clínicos de fase I a III (HUNG et al., 2022). Nesta combinação, ritonavir atua como potencializador farmacocinético, uma vez que é capaz de inibir a atividade da isoenzima 3A4 citocromo oxidase P450 (CYP3A4), otimizando a meia-vida de nirmatrelvir no organismo (LEMAITRE et al., 2022).

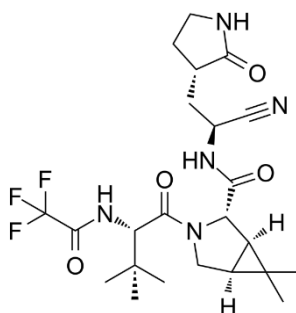


Figura 6. Estrutura química do nirmatrelvir. Fármaco aprovado como inibidor da M^{pro} de SARS-CoV-2, administrado em formulação com ritonavir, sob nome comercial Paxlovid™. Imagem gerada com ChemDraw.

Contudo, tal combinação também pode ocasionar dois cenários de preocupação no tratamento da COVID-19: (i) interações medicamentosas prejudiciais, em que a inibição de enzimas citocromo por ritonavir pode interferir na metabolização de outros (STADER; BATTEGAY; MARZOLINI, 2021), como anticoagulantes (LANGE et al., 2022), além de estatinas e benzodiazepínicos (GIRARDIN et al., 2022); e (ii) redução da eficácia de Paxlovid™, em que indutores de diferentes enzimas podem acelerar a metabolização de nirmatrelvir (HESKIN et al., 2022), como antimicrobianos para tratamento de tuberculose (p.ex.: rifampicina) (STADER; BATTEGAY; MARZOLINI, 2021) e fármacos anticonvulsivos (p.ex.: fenobarbital) (GIRARDIN et al., 2022). Ademais, ainda que eficazes contra diferentes linhagens de SARS-CoV-2 (VANGEEL et al., 2022), a eficácia limitada de nirmatrelvir no início dos sintomas (ARBEL et al., 2022), bem como as restrições de uso em

crianças e gestantes, por exemplo, em razão de interações medicamentosas do ritonavir (EMA, 2022), ressaltam a necessidade de estudos na obtenção de novos inibidores e/ou aprimoramento daqueles já existentes.

Nesse sentido, o conhecimento adquirido de cenários de dois outros vírus de RNA, o vírus influenza e o vírus da imunodeficiência humana (HIV), também ressaltam a importância da busca contínua de novas moléculas bioativas contra SARS-CoV-2. O oseltamivir (Tamiflu®) (TREANOR et al., 2000), por exemplo, ainda é a opção terapêutica de escolha para influenza, mesmo com novas linhagens virais e atualizações vacinais anuais (PETERS et al., 2001; FRY et al., 2014; JEFFERSON et al., 2014). Igualmente, diversos novos fármacos foram desenvolvidos para HIV, como inibidores de protease, transcriptase reversa (análogos de nucleosídeo e não nucleosídeos) e penetração, sendo implementados ou substituídos ao longo dos anos, em razão da seleção de resistência (mutações) na terapia antirretroviral, além da toxicidade a longo prazo (TSENG; SEET; PHILLIPS, 2015; FORSYTHE et al., 2019).

Ainda em 2022, os primeiros isolados resistentes ao nirmatrelvir *in vitro* (ZHOU et al., 2022; JOCHMANS et al., 2023) e *in vivo* (ABDELNABI et al., 2023) também já haviam sido demonstrados superando em até 100 vezes os valores de EC_{50} , principalmente associados a mutações não-sinônimas no gene da M^{pro} , levando à substituições de resíduos como T21I, L50F, E166A e L167F (ZHOU et al., 2022; JOCHMANS et al., 2023), e sendo sugeridos em seres humanos (ANDERSON; CAUBEL; RUSNAK, 2022). Essas mudanças também foram evidenciadas em estudos que indicam diferentes vias associadas a possível seleção de resistência (IKETANI et al., 2022), reiterando a importância do planejamento, busca e obtenção de novos inibidores contra SARS-CoV-2.

1.3 Estratégias no planejamento, descoberta e desenvolvimento de moléculas bioativas com atividade antiviral

O planejamento de moléculas bioativas na busca de potenciais fármacos antivirais pode ser aprimorado pela integração de estratégias computacionais e experimentais, otimizando tempo e reduzindo custos associados (SERAFIM et al., 2021a, 2021b). Considerando o cenário da COVID-19 e, frente a possibilidade de novos surtos, epidemias ou pandemias, um inibidor de amplo espectro seria de interesse como opção terapêutica para doenças causadas por outros coronavírus, incluindo possíveis mutações e seleção de resistência aos fármacos disponíveis (DENG et al., 2014; MORENS; FAUCI, 2020; IKETANI et al., 2022).

Alguns fármacos antivirais foram obtidos empregando estratégias computacionais de sucesso, como oseltamivir para tratamento de influenza (TALELE; KHEDKAR; RIGBY, 2010), boceprevir (NJOROGÉ et al., 2008), um inibidor de protease do vírus da hepatite C (HCV), além de lopinavir e ritonavir, inibidores de protease de HIV (WLODAWER, 2002). Nesse sentido, abordagens *in silico* no planejamento de fármacos baseadas na estrutura de um alvo (do inglês, *structure-based drug design*; SBDD) ou baseadas na estrutura de um ligante (do inglês, *ligand-based drug design*; LBDD) (LIMA et al., 2016), podem auxiliar na descoberta de moléculas bioativas e/ou no desenvolvimento de potenciais candidatos a fármacos contra SARS-CoV-2 e outros coronavírus (LIU; ZHENG; WANG, 2020; PANT et al., 2020; WU et al., 2020; SERAFIM et al., 2021a).

As estratégias de SBDD compreendem simulações que utilizam cálculos empíricos ou descritores baseados na estrutura de um alvo molecular, geralmente uma proteína ou DNA/RNA, sendo o acoplamento molecular (do inglês, *molecular docking*) a técnica mais comumente empregada (PINZI; RASTELLI, 2019). As técnicas de *docking* empregam diferentes algoritmos para predição de um modo de interação (pose) entre um determinado ligante (molécula potencialmente bioativa) e sua afinidade com um alvo molecular de estudo (proteínas) pontuado por uma função *score* (TORRES et al., 2019).

A segunda estratégia, LBDD, se baseia na relação entre a estrutura química de moléculas bioativas e uma determinada atividade biológica, como a inibição de uma enzima. Podem ser empregadas técnicas que comparam a estrutura de uma molécula de interesse e uma ou mais estruturas com propriedades biológicas conhecidas, por exemplo: (i) similaridade molecular, ou seja, um método para calcular e quantificar como duas ou mais moléculas são similares (MAGGIORA et al., 2014); (ii) modelos de relações quantitativas entre a estrutura química e atividade biológica (do inglês, *Quantitative Structure-Activity Relationship*; QSAR), que busca prever um valor numérico de atividade de uma molécula ou classificar substâncias como ativas ou inativas (GRAMATICA, 2020); e (iii) modelos farmacofóricos, que consistem em determinar funções ou propriedades específicas relacionadas a uma atividade biológica, comparando-as com uma molécula de interesse (LU et al., 2018).

Ambas as estratégias podem ser empregadas e combinadas dentre duas ou mais técnicas (p.ex.: *docking* e QSAR), potencialmente aumentando a confiabilidade dos métodos e predições empregadas (VÁZQUEZ et al., 2020). Ademais, simulações de dinâmica molecular (DM) e técnicas de aprendizado de máquina (do inglês, *machine learning*; ML), podem ser

empregadas junto de testes *in vitro* e triagens de substâncias sintéticas, semi-sintéticas e de produtos naturais, além do reposicionamento de fármacos, para obtenção de novos potenciais inibidores contra alvos moleculares de SARS-CoV-2 (WU et al., 2020; RUAN et al., 2021; FERREIRA et al., 2021).

1.3.1 Inibidores enzimáticos contra SARS-CoV-2

Três enzimas de SARS-CoV-2 apresentam perfis promissores no desenvolvimento de candidatos a fármacos: M^{pro}, PL^{pro} e RdRp (SERAFIM et al., 2021a). Uma vez que ambas não se encontram em células do hospedeiro, também são alvos de interesse no planejamento de inibidores seletivos. Ambas são essenciais para a multiplicação viral e conservadas dentre os coronavírus, haja visto o compartilhamento de cerca de 90 a 96% de identidade do genoma completo entre SARS-CoV-2, SARS-CoV e coronavírus de morcego (RAMAJAYAM; TAN; LIANG, 2011; ZHANG et al., 2020a), sendo um potencial alvo multi-viral ou pancoronavírus (PILLAIYAR et al., 2016; CHEN, 2020; DONG et al., 2020; IONESCU, 2020).

Diversas estruturas cristalográficas já foram resolvidas, a exemplo de monômeros e dímeros da M^{pro}, demonstrando que os principais aminoácidos envolvidos nos processos de catálise, interação com ligantes e dimerização da enzima são conservados, bem como a preferência por substratos similares (DONG et al., 2020; IONESCU, 2020). Nesta enzima, por exemplo, a clivagem ocorre entre uma glutamina (Gln) no sítio P1 e uma glicina (Gly), alanina (Ala) e serina (Ser) no sítio P1', sendo a presença de Gln essencial neste processo e oferecendo uma seletividade em potencial, uma vez que cisteína proteases humanas não apresentam essa preferência (OWEN et al., 2021; RUT et al., 2021).

O mecanismo de ação envolve duas etapas principais: (i) acilação, em que ocorre a clivagem do fragmento P1' do substrato peptídico, formando um complexo de enzima em que a cisteína (Cys145) da díade catalítica se liga de forma covalente; (ii) desacilação, em que o complexo de enzima formado é hidrolisado, liberando o fragmento P1 e recuperando o sítio enzimático para nova atividade da protease (**Figura 7**) (RAMOS-GUZMÁN; RUIZ-PERNI; TUÑÓN, 2020; CITARELLA et al., 2021).

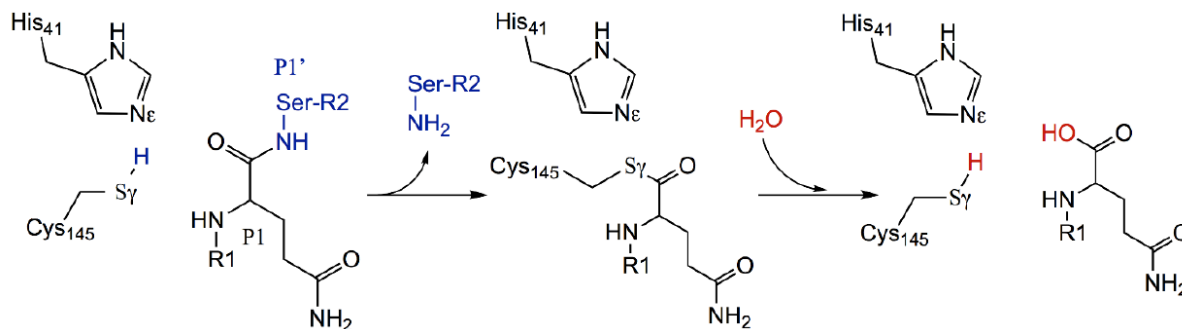


Figura 7. Etapas do mecanismo de ação da protease principal (M^{pro}) de SARS-CoV-2. Inicialmente, além de uma transferência de próton entre Cys145 e His41 na díade catalítica, ocorre a clivagem no sítio P1', gerando um complexo de enzima por interação covalente entre Cys145 e o fragmento P1. Em seguida ocorre hidrólise do complexo, liberando o fragmento e recuperando o sítio ativo da protease. Fonte: (RAMOS-GUZMÁN; RUIZ-PERNI; TUÑÓN, 2020) (*adaptado*).

Nesse sentido, inibidores covalentes reversíveis e/ou irreversíveis poderiam formar complexos impedindo a hidrólise de substratos e a consequente atividade da M^{pro} (JIN et al., 2020a, 2020b) e da PL^{pro} , outra cisteína protease (OSIPIUK et al., 2021). Estudos já mostraram a viabilidade do emprego de estratégias computacionais como *docking* e simulações de DM para triagens virtuais de moléculas contra essas enzimas em SARS-CoV-2 (JIN et al., 2020a; OSIPIUK et al., 2021; RUAN et al., 2021), e casos de sucesso dos inibidores da RdRp, molnupiravir (PAINTER et al., 2021), e da M^{pro} , nirmatrelvir (OWEN et al., 2021).

Inibidores enzimáticos peptídicos, por exemplo, se baseiam nos substratos enzimáticos e no sítio de clivagem dos alvos moleculares, usualmente apresentando variabilidade estrutural com facilidade de síntese, além de um mecanismo de ação seletivo a um único alvo ou alvos similares (PANT et al., 2020; ZHANG et al., 2020a). Contudo, fragmentos polares na estrutura desses ligantes usualmente resultam em baixa permeabilidade em membranas celulares e consequente biodisponibilidade limitada, evidenciando uma baixa eficácia em ensaios *in vitro* (LI et al., 2005).

Já os compostos peptideomiméticos, como o próprio nirmatrelvir (PILLAIYAR et al., 2022), são híbridos entre fragmentos peptídicos e, por exemplo, porções eletrofílicas, sendo uma alternativa de interesse, uma vez que podem mimetizar as funções peptídicas enquanto apresentam características farmacocinéticas mais favoráveis, como absorção (TIMIRI; SINHA; JAYAPRAKASH, 2016). Essa abordagem já demonstrou, por exemplo, resultados mais favoráveis de inibição da M^{pro} de SARS-CoV-2 por análogos peptídicos quando

combinada com estratégias computacionais (LIANG et al., 2020a; ZHANG et al., 2020a). Essas e outras moléculas bioativas, envolvendo ou não fragmentos ou uma estrutura peptídica, a exemplo de produtos naturais (NGO et al., 2020), podem ser identificadas por diversas estratégias, como triagens experimentais (*in vitro*) de moléculas em larga escala (ZANG et al., 2023) (não abordadas aqui) e triagens virtuais (*in silico*) de quimiotecas (PANT et al., 2020).

1.3.2 Triagem virtual de moléculas bioativas

Uma triagem virtual (do inglês, *virtual screening*; VS) é uma abordagem *in silico* que visa filtrar grandes conjuntos ou bancos de dados de moléculas, a fim de identificar ligantes com potencial atividade biológica. Baseadas em abordagens de LBDD e/ou SBDD, uma triagem virtual é capaz de avaliar até centenas de milhões de moléculas, acelerando uma etapa de desenvolvimento inicial (GUIDO; ANDRICOPULO; OLIVA, 2010). Uma abordagem usualmente empregada nas triagens virtuais é a caracterização e limitação do espaço químico utilizado, ou seja, a definição de parâmetros e filtros de propriedades físico-químicas e/ou estruturais que possibilitariam a seleção de melhores candidatos a fármacos, a exemplo do caráter “*druglike*” ou fármaco-similar, como em fármacos via oral com características bem definidas (LIPINSKI; HOPKINS, 2004), filtrando e selecionando moléculas a serem avaliados posteriormente (LIPINSKI, 2004). Assim, se permite analisar e solucionar possíveis problemas que normalmente resultariam em baixa afinidade e/ou atividade de um ligante e/ou inibidor (WARREN et al., 2006), por exemplo, em análises de acoplamento molecular.

Dentre esses parâmetros, a “Regra do 5 de Lipinski” (do inglês *Rule of Five*; Ro5), por exemplo, permite estimar requisitos de uma molécula para apresentar maior disponibilidade após administração oral, sendo eles: (i) massa molecular menor que 500 Da; (ii) não mais que cinco grupos doadores e (iii) 10 grupos aceptores de ligação de hidrogênio; além de (iv) log P (logaritmo do coeficiente de partição *n*-octanol/água) menor que cinco; sendo permitido violar uma das quatro condições (LIPINSKI et al., 1997). Outro exemplo, as “regras de Veber”, sugerem que a polaridade e o grau de liberdade conformacional de uma molécula são importantes para a disponibilidade oral, em que o fármaco deve apresentar uma superfície polar menor do que 140 Å² e não mais que 10 ligações rotacionáveis (VEBER et al., 2002).

Nesse sentido, para alvos como a M^{pro} de SARS-CoV-2, triagens de bibliotecas de moléculas capazes de identificar potenciais inibidores antivirais sintéticos, semi-sintéticos, produtos naturais e fármacos de reposicionamento (HAGE-MELIM et al., 2020; PANT et al., 2020; SASIDHARAN et al., 2020) podem ser otimizadas. Adicionalmente, estudos

empregando *docking* também auxiliam e já demonstraram sucesso na obtenção de inibidores contra enzimas de SARS-CoV-2 (KUMAR et al., 2020; KUMAR; SINGH; PATEL, 2020), a exemplo do lopinavir e ritonavir *in vitro* (MAHANTA et al., 2020).

Contudo, análises de *docking* apresentam acurácia discutível entre diferentes programas e também requerem uma inspeção visual de poses e conformações para auxiliar na tomada de decisões sobre interações de potenciais inibidores (FISCHER et al., 2021). Nesse sentido, a combinação dessa estratégia com simulações de DM, por exemplo, pode refinar os dados obtidos em uma triagem virtual, sendo possível estimar a frequência de interações de uma molécula com determinados resíduos de aminoácidos, além de sua permanência em um sítio da estrutura alvo, ponderando sobre a estabilidade do ligante no complexo ao longo do tempo de simulação (AMIN; SOROOR; KASRY, 2020; LIANG et al., 2020b; PANT et al., 2020; SASIDHARAN et al., 2020).

1.3.3 Simulações de dinâmica molecular (DM)

As simulações de DM consistem em um sistema de simulações que permite modelar os movimentos de moléculas e alvos moleculares a nível atômico, predizendo alguns fenômenos difíceis de se observar experimentalmente. Diferentemente do *docking*, as simulações de DM são capazes de calcular e prever as variações conformacionais de complexo alvo-ligante em função do tempo, em escalas de nano- a milissegundos (DUAN; KOLLMAN, 1998; BOWERS et al., 2006; DURRANT; MCCAMMON, 2011). Esse tipo de simulação que mimetiza uma realidade biológica pode envolver centenas de milhares de átomos, representativos de alvos moleculares e moléculas em um ambiente que considera solvente (p.ex.: água), além de íons e outras estruturas complementares, como cofatores (DUAN; KOLLMAN, 1998; BOWERS et al., 2006).

Essa estratégia obteve sucesso na busca de sítios de ligação adjacentes ao sítio ativo da integrase de HIV-1, auxiliando no desenvolvimento de inibidores desta enzima (SCHAMES 2004; DURRANT, 2011). Simulações de DM também foram capazes de prever uma cavidade de interesse no sítio de adesão da neuraminidase de vírus influenza, auxiliando, por exemplo, no desenvolvimento de fármacos contra este vírus (AMARO, 2011). Considerando SARS-CoV-2, séries de produtos naturais (SASIDHARAN et al., 2020), análogos sintéticos (LIANG et al., 2020a; FERREIRA et al., 2021) e reposicionamento de fármacos (KUMAR; SINGH; PATEL, 2020) foram avaliados como potenciais inibidores da M^{pro}.

Essa técnica também permitiu observar, por exemplo, inconsistências na flexibilidade da estrutura monomérica da M^{pro}, e estabilidade do dímero ao longo de trajetórias simuladas, identificando alguns resíduos mais frequentes de interação: His41, Gly143, His163, Glu166 e Gln189 (FERREIRA et al., 2021). Finalmente, algumas técnicas complementares também podem auxiliar este e outros métodos computacionais empregados em diferentes etapas do planejamento de moléculas bioativas, como a predição de sítios de ligação para interação de ligantes, bem como de propriedades farmacocinéticas, farmacodinâmicas e de toxicidade destes potenciais inibidores, discutidos a seguir.

1.3.4 Técnicas complementares: predições de sítios de ligação em alvos moleculares

Algumas técnicas são utilizadas para estudar os possíveis sítios de ligação de proteínas de interesse no planejamento de fármacos e para identificar regiões e resíduos favoráveis às interações moleculares de ligantes. A capacidade destas regiões interagirem, por exemplo, a uma variedade de pequenas moléculas orgânicas (moléculas-sonda) também pode ser predita por servidores como FTSite (NGAN et al., 2012) e FTMap (KOZAKOV et al., 2015). Técnicas de ML, como *random forest*, também podem auxiliar na predição de sítios, a exemplo do servidor PrankWeb (JENDELE et al., 2019).

Adicionalmente, o grau de conservação de um resíduo de aminoácido também pode ser avaliado por alinhamento de sequências múltiplas, como empregado no servidor ConSurf (ASHKENAZY et al., 2016). Esses métodos podem ser considerados análogos computacionais de determinações experimentais de macromoléculas (cristalografia de raios-X e RMN) que identificam e caracterizam sítios de ligação e quais suas possíveis interações, mapeando alvos moleculares, por exemplo, na identificação de inibidores de M^{pro} (SILVA et al., 2021; VERMA et al., 2023), juntamente de simulações de DM (FERREIRA et al., 2021), e até mesmo para outras proteínas alvo de SARS-CoV-2 (ABDELKADER et al., 2022).

1.3.5 Técnicas complementares: predições de propriedades farmacocinéticas, farmacodinâmicas e de toxicidade de ligantes

Os estudos *in silico*, além de buscar potenciais moléculas bioativas, também são empregados no intuito de filtrar ligantes com potencial toxicidade e perfil farmacocinético inadequado em etapas mais avançadas do desenvolvimento de fármacos (PEREIRA, 2007; RUIZ-GARCIA et al., 2008). Alguns métodos, por exemplo, envolvem ML, correlacionando a estrutura química, propriedades físico-químicas e moleculares com um parâmetro estudado,

como a atividade biológica, toxicidade ou perfil farmacocinético de uma molécula bioativa (YANG et al., 2019; PANTALEÃO et al., 2022).

Dentre os programas utilizados, podem ser citados os servidores admetSAR 2.0, que permite uma predição qualitativa de parâmetros farmacocinéticos como a biodisponibilidade oral, absorção intestinal e capacidade de atravessar a barreira hematoencefálica, usualmente desejáveis para potenciais fármacos (YANG et al., 2019). O ProTox-II, por sua vez, faz predições de hepatotoxicidade, mutagenicidade, imunotoxicidade e de outros efeitos adversos, como o valor teórico da dose letal mediana (DL_{50}) (BANERJEE et al., 2018). Já Pred-hERG 4.2 avalia o potencial de um ligante em interagir ou inibir canais para potássio hERG (do inglês, *human ether-a-go-go-related gene*) em células cardíacas (BRAGA et al., 2015).

Análises de absorção, distribuição, metabolismo, excreção e toxicidade (do inglês, *Absorption, Distribution, Metabolism, Excretion and Toxicity*; ADMET), também já foram empregados para selecionar inibidores da M^{pro} de SARS-CoV-2 (HAGE-MELIM et al., 2020). No geral, para todas essas análises, se sugere que mais de um algoritmo, programa ou servidor sejam empregados para aumentar a confiabilidade das predições (PANTALEÃO et al., 2022). Tomadas em conjunto, essas abordagens podem direcionar moléculas bioativas planejadas *in silico* para avaliações biológicas *in vitro*, auxiliando no planejamento, descoberta e desenvolvimento de inibidores e novos potenciais candidatos a fármacos contra enzimas de SARS-CoV-2.

2 JUSTIFICATIVA

Aproximadamente 664,8 milhões de infecções pelo coronavírus relacionado à síndrome respiratória aguda grave 2 (SARS-CoV-2) já foram notificados no mundo até 25 de janeiro de 2023, tendo a doença do coronavírus 2019 (COVID-19) ocasionado mais de 6,72 milhões de mortes notificadas em três anos. Apesar de um contínuo esforço coletivo empregando medidas sanitárias e de distanciamento social, bem como do uso de máscaras, além de um esquema vacinal de múltiplas doses implementado em massa no mundo, novas linhagens de SARS-CoV-2 e surtos de COVID-19 ainda persistem como um problema saúde pública global.

Ademais, ainda que a aprovação emergencial de fármacos antivirais (p.ex.: nirmatrelvir em combinação com ritonavir) tenha auxiliado no tratamento da COVID-19, limitações de uso por interações medicamentosas de importância, além da seleção de possíveis mutações denotam a importância da continuidade de esforços para obtenção de novas opções terapêuticas. Igualmente, considerando a possibilidade de outros coronavírus emergentes ou reemergentes, também se ressalta a necessidade de um maior número de antivirais disponíveis, frente a potenciais cenários de novos surtos, epidemias e pandemias.

Nesse sentido, se evidencia a necessidade de estratégias para planejar e avaliar potenciais moléculas bioativas com atividade contra SARS-CoV-2, a fim de se obter inibidores de alvos moleculares no planejamento de candidatos a fármacos para tratamento da COVID-19. As proteases (M^{pro} e PL^{pro}) se apresentam como alvos de interesse para o desenvolvimento de novos inibidores, em razão de sua atividade enzimática indispensável à multiplicação de SARS-CoV-2 e alta conservação estrutural dentre diferentes espécies da família *Coronaviridae*, assim um potencial pancoronavírus.

Nesse âmbito, abordagens *in silico* também podem auxiliar no desenvolvimento e otimização de estratégias para o planejamento e descoberta de novas moléculas bioativas contra SARS-CoV-2, empregando técnicas computacionais e ferramentas a fim de direcionar análises iniciais *in vitro* e, posteriormente, *in vivo*. Portanto, a presente proposta objetiva, empregando estratégias *in silico* e *in vitro* para o planejamento e triagem de potenciais inibidores, planejar e avaliar potenciais moléculas bioativas contra SARS-CoV-2, estabelecendo como alvos moleculares as duas proteases, M^{pro} e PL^{pro} .

3 OBJETIVOS

3.1 Objetivo Geral

Identificar moléculas bioativas com potencial atividade inibitória contra as proteases (M^{pro} e PL^{pro}) de SARS-CoV-2, planejadas para atuarem contra estes alvos moleculares, empregando estratégias *in silico* e validação de sua atividade *in vitro*.

3.2 Objetivos Específicos

- 3.2.1 Avaliar a capacidade preditiva de um modelo de QSAR por holograma (HQSAR), na predição de moléculas bioativas contra uma das enzimas alvo (M^{pro});
- 3.2.2 Identificar potenciais sítios de ligação preditos nas proteases (M^{pro} e PL^{pro}) de SARS-CoV-2 e outros coronavírus;
- 3.2.3 Avaliar interações dos ligantes co-cristalizados de estruturas cristalográficas com resíduos de aminoácidos empregando acoplamento molecular (*redocking*);
- 3.2.4 Construir modelos farmacofóricos utilizando as informações do mapeamento e interações das estruturas selecionadas;
- 3.2.5 Selecionar potenciais inibidores das proteases de SARS-CoV-2 em triagens virtuais de quimiotecas como BraCoLi, DrugBank, GMR, NuBBE e ZINC15;
- 3.2.6 Inspeccionar *hits* virtuais empregando análises de *docking* e dinâmica molecular;
- 3.2.7 Predizer propriedades farmacocinéticas, farmacodinâmicas e de toxicidade das moléculas selecionadas *in silico*, para desconsiderar aquelas indesejáveis de aquisição e/ou síntese;
- 3.2.8 Avaliar a inibição da atividade enzimática das moléculas selecionadas contra as proteases de SARS-CoV-2 (University of California, San Diego; UCSD);
- 3.2.9 Determinar a concentração citotóxica (CC₅₀) das moléculas selecionadas;
- 3.2.10 Determinar a potencial concentração efetiva (CE₅₀) e os respectivos índices de seletividade (IS) das moléculas selecionadas;

4. MATERIAL E MÉTODOS

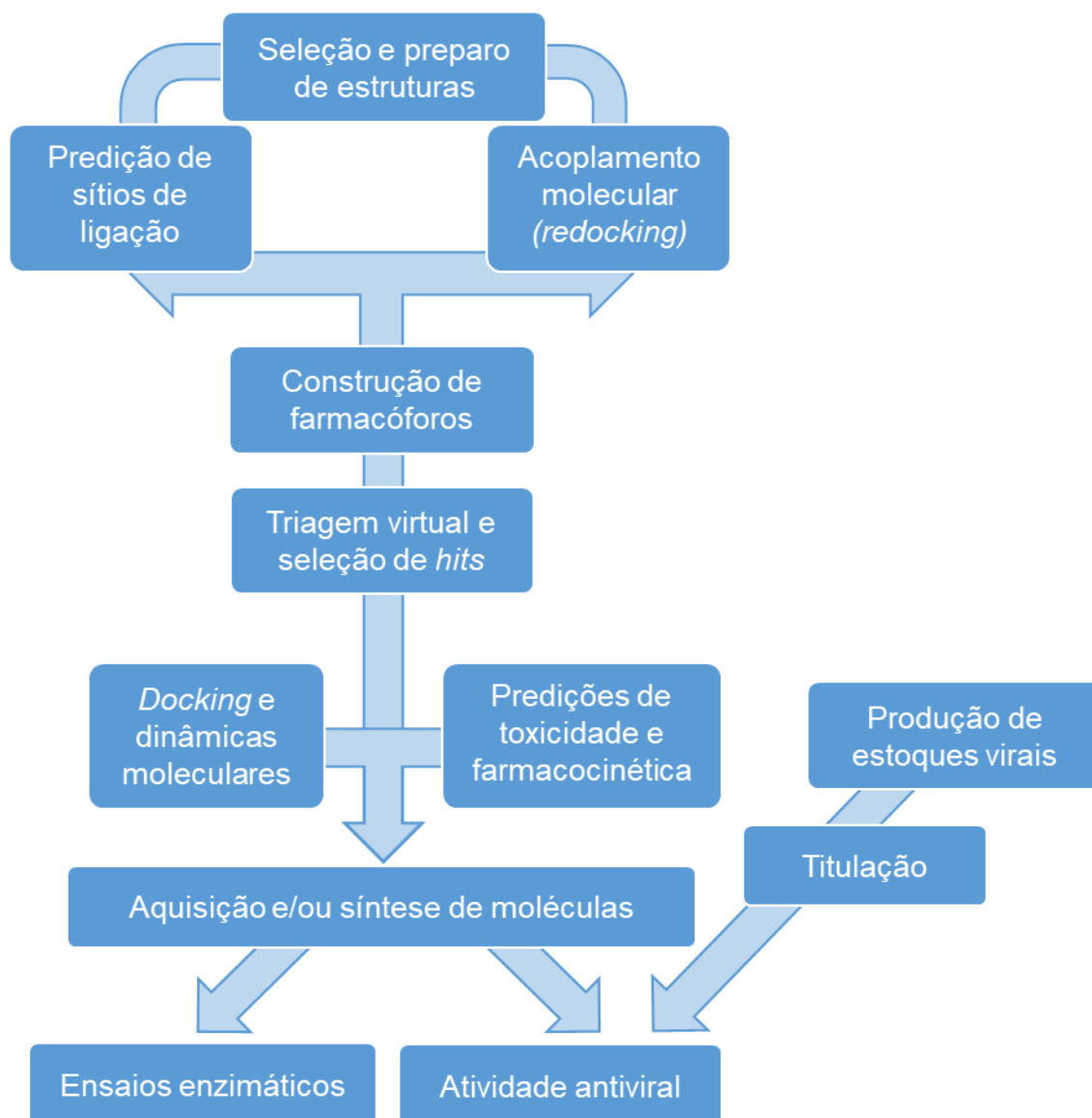


Figura 8. Fluxograma de atividades das etapas executadas no projeto de doutorado. O presente projeto seguiu duas abordagens em paralelo, iniciando com as análises computacionais das estruturas dos alvos moleculares de interesse, avaliando potenciais regiões de interação até a construção de modelos farmacofóricos para triagem virtual de moléculas bioativas. As moléculas selecionadas nesta etapa foram inspecionadas e selecionadas por análises de *docking* e simulações de dinâmica molecular para serem avaliadas em ensaios de inibição da atividade enzimática e, posteriormente, da atividade antiviral (fluxograma abaixo). O referido projeto apresenta cadastro vinculado ao Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SisGen), sob número de registro **AD750A9**.

4.1 Análises *in silico*

4.1.1 Construção dos conjuntos de dados

As estruturas tridimensionais (3D) das moléculas de cada conjunto ou bancos de dados foram construídos no programa BIOVIA Discovery Studio (BIOVIA, USA, 2017). Estas foram então submetidas ao programa OMEGA 4.1.1 (OpenEye Scientific Software, USA, 2021) para cálculo do confôrmero de menor energia de cada molécula (comando “*-maxconfs 1*”), ou seja, a unidade conformacional de maior estabilidade dentre um número máximo de conformações. Posteriormente, todas as estruturas tiveram seus estados de ionização corrigidos de acordo com o pH fisiológico (7,4), utilizando a aplicação fixpka no pacote computacional QUACPAC 2.1.2 (OpenEye Scientific Software, USA, 2021), assim determinando um único estado de ionização mais provável de cada molécula de acordo com o pH em ensaios experimentais subsequentes.

4.1.2 Relação quantitativa entre a estrutura química e a atividade biológica por holograma (HQSAR) e predição de atividade biológica

O HQSAR também foi inicialmente empregado a fim de obter informações sobre a contribuição de fragmentos moleculares de inibidores enzimáticos existentes na literatura. O método foi embasado na combinação de descritores moleculares, átomos (A), ligações (B), conectividade (C), átomos de hidrogênio (H), quiralidade (Q) e átomos doadores ou aceptores (DA), para geração dos modelos utilizando a estrutura bidimensional (2D) de moléculas. O tamanho do fragmento foi fixado em relação ao número de átomos (4 a 7), além de um tamanho mínimo e máximo de extensão dos hologramas (50 a 500 *bins*), organizados em números primos. O modelo com maior valor de q^2 obtido nessa fase de otimização foi submetido a uma variação de tamanho dos fragmentos, em intervalos de quatro átomos (1 a 4 até 7 a 10), mantendo os descritores do modelo.

Foram selecionados seis modelos, obtendo uma otimização dos resultados na primeira etapa, buscando obter um modelo ótimo estatisticamente mais representativo, que correlacionasse o holograma gerado com a atividade biológica (inibição enzimática) das moléculas. Assim, mapas de contribuições (positivos, neutros ou negativos) dos fragmentos moleculares moléculas utilizadas no modelo foram gerados, indicando regiões potencialmente fundamentais para atividade biológica (CHAI; DRAXLER, 2014; CHIRICO; GRAMATICA, 2011; GRAMATICA; SANGION, 2016), neste caso, a inibição da M^{PRO} de SARS-CoV-2.

Finalmente, o cálculo das métricas da validação externa do modelo foi feito no pacote computacional KNIME® (BERTHOLD et al., 2007). Adicionalmente, uma predição da atividade biológica das moléculas foi feita empregando um modelo de *random forest* sobre cálculos de diferentes descritores moleculares (VERÍSSIMO et al., 2019) no software PaDEL (YAP, 2011). Dessa forma, descritores moleculares 2D, *fingerprints* e propriedades físico-químicas das estruturas do HQSAR foram avaliados para prever demais moléculas com potencial atividade biológica em outros bancos de dados.

4.1.3 Seleção e preparo das estruturas das proteases de SARS-CoV-2

Uma busca no banco de dados de proteínas *Protein Data Bank* (PDB) (BERMAN et al., 2000), foi feita utilizando os seguintes termos como palavras-chave: “SARS-CoV-2”, “nsp3”/“PLpro”, “nsp5”/“Mpro”, “*protease*”. Quatro critérios para seleção das estruturas mais apropriadas para o estudo foram estabelecidos, sendo: (i) valor da resolução < 2.0 Å, que indica estruturas de com maior qualidade de resolução, possibilitando maior confiabilidade da posição dos átomos (AZEVEDO et al., 2022); (ii) presença de inibidores co-cristalizados, a fim de compará-los com ligantes planejados, bem como de obter um estado conformacional apropriado para as interações; (iii) ausência de mutações, buscando sequências mais similares e/ou fidedignas àquelas dos modelos determinados experimentalmente; e (iv) cadeias completas, dispensando uma construção por homologia. O programa PyMOL v0.99c (DeLano Scientific, CA, USA, 2006) foi usado para análises dos resultados e elaboração das imagens.

4.1.4 Predição dos sítios de ligação das enzimas

As estruturas selecionadas da M^{pro} e PL^{pro}, no formato PDB, foram submetidas aos servidores FTSite e FTMap (KOZAKOV et al., 2015), buscando prever cavidades ou regiões susceptíveis a atuarem como sítios de ligação, bem como quais os tipos de interações seriam propícios a ocorrerem, respectivamente. O servidor FTSite resultou nos sítios de ligação e resíduos de aminoácidos das enzimas, enquanto o servidor FTMap resultou nos grupamentos de moléculas-sonda divididos por sítios de ligação, as quais interagem com os resíduos das enzimas, mapeando os sítios nas proteases. Assim, foram estimados os aminoácidos que fizeram mais interações com alguma molécula-sonda próxima, bem como quais sondas poderiam ser mais relevantes estando presentes em um possível inibidor.

Adicionalmente, o servidor PrankWeb também foi empregado para predições de potenciais sítios de ligação, análise que é baseada no modelo P2Rank, um conjunto de preditores utilizando ML capaz de mapear potenciais resíduos de interação em alvos moleculares de interesse (JENDELE et al., 2019). Potenciais pontos de interação de ligantes na superfície das proteases foram preditos e as propriedades físico-químicas e geométricas dos resíduos circundantes calculadas e agrupadas, resultando em uma pontuação e classificação individual para cada agrupamento (*cluster*), ranqueando sítios de interesse.

4.1.5 Acoplamento molecular (*docking*) de ligantes e potenciais moléculas bioativas

Os estudos de interação entre os ligantes selecionados e os alvos moleculares foram feitos utilizando o método de acoplamento molecular, empregando também a reprodução do *docking* do ligante co-cristalizado, o chamado *redocking*. Este é um importante passo para avaliar a confiabilidade do algoritmo genético (do inglês, *genetic algorithm*; GA) de um programa de *docking* na predição de poses experimentais. O sítio de ligação das proteases foi definido ao redor dos respectivos ligantes co-cristalizados, utilizando os programas Surflex-Dock (JAIN, 2003) e GOLD 5.1 (VERDONK et al., 2003), que possuem algoritmos distintos. Os parâmetros intrínsecos dos programas, como funções de pontuação, flexibilidade de resíduos e átomos, além do número de poses foram avaliados previamente buscando reproduzir um modo de ligação experimental mais adequado dos ligantes co-cristalizados (VALLONE et al., 2018) e, posteriormente, dos ligantes selecionados nas triagens.

Ambas as proteases foram preparadas nos dois programas adicionando átomos de hidrogênio, removendo moléculas de água não estruturais, extraindo ligantes co-cristalizados e outras subestruturas (p.ex.: cofatores) e calculando cargas atômicas parciais em cada programa (*default*). Os parâmetros utilizados no Surflex-Dock foram: *bloat* (0.5); *threshold* (0.5); tamanho (0.0); movimento de hidrogênios/átomos de hidrogênio (não); poses otimizadas (não); conformações iniciais adicionais (8); *grid* de busca (6 Å); máximo de conformações por fragmento (60); máximo de ligações rotacionáveis (100); densidade de busca (6); número de *spins* (24); máximo de poses (20); e RMSD mínimo entre poses (0.5). No GOLD, por sua vez, as opções e parâmetros foram: (i) todas as ligações rotacionáveis foram fixadas; (ii) interação definida com o ligante co-cristalizado e todos os átomos dentro 6 Å; (iii) *chemscore_kinase* como *template*; (iv) 200 *GA runs* para *redocking* e 20 para *docking*; (v) CHEMPLP como função *score*, sem encerramento precoce (< 0.5 Å entre poses); e (vi) opção de busca definida como lenta.

O *redocking* do ligante co-cristalizado, por sua vez, forneceu dois resultados de importância: um modo de ligação de referência das interações intermoleculares, necessárias para inibir as enzimas, considerando valores de $\text{RMSD} \leq 2.0 \text{ \AA}$, a fim de se comparar tais interações com ligantes avaliados, além de uma pontuação referente à afinidade do ligante com a enzima estudada (ranqueamento das moléculas). Assim, informações relevantes de cada método foram utilizadas na construção dos modelos farmacofóricos em etapas posteriores, enquanto os resultados do *docking* também foram empregados para a análise dos *hits* obtidos nas etapas das triagens virtuais subsequentes.

4.1.6 Construção de modelos farmacofóricos e triagem virtual de bancos de dados

Utilizando as informações da predição de sítios das proteases e das análises de *docking*, modelos farmacofóricos foram construídos no programa UNITY implementado na plataforma computacional SYBYL-X 2.1. Os resultados obtidos, juntamente de simulações de dinâmicas moleculares, possibilitaram corroborar a seleção de resíduos de interesse, definindo os centroides de interação molecular em cada modelo farmacofórico, permitindo filtrar moléculas de interesse como potenciais inibidores das enzimas (ARUN et al., 2020). Tais centroides foram definidos com 1,0 \AA de tolerância.

As quimiotecas empregadas na triagem virtual foram: (i) Brazilian Compound Library (BraCoLi) (VERÍSSIMO et al., 2022), que contempla 1.176 substâncias sintéticas catalogadas; (ii) DrugBank, banco contendo mais de 7.800 fármacos aprovados pelo FDA ou em fase experimental e disponíveis comercialmente; (iii) NuBBE, Núcleo de Bioensaios, Biossíntese e Ecofisiologia da UNESP (PILON et al., 2017), contendo 2.147 produtos naturais e derivados; (iv) GMR (“Gerwick, McKerrow and Roush”), uma quimioteca de isolados de produtos naturais e moléculas sintéticas (potenciais inibidores de catepsinas) da Universidade da Califórnia (ANG et al., 2011), contendo, até julho de 2022, 1536 moléculas; e (v) um subconjunto do banco de dados ZINC15 (<https://zinc.docking.org/>), com mais de 6,8 milhões de moléculas utilizadas no presente trabalho, selecionadas baseado na disponibilidade de aquisição via MolPort. O número total de candidatos corresponde às moléculas disponíveis em cada quimioteca até a execução deste protocolo de triagem virtual.

A fim de se obter moléculas com maior biodisponibilidade na etapa de triagem inicial, um filtro foi estabelecido para o conjunto de dados ZINC15, baseado em modificações das regras dos cinco de Lipinski (LIPINSKI, 1997) e nas “regras de Veber” (VEBER, 2002).

Alguns critérios específicos foram modificados (centros quirais, carga total, solubilidade e elementos permitidos) em razão dos métodos aplicados no presente trabalho (**Tabela 2**).

Tabela 2. Parâmetros definidos para os modelos de triagem virtual de ZINC15 (Molport)

Parâmetros	Condições
Elementos permitidos	C, O, N, H, P, S, F, Cl, Br, I
Doador de hidrogênio (HBD)	0 a 5
Aceptor de hidrogênio (HBA)	0 a 10
Log P	-2,0 a 5,5
Massa molecular (MM)	150 a 550
Carga total	-2 a +2
Centro quiral	≤ 2
Ligações rotacionáveis	≤ 10
Área de superfície polar (PSA)	0 a 300
Solubilidade	Baixa (limitado a DMSO) até completamente solúvel

4.1.7 Simulações de dinâmica molecular (DM)

As estruturas das proteases selecionadas foram preparadas para as simulações de DM ajustando os estados de ionização de resíduos e corrigindo quaisquer átomos ausentes nas cadeias laterais (PrepWiz, Maestro v2019.1). Os segmentos ausentes das enzimas foram modelados usando Prime (VOEGLER SMITH; HALL, 2001), removendo sequências incompletas e íons sulfato, e moléculas de água com três ou mais ligações de hidrogênio foram retidas. Posteriormente, o sistema teve seu estado de protonação ajustado usando o programa Epik (SHELLEY et al., 2007), minimizando a energia (JACOBSON et al., 2004).

As simulações foram feitas usando o programa Desmond (BOWERS et al., 2006) com campos de força OPLS4, aprimorando a acurácia de modelos incluindo íons e frações contendo enxofre (LU et al., 2021). Os sistemas simulados, critérios geométricos, escalas de tempo e execuções de simulação foram empregados conforme descrito por Kronenberger et al., (2021) (KRONENBERGER et al., 2021) e Maltarollo et al., (2022) (MALTAROLLO et al., 2022). As simulações foram avaliadas em cinco replicatas de 200 ns (5 x 200 ns), totalizando 1 μ s, em dois momentos: (i) nos estágios iniciais com os ligantes co-cristalizados, para avaliação do perfil de interações entre cada inibidor com as proteases alvo (ROCHA et al., 2021); e (ii) na seleção e avaliação dos *hits* virtuais de interesse.

NOTA: as simulações e análises nesta etapa foram feitas em colaboração com o Dr. Thales Kronenberger, da Universidade de Tübingen, Alemanha.

4.1.8 Predições de propriedades farmacocinéticas, farmacodinâmicas e de toxicidade

As propriedades das moléculas referentes a algumas características farmacocinéticas (p.ex.: absorção, distribuição, metabolismo e eliminação), farmacodinâmicas (p.ex.: interação com receptores) e toxicológicas (p.ex.: hepatotoxicidade) foram preditas empregando os servidores admetSAR 2.0 (YANG et al., 2019) (<http://lmm.d.ecust.edu.cn/admetsar2/>), ProTox-II (BANERJEE et al., 2018) (http://tox.charite.de/protox_II/), e Pred-hERG 4.2 (BRAGA et al., 2015) (<http://predherg.labmol.com.br/>). As estruturas das moléculas foram submetidas aos servidores no formato SMILES (do inglês, *Simplified Molecular Input Line Entry Specification*), retornando os dados de predição baseados em conjuntos de moléculas com uma ou mais características associadas às propriedades farmacocinéticas, farmacodinâmicas e/ou de toxicidade.

4.2 Ensaios *in vitro*

4.2.1 Ensaios de inibição da atividade enzimática da M^{pro} de SARS-CoV-2

M^{pro} recombinante de SARS-CoV-2 foi expressa em *Escherichia coli* e purificada em uma coluna de afinidade contendo níquel, conforme previamente descrito por Mellot et al., (2021) (MELLOTT et al., 2021). A atividade inibitória das moléculas foi avaliada conforme descrito previamente por Ashhurst et al., (2020) (ASHHURST et al., 2021). Brevemente, 10 µM de cada molécula foram incubados com 100 nM da enzima por 15 minutos em microplacas (fundo preto) de 384 poços, em tampão 50 mM de HEPES (pH 7.5), 150 mM de NaCl, 1 mM de EDTA e 0,01% de Tween 20. Subsequente, 20 µM do substrato fluorogênico, Ac-Abu-Tle-Leu-Gln-MCA (Biosynth, FA178674), diluído em tampão foi adicionado, e o monitoramento da fluorescência foi conduzido por até 2h a 37 °C em aparelho Synergy HTX (Biotek, EUA), com faixa de excitação de 360 nm e emissão de 460 nm. As concentrações finais de enzima e substrato foram de 50 nM e 10 µM, respectivamente, e até 0,01% de DMSO. A atividade enzimática foi normalizada em comparação ao controle de DMSO (poços ausentes de moléculas ou inibidores) e os percentuais de inibição (%) calculados. Nirmatrelvir (100 nM) foi utilizado como controle positivo de inibição.

Ensaios para determinação das curvas de dose-resposta foram feitos para as moléculas que inibissem ao menos 50% da atividade enzimática na triagem inicial. Pelo menos nove concentrações foram avaliadas (0,01 até 40 µM). Os valores da concentração inibitória de 50% (IC₅₀) foram calculados com regressão não linear ($r^2 > 0.9$) e o desvio-padrão de cada ponto foi calculado, sendo menor que 10%. Os resultados foram analisados no programa

GraphPad Prism 9.0 (GraphPad Software, San Diego, California, EUA). Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

NOTA: Os ensaios enzimáticos foram conduzidos na University of California San Diego (UCSD), sob supervisão dos Dr. Anthony J. O'Donoghue e Conor R. Caffrey.

4.2.2 Ensaios de inibição da atividade enzimática da M^{pro} de SARS-CoV e MERS-CoV

A atividade inibitória das moléculas foi avaliada contra M^{pro} recombinantes de SARS-CoV e MERS-CoV (R&D Systems, E720 e E719, respectivamente), sob as mesmas condições descritas no **item 4.2.1** para as duas enzimas (PILLAIYAR et al., 2022). As concentrações finais de enzima e substrato foram de 50 nM e 10 μ M, respectivamente, e até 0,02% de DMSO. Pelo menos oito concentrações foram avaliadas (0,02 até 10 μ M) para moléculas que apresentaram ao menos 50% de inibição na triagem inicial (10 μ M). Os valores de IC₅₀ foram calculados da mesma maneira e os dados analisados com GraphPad Prism 9.0 (GraphPad Software, San Diego, California, EUA). Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

4.2.3 Ensaios de inibição da atividade enzimática da PL^{pro} de SARS-CoV-2

Na avaliação da atividade inibitória das moléculas contra a PL^{pro} recombinante de SARS-CoV-2 (Acro Biosystems, PAE-C5184), 100 nM da enzima foram pré-incubadas por 15 minutos com 10 μ M de cada molécula em microplacas (fundo preto) de 384 poços, em tampão 50 mM de HEPES (pH 6.5), 150 mM de NaCl, 0,01% de Tween 20, e 0,1 mM de ditioneitol (DTT). A clivagem do substrato fluorogênico (100 μ M) Z-Arg-Leu-Arg-Gly-Gly-AMC (Bachem, 369 I1690) foi monitorada por até 1h 37 °C no aparelho Synergy HTX (Biotek, EUA), com excitação/emissão de 360/460 nm. As concentrações finais de enzima e substrato foram de 50 nM e 50 μ M, respectivamente, e até 0,01% de DMSO. A triagem inicial foi feita com 10 μ M de cada molécula e os percentuais de inibição (%) calculados em comparação ao controle de DMSO (poços ausentes de moléculas ou inibidores). O inibidor de protease GRL-0617 (10 μ M) foi utilizado como controle positivo de inibição. Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

4.2.4 Vírus e linhagens celulares

Um coronavírus que infecta camundongos, classe de risco 2 (CR-2), Mouse hepatitis virus 3 (MHV-3), foi gentilmente cedido pela Viriontech, Biominas, Minas Gerais, Brasil.

Este organismo é considerado um dos modelos para estudos em nível de biossegurança 2 (NB2), capaz de mimetizar as condições debilitantes de SARS-CoV-2 em camundongos, sendo empregado, por exemplo, para ensaios antivirais, de resistência e inativação viral, além de diminuir restrições de biossegurança para o manuseio e reduzir custos que estariam associados em um laboratório nível de biossegurança 3 (NB3) (PEREIRA OLIVEIRA; KROON, 2021).

A produção dos estoques virais, titulação a avaliação de atividade antiviral, bem como ensaios de citotoxicidade, foram feitas em uma linhagem celular de fibroblasto de tecido adiposo murino, L929 (ATCC® CCL-1™). As células foram cultivadas em meio Eagle modificado por Dulbecco (do inglês, *Dulbecco's modified Eagle's medium*; DMEM) (Cultilab, Brasil). O meio foi suplementado com 5% de soro fetal bovino (SFB) (Cultilab, Brasil), 100 UI/mL de penicilina (Cellofarm, Brasil), 100 µg/mL de estreptomicina (Merck, Alemanha) e 0,25 µg/mL de anfotericina B (Cultilab, Brasil). As células permaneceram em cultivo a 37 °C e atmosfera de 5% de CO₂, sendo homogeneizadas em tripsina acrescida de ácido etilenodiaminoacético (EDTA) (Sinapse Biotecnologia, Brasil) e contadas para realização dos ensaios celulares, ambos conduzidos no Laboratório de Vírus da UFMG. Os ensaios com SARS-CoV-2 serão estritamente conduzidos em ambiente com nível de biossegurança 3 (NB-3), na estrutura implementada no Instituto de Ciências Biológicas (ICB) da UFMG. Os procedimentos seguirão estritamente os protocolos de biossegurança recomendados, realizados por operadores vinculados ao laboratório.

4.2.5 Propagação viral

Células L929 foram cultivadas em meio DMEM suplementado com 5% de SFB e incubadas a 37 °C e 5% de CO₂, em frascos de cultura de 75 cm². Logo após 24h de incubação, as monocamadas celulares com aproximadamente 60 a 80% de confluência foram lavadas com salina tamponada com tampão fosfato (do inglês, *phospate-buffered saline*; PBS) para remoção de restos celulares e componentes de soro, sendo infectadas com uma multiplicidade de infecção (MOI) de 0,001 do estoque viral (*pool* semente) MHV-3 em 1,0 mL de meio DMEM sem SFB. Logo após 1h adsorção e incubação em estufa a 37 °C e 5% de CO₂, com homogeneização delicada dos frascos a cada 10 minutos, foram adicionados 15 mL de meio DMEM com 2% de SFB aos frascos e incubados sob as mesmas condições. Frascos foram observados diariamente em microscópio óptico para visualização da formação de efeitos citopáticos. Observados efeitos em aproximadamente 80% das células (cerca de 72h),

o sobrenadante foi removido e centrifugado a 2.016 g por 10 minutos a 4 °C em uma centrífuga RT6000B (Sorvall, Thermo Scientific, EUA), para remoção dos restos celulares. O material foi distribuído em microtubos tipo Eppendorf de 1,5 mL, armazenado a -70°C e, posteriormente, titulado.

4.2.6 Titulação viral

Células L929 foram cultivadas em microplacas de 12 poços contendo $1,5 \times 10^5$ células/poço e incubadas a 37°C e 5% de CO₂ por 24h. Logo após, as células foram infectadas com 100 µL de diluições seriadas do vírus em proporção 1:10 (de 10^{-1} a 10^{-5}) em DMEM com 1% de SFB, seguidas de 1h de adsorção a 37°C e 5% de CO₂, com homogeneização delicada das placas a cada 10 minutos. O meio foi desprezado e foram adicionados 1,5 mL de meio M199 (Cultilab, Brasil) suplementado com antibióticos e 2% de SFB, além de 1,0% de carboximetilcelulose (CMC) (Synth, Brasil). As microplacas foram incubadas por cinco dias em estufa a 37°C e 5% de CO₂ e então fixadas com formalina 10% *overnight*, gentilmente lavadas e coradas com cristal violeta 1% por 20 minutos. As placas foram lavadas e colocadas para secar, contando a olho nu as placas de lise formadas. Poços contendo cerca de 10 a 100 placas de lise foram considerados para cálculo do título viral em unidades formadoras de placa (UFP/mL), corrigindo o resultado pelo fator de correção do volume de adsorção (10^1) e o inverso da diluição seriada utilizada para contagem.

4.2.7 Avaliação da concentração citotóxica 50% (CC₅₀)

Células L929 foram cultivadas em microplacas de 96 poços contendo $4,0 \times 10^4$ células/poço, em meio DMEM com 1% de SFB, e incubadas a 37°C e 5% de CO₂ por 24h. Logo após a incubação, o meio foi descartado e foram adicionados 200 µL de DMEM com 1% de SFB contendo diluições seriadas das substâncias (100 a 0,156 µM). Poços contendo uma diluição seriada de DMSO (na mesma proporção) foram utilizados como controle do veículo, contendo apenas meio como controle de viabilidade celular e, aqueles sem células, como controle de esterilidade e/ou branco. Logo após 72h de incubação sob as mesmas condições, a viabilidade celular foi avaliada empregando o teste colorimétrico de brometo de 3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazólio (MTT) (MOSMANN, 1983).

O meio foi removido e 100 µL de uma solução de MTT (ThermoFischer Scientific, EUA) a 0,5 mg/mL em DMEM com 1% de SFB foi adicionado a cada poço, seguido de incubação por 3h sob as mesmas condições. Este meio foi então removido e foram

adicionados 100 μL de DMSO para dissolução dos cristais de formazana. As placas permaneceram em agitação por 20 minutos e a leitura da absorbância foi feita em aparelho espectrofotômetro VersaMax (Molecular Devices, EUA), sob comprimento de onda de 570 nm. Os percentuais de inibição da viabilidade celular foram calculados utilizando o controle de veículo, determinando os valores da concentração citotóxica 50% (CC_{50}). Regressão linear foi utilizada para as análises ($r^2 > 0,9$). Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

4.2.8 Avaliação da concentração efetiva 50% (EC_{50}) e índices de seletividade (IS)

Monocamadas celulares (L929) foram cultivadas em microplacas de 96 poços contendo $4,0 \times 10^4$ células/poço em meio DMEM com 1% de SFB e incubadas a 37°C e 5% de CO_2 por 24h. Diluições seriadas (100 μL) das substâncias (até oito diluições abaixo dos valores de CC_{50}) foram incubadas sob as mesmas condições por 72h, juntamente de 100 μL das suspensões virais de MHV-3 em uma MOI de 0,1. Células infectadas com vírus foram incubadas apenas com meio contendo DMSO, na proporção das diluições das substâncias, foram considerados como controles. O tratamento com MTT seguiu conforme **item 4.2.7**. Ribavirina foi utilizado como controle positivo de inibição (10 μM). A concentração efetiva 50% (CE_{50}) foi calculada a partir da razão entre poços com células infectadas contendo as diluições das substâncias e poços com células infectadas contendo as diluições dos controles de DMSO. Índices de seletividade (IS) foram calculados pela razão entre os valores de CC_{50} e EC_{50} . Regressão linear foi utilizada para as análises ($r^2 > 0,9$). Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

NOTA: Estudos demonstraram a atividade antiviral de extratos e isolados triterpenos contra SARS-CoV-2 (YI et al., 2022). Dentre estes, moléculas isoladas (ALHADRAMI et al., 2021) e análogos sintéticos (BAEV et al., 2022) também já foram preditos e tiveram sua atividade inibitória contra M^{pro} de SARS-CoV-2 confirmada, utilizando estratégias computacionais similares àquelas no presente trabalho. Abordagens combinando *docking*, simulações de MD e triagens virtuais também já foram empregadas para identificação de triterpenos como potenciais inibidores (GIOFRÈ et al., 2021; WU et al., 2021). Nesse âmbito e, mediante as restrições ainda vigentes da parceria entre UCSD e UFMG deste projeto para testes com SARS-CoV-2 no NB3 da UFMG, uma série de extratos e triterpenos isolados (propositalmente descaracterizados) do Departamento de Química da UFMG foram incluídos para avaliação da atividade antiviral contra MHV-3, conforme descrito nos itens acima.

5 RESULTADOS E DISCUSSÃO

5.1 HQSAR de inibidores da M^{pro} de SARS-CoV-2: triagem de moléculas bioativas e avaliação da inibição da atividade enzimática

Na construção dos modelos de HQSAR foram selecionadas 25 moléculas baseadas no fármaco perampanel, com IC₅₀ entre 0,018 e 9,99 μM contra M^{pro} de SARS-CoV-2 (ZHANG et al., 2021). Alguns fatores favoreceram a seleção destas moléculas para as análises de HQSAR, como a estrutura geral conservada, o tamanho relativamente pequeno das estruturas e a variedade de grupos substituintes presentes nos análogos (**Figura 9**).

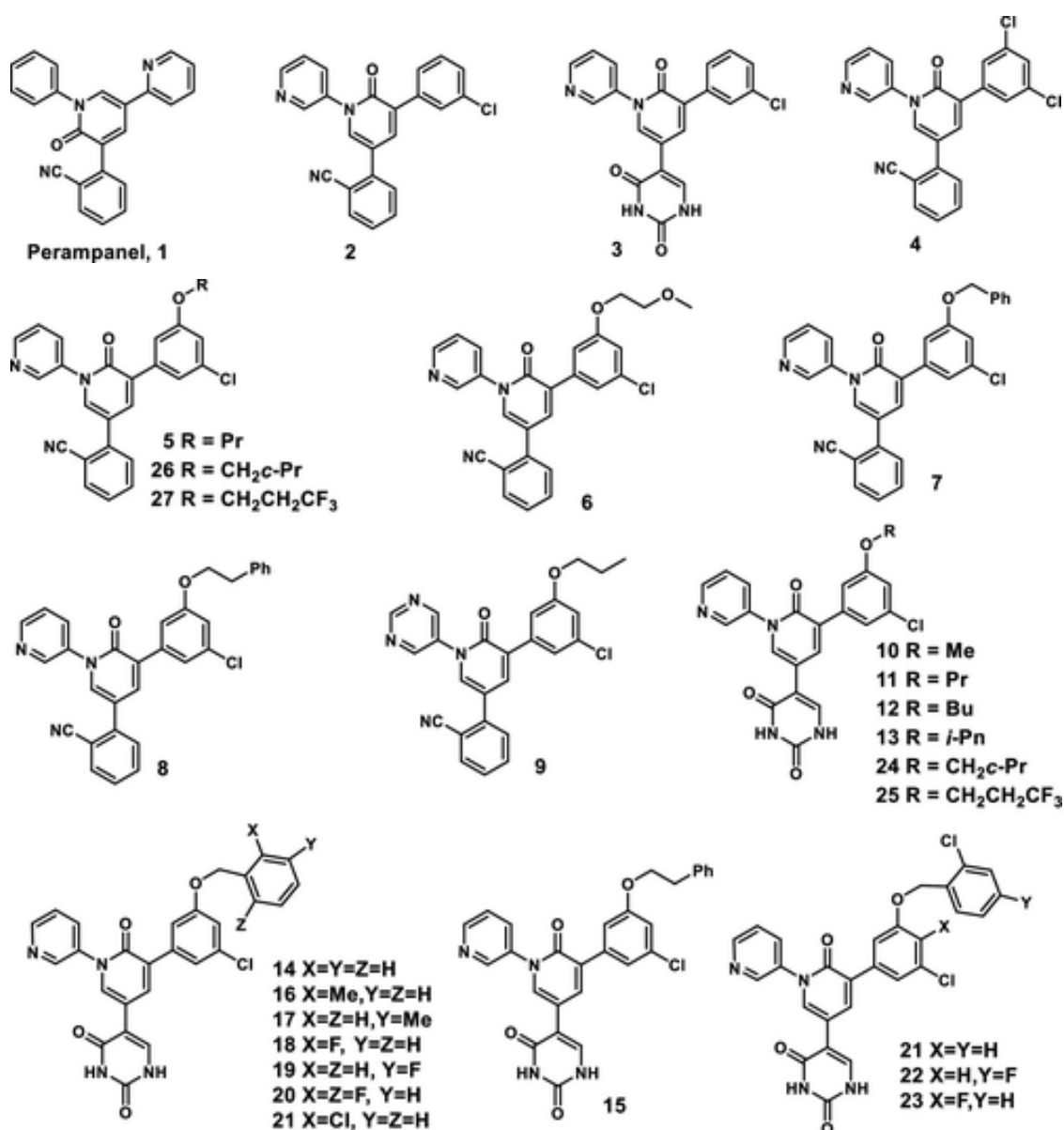


Figura 9. Estrutura geral das moléculas utilizadas na construção do conjunto de dados de HQSAR. Estruturas das 27 moléculas (ZHANG et al., 2021) e respectivos grupos substituintes (R, X, Y e Z) dos análogos de perampanel (1). As moléculas 1 e 9 não foram consideradas, já que não tiveram valores de IC₅₀ determinados.

Cada uma das 25 moléculas teve sua estrutura 3D construída, sendo individualmente conferidas quanto à sua estereoquímica. Esta etapa denota uma importância para análises computacionais, uma vez que a estereoquímica pode estar diretamente relacionada à atividade biológica de uma molécula e, além disso, a configuração do estereocentro, ou seja, a orientação R (sentido horário) ou S (sentido anti-horário) dos grupos substituintes, quando incorreta, pode induzir um modelo a erros de predição (WINKLER, 2002).

Seguidamente, as moléculas foram submetidas a um tratamento no programa OMEGA 4.1.1, obtendo os confôrmeros de menor energia de cada molécula, ou seja, a conformação mais estável. Essa etapa é importante, por exemplo, em análises computacionais como o HQSAR, uma vez que, quanto maior o número de moléculas e confôrmeros em um banco de dados, maior são as variações de coordenadas atômicas de hidrogênios nas moléculas (BORODINA et al., 2007). Finalmente, cada molécula teve seu estado de protonação corrigido de acordo com o pH 7,4, utilizando a aplicação fixpka no pacote computacional QUACPAC 2.1.2. Um estado iônico mais favorável contribui para um perfil preferível de possíveis interações intermoleculares, além de uma maior confiabilidade ao modelo avaliado, favorecendo, por exemplo, análises posteriores de *docking* (BOLCATO et al., 2019).

As 25 moléculas tiveram seus valores de IC₅₀ convertidos em pIC₅₀, ou seja, normalizados com log negativo da molaridade, assim facilitando a comparação da potência relativa de cada molécula. Estas foram então divididas aleatoriamente em dois conjuntos, treinamento (n = 20) e teste (n = 5). Adicionalmente, um segundo conjunto de treinamento foi gerado em triplicata (n = 60), uma estratégia proposta para aumentar a robustez preditiva de conjuntos de dados reduzidos (KRONENBERGER et al., 2018), considerando os valores de pIC₅₀ somados ou subtraídos de cada desvio-padrão (SD).

Os modelos de HQSAR foram criados fixando o tamanho do fragmento em relação ao número de átomos (4 a 7) (**Tabela 3**), sendo selecionado o modelo com maior valor de q^2 obedecendo a alguns critérios de validação de QSAR, como $q^2 > 0,5$ (GOLBRAIKH; TROPSHA, 2002), juntamente de valores mínimos de $r^2 > 0,6$ (idealmente superiores 0,65 e não muito distantes de q^2) (GRAMATICA; SANGION, 2016; TROPSHA; GRAMATICA; GOMBAR, 2003). Nesse sentido, o modelo selecionado superou os valores mínimos de q^2 (0,885) e ideais de r^2 (0,977), sendo ainda mais robusto no modelo de triplicata, com os valores de q^2 (0,964) e r^2 (0,975), respectivamente.

Tabela 3. Relação quantitativa entre a estrutura química e atividade biológica por holograma (HQSAR) do conjunto de dados em triplicata de inibidores da M^{pro} de SARS-CoV-2.

Distinção dos hologramas (4 a 7 átomos)	q ²	r ²	SEV	SEE	BL	COMP
A/B	0.954	0.976	0.165	0.119	353	6
A/B/C	0.951	0.974	0.169	0.123	257	6
A/B/C/H	0.962	0.975	0.148	0.121	83	6
A/B/C/H/Ch	0.964	0.975	0.145	0.120	83	6
A/B/H	0.959	0.973	0.154	0.126	53	6
A/B/C/Ch	0.954	0.977	0.164	0.116	353	6
A/B/DA	0.957	0.975	0.158	0.121	151	6
A/B/C/DA	0.959	0.975	0.155	0.121	59	6
A/B/H/DA	0.961	0.977	0.151	0.117	199	6
A/B/C/Ch/DA	0.960	0.976	0.152	0.119	59	6
A/B/H/Ch/DA	0.960	0.975	0.153	0.120	59	6
A/B/C/H/Ch/DA	0.959	0.975	0.154	0.120	151	6
A/B/C/H/DA	0.954	0.927	0.163	0.129	151	6

A: átomos; B: ligações; C: conectividade; HA: átomos de hidrogênio; Ch: quiralidade; DA: doadores e aceptores. Tamanho de fragmentos 4 a 7 átomos. q²: coeficiente de validação; r²: coeficiente de correlação não-validado; SEE: erro padrão da estimativa; SEV: erro padrão da validação; BL: extensão do holograma; COMP: componentes principais. O modelo selecionado (maior valor de q²) está destacado em negrito. Análises foram feitas no programa SYBYL-X 2.1.

Esses dois parâmetros (q² e r²) corroboram diferentes métricas de correlação por regressão, a exemplo de testes de coeficientes de validação de q² (coeficiente de correlação preditiva quadrático) e médias de r², ambos importantes individualmente e em concordância, testificando a robustez de um modelo nos métodos de QSAR (CHAI; DRAXLER, 2014; CHIRICO; GRAMATICA, 2011; KIRALJ; FERREIRA, 2009). No presente trabalho, a análise de HQSAR com tamanho fixo dos fragmentos (4 a 7 átomos) apresentou átomos, ligações, conectividade, átomos de hidrogênio e quiralidade (A, B, C, H e Ch) como o modelo mais robusto, com maior valor da métrica de correlação interna, q² (0,964).

Esse resultado indica que a conectividade, ou seja, o estado de hibridização (C) dos átomos na molécula, é importante para prever a atividade biológica, assim como a estereoquímica de uma molécula (orientação R ou S), representada no modelo pela quiralidade (Ch). Ademais, apresentar uma distinção além de átomos somente, por exemplo,

átomos de hidrogênio, torna este um modelo mais simples para predições. As ligações e a estereoquímica também podem sugerir uma correlação importante com a conectividade de átomos, o que também reforça a importância da estrutura e de orientações das moléculas, junto dos respectivos grupos substituintes em cada estrutura (GRAMATICA, 2020).

Pensando nisso, o modelo de maior valor de q^2 também foi submetido a variações dos tamanhos dos fragmentos, em intervalos de 4 átomos (1 a 4 até 7 a 10), mantendo sempre os descritores do modelo (A/B/C/H/Ch). Dentre todos seis modelos gerados, nenhum superou valores dos parâmetros empregados (Tabela 4).

Tabela 4. Variação do tamanho (quatro átomos) dos fragmentos do modelo HQSAR selecionado.

Distinção dos hologramas	Tamanho	q^2	r^2	SEV	SEE	BL	COMP
A/B/C/H/Ch	1_4	0.931	0.951	0.201	0.170	83	6
A/B/C/H/Ch	2_5	0.950	0.967	0.171	0.140	83	6
A/B/C/H/Ch	3_6	0.958	0.970	0.156	0.132	83	6
A/B/C/H/Ch	5_8	0.952	0.968	0.168	0.137	83	6
A/B/C/H/Ch	6_9	0.940	0.963	0.186	0.146	83	6
A/B/C/H/Ch	7_10	0.942	0.963	0.184	0.146	83	6

A: átomos; B: ligações; C: conectividade; HA: átomos de hidrogênio; Ch: quiralidade; DA: doadores e aceptores. Tamanho de fragmentos de quatro átomos. q^2 : coeficiente de validação; r^2 : coeficiente de correlação não-validado; SEE: erro padrão da estimativa; SEV: erro padrão da validação; BL: extensão do holograma; COMP: componentes principais. Análises foram feitas no programa SYBYL-X 2.1.

Essa abordagem permite avaliar todos os fragmentos possíveis no holograma, incluindo sobreposições e a contagem de cada vez em que um fragmento ocorre. Dessa forma, com as possibilidades de fragmentos com tamanhos entre átomos cobertas no *fingerprint*, os hologramas codificam as informações químicas das moléculas, incluindo sua topologia, representando dados mais robustos de correlação de um modelo de predição (HERITAGE; LOWIS, 1999; GARCIA; HONÓRIO, 2011).

Uma vez que q^2 é uma métrica que isoladamente pode subestimar a qualidade preditiva de um modelo ao avaliar apenas as moléculas que foram utilizadas na construção do modelo, demais validações são necessárias para se obter um modelo confiável (GOLBRAIKH, 2002). Nesse caso, uma validação externa e outros métodos de validações cruzadas são importantes parâmetros de caracterização e validação da capacidade preditiva entre diferentes modelos de QSAR (GRAMATICA; SANGION, 2016).

Sendo assim, o modelo em triplicata selecionado foi submetido a uma validação externa no pacote computacional KNIME®. Mapas de contribuição química foram gerados (**Figura 10**), evidenciando regiões de contribuições previstas nas moléculas, sendo positivas (verde e amarelo), neutras (branco) e/ou negativas (laranja e vermelho) para a atividade biológica contra o alvo proposto. As moléculas não apresentaram regiões previstas como contribuição negativa e ressaltam a importância (contribuição positiva; em verde) dos anéis aromáticos laterais (fenil), bem como das piridinas, além dos substituintes cloro e flúor nas moléculas 23 e 21m (21 – SD).

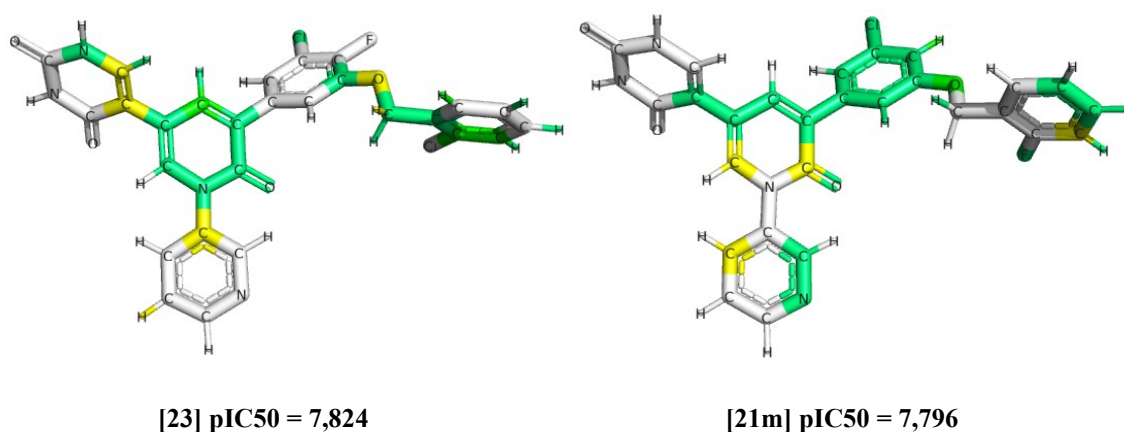


Figura 10. Mapas de contribuição química das moléculas com maiores valores de pIC₅₀ (23 e 21m). Regiões de contribuições positivas (verde e amarelo), neutras (branco) e/ou negativas (laranja e vermelho) são destacadas. Molécula 23 (esquerda) e 21m = 21 – SD (direita) apresentam os maiores valores de pIC₅₀. Códigos dos átomos de carbono, cloro, flúor, hidrogênio, nitrogênio e oxigênio foram destacados em preto. Imagens foram geradas no programa SYBYL-X 2.1.

Esses anéis aromáticos já foram sugeridos em interações com sítios de interesse na M^{pro} de SARS-CoV-2, a exemplo de resíduos alifáticos em bolsos hidrofóbicos, como demonstrado em interações com os resíduos Leu167 e Pro168 (ZHANG et al., 2021). Ademais, a presença de átomos de nitrogênio (piridina) na molécula **23**, bem como dos substituintes cloro e flúor ligados (clorofenil e fluorofenil) da molécula **21m**, também podem ser favoráveis, por exemplo, em ligações de hidrogênio. Estes substituintes, por sua vez, foram preditos em ligações de hidrogênio com His163 e Glu166, além de interações com anel imidazólico de His41, sendo as moléculas **21m** e **23** inibidores com IC₅₀ determinada ($0,018 \pm 0,002$ e $0,020 \pm 0,005$ μM , respectivamente) contra a enzima (ZHANG et al., 2021).

Aljuhani et al., (2022), também avaliaram pequenas moléculas derivadas e análogos de piridina, incluindo a presença de substituintes cloro no anel aromático, na atividade antiviral contra SARS-CoV e SARS-CoV-2 em células Vero, além da atividade inibitória contra M^{pro}

de SARS-CoV-2. Análises de *docking* demonstraram interações preditas com His41, Gly143, Cys145, His163 e Glu166, similares àquelas observadas no estudo de Zhang et al., (2021). Adicionalmente, valores de IC₅₀ de até 0,67 µM foram obtidos contra SARS-CoV-2, além de valores de EC₅₀ de até 0,021 µM e 0,03 µM nos ensaios antivirais contra SARS-CoV-2 e SARS-CoV, respectivamente (ALJUHANI et al., 2022). Luo et al., (2022), também identificaram contribuições positivas de anéis piridina, substituintes cloro e flúor de inibidores da M^{pro} de SARS-CoV-2 em análises de HQSAR, além de prever as posições dos substituintes nos anéis aromáticos com contribuições mais negativas para a atividade inibitória das moléculas (LUO et al., 2022).

Sabendo da importância de determinados fragmentos estruturais e, considerando essa capacidade do HQSAR em prever estruturas ou grupos substituintes de interesse em moléculas bioativas com atividade enzimática contra a M^{pro} de SARS-CoV-2, algumas características das estruturas avaliadas foram então consideradas em uma avaliação de descritores moleculares (**Tabela 5**) empregando o software PaDEL (YAP, 2011). Dessa forma, ao se construir e avaliar descritores moleculares 2D, *fingerprints* e propriedades físico-químicas das estruturas do HQSAR, seria possível prever demais moléculas com potencial atividade biológica em outros bancos de dados, como a quimioteca BraCoLi da UFMG.

Tabela 5. Características dos descritores moleculares utilizados no software PaDEL

Descritor	Características
Klekotha	Presença de subestruturas do HQSAR (A, B, C, Ch, H e DA)
MACCS	<i>keys</i> = número de moléculas de cada átomo
PubChem	Presença de fragmentos moleculares
AtomPairs2D	Distância entre um par de átomos no plano 2D
<i>Substructure</i>	Total de subestruturas
<i>count</i>	Análise quantitativa

MACCS: *Molecular ACCess System*

Os descritores calculados passaram por uma validação externa com um fluxo de trabalho empregado no pacote computacional KNIME®, sendo classificados de acordo com os valores do coeficiente de correlação preditiva quadrático (CCC), de q^2 (correlação interna) e da média de r^2 (r^2_{avg}), ou seja, a média da correlação externa. Nesta análise, também são avaliados os parâmetros do modelo de *random forest*, sendo o número de iterações ou níveis máximos (do inglês, *MaxLevels*) e o número de modelos (do inglês, *Nmodels*), variando de 10

a 50 níveis e 10.000 a 50.000 modelos. PubChem e AtomPairs2DCount apresentaram os maiores valores preditivos nas três métricas avaliadas, sendo maiores que 0,86 para CCC e 0,75 para q^2 , superando todos os demais descritores em pelo menos 10% (0,100). Assim, ambos foram considerados para serem empregados nos modelos de triagem baseados em resultados do HQSAR, sendo validados comparativamente quanto aos valores preditivos de classificação (**Figura 11**).

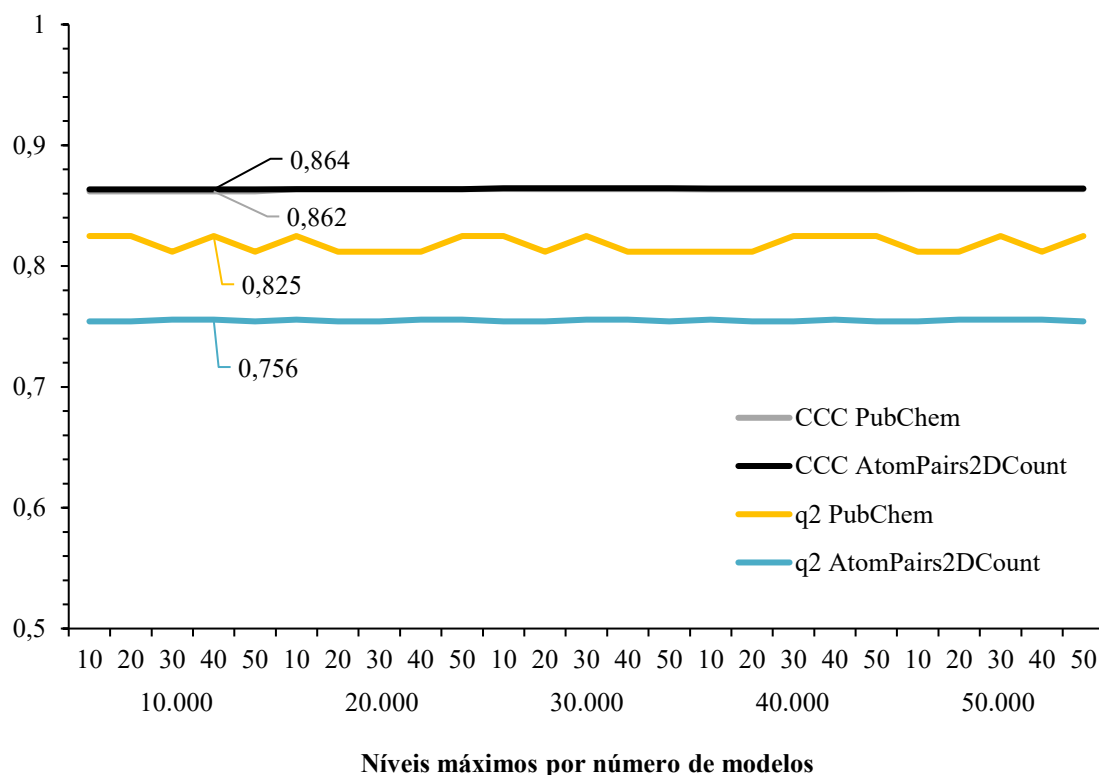


Figura 11. Seleção de descritores moleculares empregando aleatorização de N modelos. Descritores com maiores valores preditivos de classificação comparando parâmetros de q^2 e CCC entre PubChem (0,825 e 0,862) e AtomPairs2DCount (0,756 e 0,864).

Veríssimo et al., (2019), empregaram um modelo similar de QSAR baseado em *random forest* na predição da atividade biológica de derivados de nitroimidazol planejados contra cepas susceptíveis e resistentes de *Trichomonas vaginalis* (VERÍSSIMO et al., 2019). Neste estudo também foram avaliadas variações de *MaxLevels* e *Nmodels* na obtenção de modelos com valores de q^2 entre 0,619 e 0,697, similares àqueles obtido no presente trabalho (0,825 de PubChem e 0,756 de AtomPairs2DCount).

Finalmente, a triagem da quimioteca BraCoLi foi feita pelo método de *random forest* no pacote computacional KNIME®, obtendo 24 moléculas oriundas dos dois modelos (**Apêndice B**), 20 da predição com o descritor PubChem e quatro com AtomPairs2DCount,

estabelecendo um valor de corte dos valores de pIC_{50} em 6,5 (correspondente a 0,3 μ M). Estas seguiram para as predições de toxicidade e farmacocinética e aos ensaios de inibição da atividade enzimática da M^{pro} de SARS-CoV-2.

Previamente aos ensaios de inibição da atividade enzimática, as moléculas foram avaliadas em três servidores para avaliação de propriedades farmacocinéticas, farmacodinâmicas e de toxicidade: Pred-Herg 4.2, ProTox-II e admetSAR 2.0. Dentre as 24 moléculas (aqui denominadas QF1 a QF24) avaliadas com Pred-hERG 4.2, nenhuma foi predita como potencial bloqueador de canais para potássio. Considerando o modelo de admetSAR 2.0 (**Tabela 6**), todas as moléculas apresentaram predição positiva para absorção intestinal e permeabilidade da barreira hematoencefálica, além de interações com plasma entre 50,46 e 88,64%. Apenas duas moléculas, QF21 e QF22, não apresentaram uma predição positiva para hepatotoxicidade.

Tabela 6. Predições de toxicidade e farmacocinética das 24 moléculas com admetSAR 2.0

QF	Hepatotoxicidade	Absorção intestinal	B.B.B.	Biodisponibilidade (via oral)	Interação plasmática
1	+	+	+	+	88,64 %
2	+	+	+	-	75,24 %
3	+	+	+	+	85,76 %
4	+	+	+	-	75,24 %
5	+	+	+	-	69,27 %
6	+	+	+	-	74,18 %
7	+	+	+	-	74,18 %
8	+	+	+	+	50,46 %
9	+	+	+	+	54,38 %
10	+	+	+	-	72,82 %
11	+	+	+	+	51,96 %
12	+	+	+	+	77,15 %
13	+	+	+	+	58,45 %
14	+	+	+	+	73,67 %
15	+	+	+	+	60,73 %
16	+	+	+	+	63,80 %
17	+	+	+	-	63,71 %
18	+	+	+	-	63,71 %
19	+	+	+	-	63,06 %
20	+	+	+	+	71,55 %
21	-	+	+	+	59,71 %
22	-	+	+	+	69,89 %
23	+	+	+	+	50,67 %
24	+	+	+	-	73,74 %

+ predição positiva (verde-água); - predição negativa (laranja). B.B.B.: barreira hematoencefálica.

O ProTox-II também foi empregado para predição de toxicidade oral das 24 moléculas, com apenas uma molécula predita como ativa para imunotoxicidade. Todas foram preditas como classe V ou superior para dose letal mediana (DL_{50}), ou seja, prejudiciais se ingeridas ($2000 < DL_{50} \leq 5000$ mg/kg), de acordo com o sistema globalmente harmonizado de classificação e rotulagem de produtos químicos (do inglês, *globally harmonized system of classification and labelling of Chemicals*; GHS) (TA, 2021). Caso alguma das moléculas apresentasse atividade inibitória ($> 50\%$) contra M^{pro} de SARS-CoV-2, essas moléculas deveriam ser avaliadas com maior cautela, por exemplo, nos ensaios celulares.

Finalmente, as 24 moléculas triadas foram avaliadas quanto a sua atividade inibitória da atividade enzimática da M^{pro} de SARS-CoV-2. Os ensaios enzimáticos consistem, inicialmente, na leitura da clivagem de um substrato fluorogênico durante determinado período. Na avaliação da inibição da atividade de uma enzima, como a M^{pro} , a diferença de leitura entre um controle e as moléculas (potenciais inibidores) avaliadas permitem calcular percentuais de inibição (%), assim determinando uma atividade inibitória. Dentre as 24 moléculas selecionadas das triagens virtuais, no entanto, nenhuma daquelas testadas apresentou inibição superior a 50% quando avaliadas a 10 μ M (**Tabela 7**).

Tabela 7. Atividade inibitória das moléculas triadas por HQSAR contra M^{pro} de SARS-CoV-2

Molécula	% inibição (10 μ M ^a)	Molécula	% inibição (10 μ M ^a)	Molécula	% inibição (10 μ M ^a)
QF1	0,0 \pm 0,0	QF9	ND	QF17	0,0 \pm 0,0
QF2	0,0 \pm 0,0	QF10	ND	QF18	0,0 \pm 0,0
QF3	0,0 \pm 0,0	QF11	0,0 \pm 0,0	QF19	0,0 \pm 0,0
QF4	5,0 \pm 2,9	QF12	0,0 \pm 0,0	QF20	0,0 \pm 0,0
QF5	1,0 \pm 0,6	QF13	0,0 \pm 0,0	QF21	0,0 \pm 0,0
QF6	0,8 \pm 0,8	QF14	0,0 \pm 0,0	QF22	ND
QF7	0,0 \pm 0,0	QF15	0,0 \pm 0,0	QF23	ND
QF8	5,0 \pm 1,8	QF16	0,0 \pm 0,0	QF24	ND

^a O percentual de inibição é representado como a média e o erro do desvio-padrão calculados entre um ensaio independente em triplicata. O erro é apresentado como a razão entre o desvio-padrão e a raiz quadrada do número de replicatas/leituras ($n = 3$). ND: não determinada.

Alguns fatores podem explicar a baixa acurácia preditiva do modelo, culminando na ausência de atividade inibitória ($> 50\%$) das moléculas frente à M^{pro} de SARS-CoV-2. A definição de um domínio de aplicabilidade, ou seja, o espaço químico definido pelo conjunto de moléculas do conjunto treinamento, por exemplo, é uma característica importante em modelos de QSAR (MATHEA; KLINGSPOHN; BAUMANN, 2016), uma vez que as

predições são sugeridas como confiáveis quando as moléculas estão inseridas em um determinado domínio (ALVES et al., 2018). Ainda, o domínio de aplicabilidade também é restrito ao próprio tamanho e à diversidade do conjunto treinamento utilizado e, estando uma molécula fora de um domínio, a predição pode ser incorreta, reiterando a importância dessa etapa para aceitabilidade de um modelo de QSAR e para predição de moléculas de outras quimiotecas (TROPSHA, 2010).

Ademais, modificações nas variáveis empregadas no método de HQSAR poderiam aumentar a robustez da técnica para as predições, como a seleção dos três ou cinco modelos com maiores valores de q^2 na primeira etapa, além de mais variações dos tamanhos dos fragmentos dos modelos aqui selecionados, por exemplo, em intervalos de 2, 3 e 4 átomos (TROPSHA; GRAMATICA; GOMBAR, 2003; GRAMATICA, 2020). Igualmente, validações complementares do modelo avaliado poderiam incrementar a confiabilidade do método para as predições, empregando alguns testes de validação independentes. Nesse sentido, poderiam ser empregados, por exemplo: (i) um coeficiente de validação cruzada, em que se excluiria uma molécula de um conjunto, construindo um novo modelo, auxiliando na predição do valor desta variável dependente do modelo selecionado previamente; uma (ii) aleatorização de Y, em que se avaliaria a robustez (q^2) e o erro de predição (SEV) na probabilidade de ocorrer uma correlação dependente de cada variável em um modelo; e (iii) *leave many out* (LMO) ou *leave-N-out*, no qual as moléculas do conjunto treinamento seriam divididas em subgrupos múltiplos de 5, avaliando a robustez do modelo em função de diferentes tamanhos de subgrupos de validação (KIRALJ; FERREIRA, 2009).

Sendo assim, tomados em conjunto, uma estratégia baseada no planejamento de fármacos baseados na estrutura de um ligante (LBDD) (LIMA et al., 2016), conforme discutido anteriormente, pode auxiliar na descoberta de moléculas bioativas contra SARS-CoV-2 e outros coronavírus (PANT et al., 2020; WU et al., 2020; SERAFIM et al., 2021a). Contudo, considerando os resultados dos ensaios enzimáticos no presente trabalho, as análises de HQSAR, de forma independente, não foram capazes de prever com acurácia moléculas bioativas contra a M^{Pro} de SARS-CoV-2, pelo menos nas concentrações avaliadas (10 μ M). Diante desses resultados, uma estratégia baseada no planejamento de fármacos baseadas na estrutura de um alvo (SBDD) também foi empregada, sendo apresentada nos tópicos a seguir.

5.2 Estratégia de planejamento de moléculas bioativas baseado na estrutura de alvos

5.2.1 Seleção de estruturas de enzimas no PDB

Aproximadamente 405 estruturas cristalográficas da M^{pro} de SARS-CoV-2, e 39 da PL^{pro}, se encontravam depositadas no PDB até janeiro de 2021. Dentre elas, uma estrutura de cada enzima foi selecionada para as análises estruturais, sendo PDB ID: 7D1M (FU et al., 2020) para M^{pro} e PDB ID: 6WX4 (RUT et al., 2020) para PL^{pro}. As duas estruturas cristalográficas atendiam aos critérios estabelecidos anteriormente, como resolução < 2.0 Å e presença um inibidor co-cristalizado (**Tabela 8**), sendo as duas com menor valor de resolução na ocasião. Ademais, ambas apresentavam estruturas aqui denominadas “estruturas-irmãs”, ou seja, a mesma enzima expressa e determinada experimentalmente com diferentes ligantes. Os resíduos ausentes na PL^{pro} se referem ao final da cadeia lateral, que até o presente momento não foi demonstrada interagindo ou tendo relação nas interações com inibidores.

Tabela 8. Características das estruturas cristalográficas das proteases de SARS-CoV-2 selecionadas no banco de dados *Protein Data Bank* (PDB, disponível em <https://www.rcsb.org/>)

Informação/Estrutura	M ^{pro} (PDB ID: 7D1M)	PL ^{pro} (PDB ID: 6WX4)
Resolução	1.5 Å	1.66 Å
Mutações	-	-
Fragmentos ausentes	1 (cadeia A); 1-2 (cadeia B)	321-324
Presença de inibidor	GC376 (IC ₅₀ = 0,15 + 0,03 µM)	VIR251 (IC ₅₀ = ~50 µM)
Método	Difração de raios-X	Difração de raios-X
Modelo de expressão	<i>Escherichia coli</i> K-12	<i>Escherichia coli</i> BL21
Regiões favoráveis (<i>Ramachandram</i>)	> 99,9%	> 99,9%
Referência	FU et al., 2020	RUT et al., 2020

DOI: Digital Object Identifier.

Visando complementar as análises estruturais e expandir a abordagem na obtenção de potenciais inibidores “pancoronavírus” (CHEN, 2020; SERAFIM et al., 2021b), estruturas das proteases de distintos coronavírus que infectam seres humanos e outros mamíferos, além de aves, foram incluídas (**Tabela 9**). As análises destes alvos seguiram o mesmo fluxograma disposto nos resultados dos subtópicos a seguir, a fim de se evidenciar resíduos conservados entre as estruturas de todos os coronavírus disponíveis no PDB em fevereiro de 2021.

Tabela 9. Estruturas cristalográficas de enzimas de demais coronavírus selecionadas no banco de dados *Protein Data Bank* (PDB, disponível em <https://www.rcsb.org/>)

Organismo/Estrutura	M ^{pro} (PDB ID / Å)	PL ^{pro} (PDB ID / Å)
SARS-CoV-2	7D1M / 1.5 Å	6WX4 / 1.66 Å
SARS-CoV	6W2A / 1.65 Å	4OW0 / 2.1 Å
MERS-CoV	6VH3 / 2.2 Å	4RNA / 1.79 Å
HCoV-229E	2ZU2 / 1.8 Å	-
HCoV-HKU1	3D23 / 2.5 Å	-
HCoV-NL63	5GWY / 2.85 Å	-
Morcego (HKU4)	2YNB* / 1.96 Å	-
Camundongo (MHV)	6JJI / 2.65 Å	-
Porco (TGEV)	1P9U / 2.37 Å	3MP2 / 2.5 Å
Porco (PDCOV)	-	6LN0 / 2.46 Å
Galinha (IBV)	-	4X2Z / 2.15 Å

HCoV: Human coronavirus; MHV: Mouse hepatitis virus; TGEV: Transmissible gastroenteritis virus; PDCOV: Porcine deltacoronavirus; IBV: Avian infectious bronchitis virus. *a ser publicada

Dentre essas estruturas, especificamente a PL^{pro} de Transmissible gastroenteritis virus (TGEV), apresenta diferenças estruturais, inclusive em relação ao seu sítio catalítico, quando comparada às proteases dos demais coronavírus que infectam mamíferos, como SARS-CoV e HCoV-229E (WOJDYLA et al., 2010). Esses fatores foram levados em consideração para as análises estruturais nos tópicos a seguir, uma vez que uma mudança no sítio catalítico implicaria em diferenças nas predições de sítios de ligação e de interações com ligantes.

5.2.2 Mapeamento das proteases de SARS-CoV-2 e de outros coronavírus

A predição e caracterização dos sítios de ligação das proteases de SARS-CoV-2 foi conduzida inicialmente nos servidores FTSite e FTMap. O servidor FTSite realizou a predição de três possíveis sítios de ligação nas enzimas, em que se ressalta a importância de cavidades próximas às regiões catalíticas (sítios ativos), em que um ligante poderia interagir para atuar como um efetivo inibidor. O arquivo resultante do FTMap, por sua vez, apresenta agrupamentos de moléculas-sonda divididos por sítios de ligação, as quais interagem com os resíduos daqueles sítios delimitados no FTSite. Os resultados demonstram a consonância entre os resultados de agrupamentos de moléculas-sonda no FTMap e dos sítios preditos no FTSite, como demonstrado para PL^{pro} de SARS-CoV-2 (**Figura 12**), em que são evidenciados, por exemplo, três grupamentos de sondas em cada um dos três sítios de ligação preditos, respectivamente.

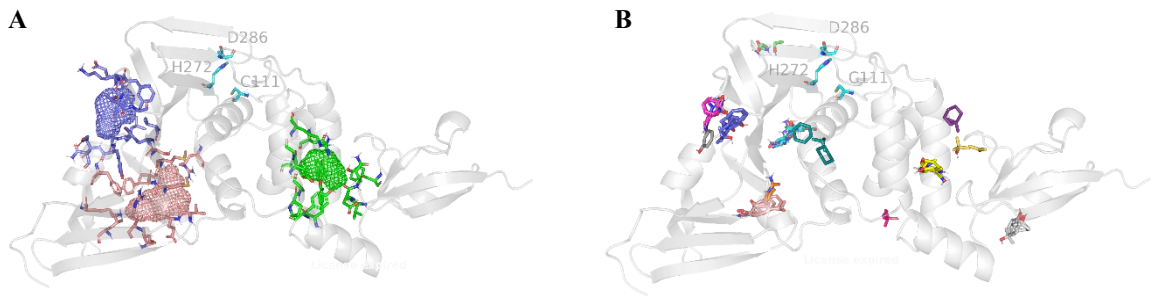


Figura 12. Potenciais sítios de ligação e grupamentos de moléculas-sonda preditos na PL^{pro} de SARS-CoV-2. (A) Predição de sítios de ligação calculadas pelo servidor FTSite na PL^{pro} de SARS-COV-2 (cinza). Sítio 1 (rosa), sítio 2 (verde) e sítio 3 (azul) representados em redes tridimensionais, com os respectivos aminoácidos destacados. (B) Agrupamentos de moléculas-sonda calculadas pelo servidor FTMap. Grupamentos 00 (ciano), 03 (salmão) e 06 (laranja) respectivos ao sítio 1; grupamentos 02 (amarelo), 10 (bege) e 11 (roxo) respectivos ao sítio 2; e grupamentos 01 (magenta), 05 (lilás) e 12 (cinza), respectivos ao sítio 3. Resíduos catalíticos Cys111, His272 e Asp286 destacados (ciano). Imagens foram geradas com o programa PyMOL (v0.99c).

Igualmente, os resultados entre os agrupamentos das sondas e os sítios com potenciais resíduos favoráveis de interação da M^{pro} também se corroboraram entre os servidores. Foram cinco grupamentos de sondas no sítio 1, três no sítio 2 e mais dois no sítio 3 (**Figura 13**).

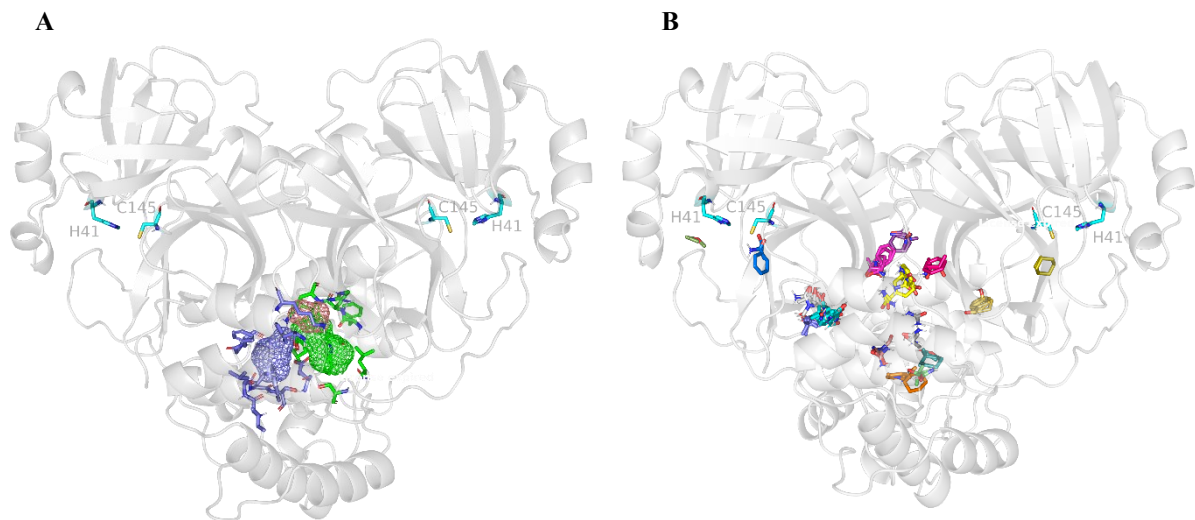


Figura 13. Potenciais sítios de ligação e grupamentos de moléculas-sonda preditos na M^{pro} de SARS-CoV-2. (A) Predição de sítios de ligação calculadas pelo servidor FTSite na M^{pro} de SARS-COV-2 (cinza). Sítio 1 (rosa), sítio 2 (verde) e sítio 3 (azul) representados em redes tridimensionais. (B) Agrupamentos de moléculas-sonda calculadas pelo servidor FTMap. Grupamentos 02 (amarelo), 03 (salmão), 06 (laranja), 07 (lima) e 12 (cinza) respectivos ao sítio 1; grupamentos 00 (ciano), 04 (branco) e 05 (azul) respectivos ao sítio 2; e grupamentos 08 (verde azulado) e 10 (bege), respectivos ao sítio 3. Resíduos catalíticos His41 e Cys145 destacados (ciano). Imagens foram geradas com o programa PyMOL (v0.99c).

O mesmo método foi empregado às enzimas dos demais coronavírus com estruturas disponíveis. Igualmente, sítios e agrupamentos similares foram preditos entre as diferentes M^{pro} no servidor FTSite (**Figura 14**).

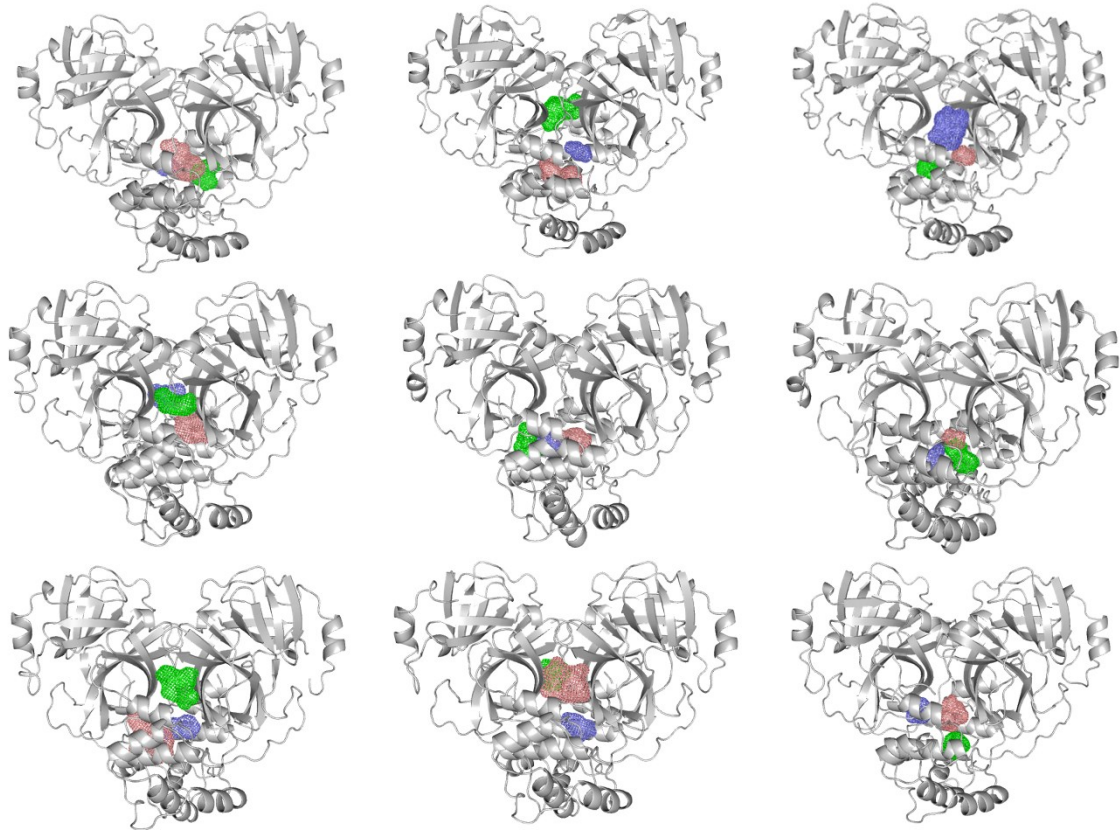


Figura 14. Potenciais sítios de ligação preditos nas M^{pro} de diferentes coronavírus. Predição das cavidades sugeridas como sítios de ligação calculadas pelo servidor FTSite, sendo sítio 1 (rosa), sítio 2 (verde) e sítio 3 (azul) representados em redes tridimensionais (esquerda para direita): SARS-CoV, MERS-CoV, SARS-CoV-2, HCoV-229E, HCoV-HKU1, HCoV-NL63, HKU4 (morcego), MHV (camundongo) e TGEV (porco), coloridas em prata. Nestas análises não foram evidenciados resíduos do sítio catalítico preditos nas estruturas das enzimas. Imagens foram geradas com o programa PyMOL (v0.99c).

Nas predições dos dois programas não foi possível identificar ao menos os resíduos dos sítios catalíticos das proteases, His41 e Cys145 na M^{pro} , e Cys111, His272 e Asp286 na PL^{pro} . No entanto, ao simular a predição das M^{pro} com somente uma das cadeias, ou seja, avaliando um único monômero, as predições foram significativamente diferentes das anteriores. Diversos resíduos de interesse foram identificados concomitantes nesta protease de SARS-CoV-2 (**Figura 15**) e dos demais coronavírus, a exemplo de Met49, Phe140, Asn142, Gly143, Ser144, His163, His164, Met165 e Glu166, incluindo aqueles da díade catalítica, His41 e Cys145.

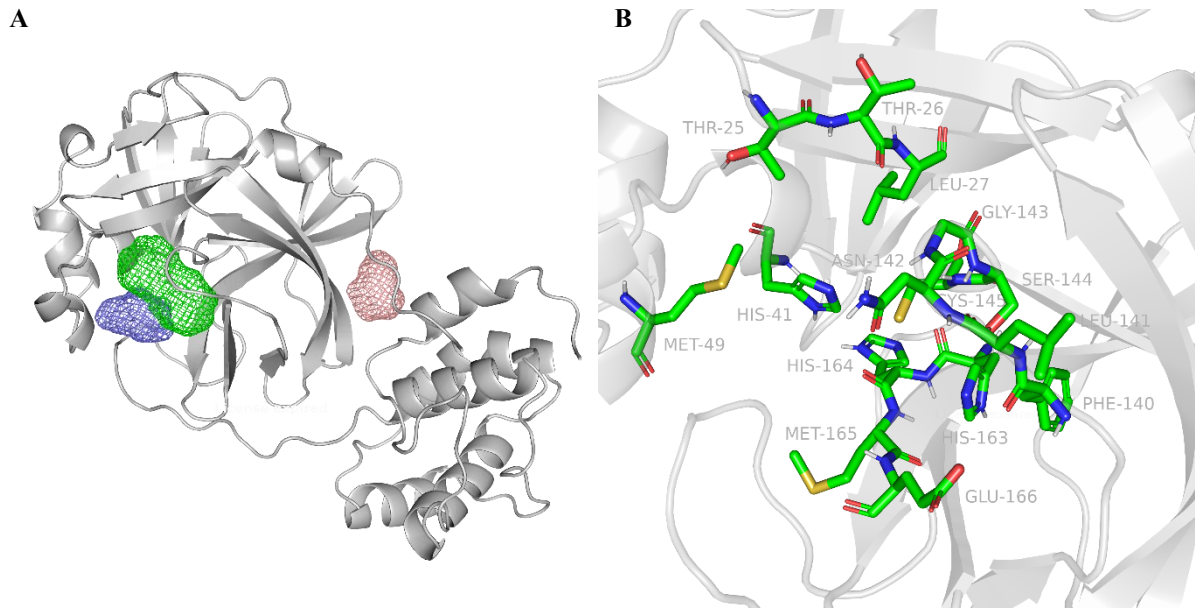


Figura 15. Predições de sítios de ligação com FTSite em um monômero da M^{pro} de SARS-CoV-2. As predições empregando FTSite resultaram em (A) três sítios no monômero (cadeia A) da M^{pro} de SARS-CoV-2, com destaque no (B) sítio 2 (verde). Nestas análises foram preditos os resíduos do sítio catalítico, His41 e Cys145, além de Thr25, Thr26, Leu27, His41, Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165 e Glu166. Imagens foram geradas com o programa PyMOL (v0.99c).

Silva et al., (2021), também identificaram três sítios correspondentes na M^{pro} de SARS-CoV-2, agrupando um total de 16 moléculas-sonda interagindo com resíduos de interesse, por exemplo, His41 e Cys145 (SILVA et al., 2021), quando avaliado o monômero da protease, corroborando os dados evidenciados no presente trabalho. Evidenciando essas diferenças nas análises das primeiras predições, o servidor PrankWeb foi incluído para análise de possíveis sítios de interação nos alvos moleculares, a fim de identificar os resíduos de interesse com maior confiabilidade. Igualmente para esta etapa, não somente as proteases de SARS-CoV-2 foram incluídas, mas também as enzimas com estrutura cristalográfica disponível dos demais coronavírus.

Sítios similares foram preditos nas análises de PrankWeb, agora evidenciando regiões que apresentaram os resíduos dos sítios catalíticos das enzimas, a exemplo das M^{pro} dos diferentes coronavírus (**Figura 16**). Nestas, um sítio de interesse incluindo a díade catalítica foi predito em todas as estruturas avaliadas, sempre na porção esquerda do monômero da cadeia A, sendo representados nas cores azul ou vermelho. Demais sítios preditos foram destacados com cores padrão do programa (vermelho, salmão, marrom, magenta, verde, laranja, branco, azul e amarelo).

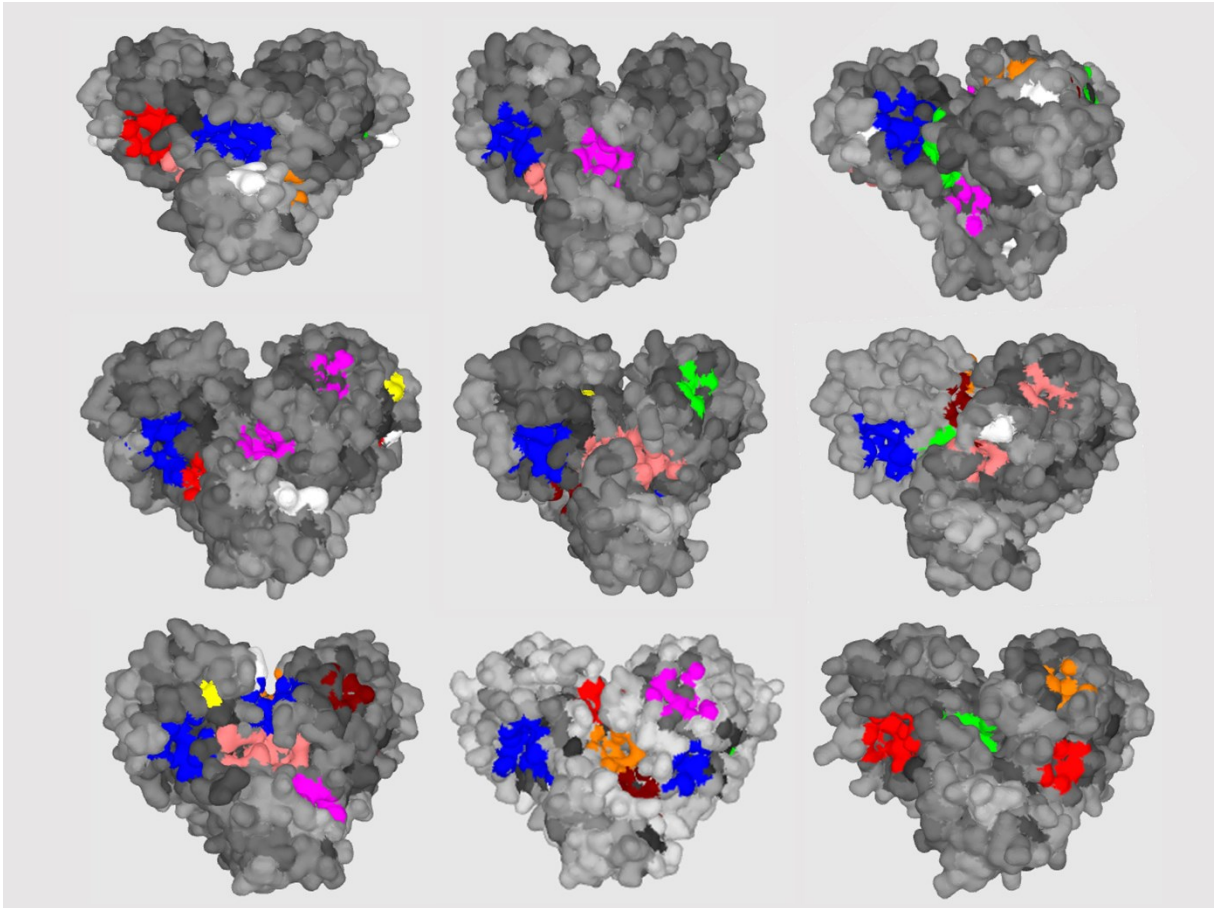


Figura 16. Predições de potenciais sítios de ligação com PrankWeb nas M^{Pro} de diferentes coronavírus. As predições empregando PrankWeb resultaram em diversos potenciais sítios nas M^{Pro} de (esquerda para direita): SARS-CoV, MERS-CoV, SARS-CoV-2, HCoV-229E, HCoV-HKU1, HCoV-NL63, HKU4 (morcego), MHV (camundongo) e TGEV (porco). O sítio de interesse, que inclui a díade catalítica, se encontra na porção esquerda do monômero (cadeia A), representados de azul ou vermelho, concomitantes em todas as estruturas das enzimas. Demais sítios preditos foram destacados (vermelho, salmão, marrom, magenta, verde, laranja, branco, azul e amarelo). Imagens foram geradas diretamente do servidor PrankWeb.

O método do servidor empregado se baseia em um modelo de ferramentas P2Rank, que utiliza um conjunto de preditores baseados em ML, apresentando assim uma maior robustez e acurácia preditiva ao comparar distintos estados conformacionais nos alvos, agrupando as regiões de interesse (JENDELE et al., 2019). Cada agrupamento predito recebe uma pontuação de classificação individual, ranqueando e priorizando os sítios de ligação. Estes resultados corroboram as análises do monômero da M^{Pro} , sendo considerados para a construção do modelo farmacofórico nas etapas subsequentes no presente trabalho. Novamente, resíduos como Thr25, Thr26, His41, Phe140, Gly143, Ser144, Cys145, His163, His164, Glu166 e Gln189 foram preditos entre todas as estruturas (**Tabela 10**), representados no ranqueamento dos três sítios (quando aplicável) com maiores pontuações para as M^{Pro} .

Tabela 10. Potenciais resíduos de interação preditos com PrankWeb nas diferentes M^{pro}

Alvo	Sítio #1 Score (confiabilidade, %)	Sítio #2 Score (confiabilidade, %)	Sítio #3 Score (confiabilidade, %)
7D1M	15.78 (60,5%)	14.71 (56,9%)	11.67 (44,5%)
AA	B_140 B_141 B_142 B_143 B_144 B_145 B_163 B_164 B_165 B_166 B_187 B_188 B_189 B_25 B_26 B_27 B_41 B_49	A_140 A_141 A_142 A_143 A_145 A_163 A_164 A_165 A_166 A_187 A_188 A_189 A_25 A_26 A_27 A_41 A_49	A_2 A_214 A_299 A_3 A_302 A_303 A_304 A_306 A_4 B_137 B_138 B_139 B_140 B_141 B_166 B_169 B_170 B_171 B_172
6W2A	13.19 (51,0%)	12.65 (48,9%)	12.11 (46,6%)
AA	A_116 A_118 A_123 A_124 A_139 A_141 B_111 B_127 B_151 B_152 B_292 B_294 B_295 B_298 B_299 B_6 B_7 B_8	A_140 A_141 A_142 A_143 A_144 A_145 A_163 A_164 A_165 A_166 A_172 A_187 A_188 A_189 A_25 A_26 A_27 A_41 A_49	B_140 B_141 B_142 B_143 B_144 B_145 B_163 B_164 B_165 B_166 B_187 B_188 B_189 B_25 B_26 B_27 B_41 B_49
6VH3	13.20 (51,1%)	10.49 (38,9%)	9.44 (33,8%)
AA	B_143 B_144 B_145 B_146 B_147 B_148 B_166 B_167 B_168 B_169 B_190 B_191 B_192 B_24 B_25 B_26 B_41 B_49	A_119 A_121 A_126 A_142 A_144 B_114 B_115 B_116 B_130 B_154 B_155 B_156 B_292 B_294 B_295 B_298 B_299 B_302 B_6 B_7 B_8	A_143 A_144 A_145 A_146 A_147 A_148 A_166 A_167 A_168 A_169 A_175 A_190 A_191 A_192 A_25 A_26 A_41 A_49
2ZU2	11.74 (44,8%)	10.97 (41,1%)	
AA	A_140 A_141 A_142 A_143 A_144 A_145 A_163 A_164 A_165 A_166 A_187 A_188 A_189 A_25 A_26 A_27 A_41 A_49 A_52	B_140 B_141 B_142 B_143 B_145 B_163 B_164 B_165 B_166 B_187 B_188 B_189 B_25 B_26 B_41 B_44 B_49	
3D23	13.05 (50,5%)	9.98 (36,6%)	
AA	B_140 B_141 B_142 B_144 B_162 B_163 B_164 B_165 B_186 B_187 B_188 B_25 B_26 B_27 B_41 B_47 B_51	A_139 A_140 A_141 A_142 A_143 A_144 A_162 A_163 A_164 A_165 A_186 A_187 A_188 A_25 A_26 A_41 A_47 A_51	
5GWY	15.37 (59,2%)	13.49 (52,2%)	10.86 (40,6%)
AA	A_139 A_140 A_141 A_142 A_144 A_163 A_164 A_165 A_166 A_187 A_188 A_189 A_25 A_26 A_27 A_41 A_44 A_46 A_47 A_51 B_2	A_127 A_128 A_130 A_136 A_197 A_199 A_200 A_279 A_280 A_281 A_283 A_284 A_285 A_286 A_287 A_288 A_3 A_4 A_5 B_136 B_277 B_279 B_280 B_281 B_282 B_283 B_287 B_3 B_4 B_5	B_139 B_140 B_141 B_142 B_144 B_163 B_164 B_165 B_166 B_187 B_188 B_189 B_25 B_26 B_27 B_41 B_47
2YNB	14.14 (54,8%)	13.01 (50,4%)	
AA	B_144 B_145 B_146 B_147 B_148 B_166 B_167 B_168 B_169 B_190 B_191 B_192 B_25 B_26 B_27 B_41 B_49	A_143 A_144 A_145 A_146 A_147 A_148 A_166 A_167 A_168 A_169 A_190 A_191 A_192 A_25 A_26 A_41 A_46 A_49	

AA: resíduos de aminoácidos; A_ : cadeia A; B_ : cadeia B; 7D1M: SARS-CoV-2; 6W2A: SARS-CoV; 6VH3: MERS-CoV; 2ZU2: HCoV-229E; 3D23: HCoV-HKU1; 5GWY: HCoV-NL63; 2YNB: HKU4 (morcego); 6JIJ: MHV (camundongo); e 1P9U: TGEV (porco).

Tabela 10. Potenciais resíduos de interação preditos com PrankWeb nas diferentes M^{pro} (continuação)

Alvo	Sítio #1 Score (confiabilidade, %)	Sítio #2 Score (confiabilidade, %)	Sítio #3 Score (confiabilidade, %)
6JIJ	16.81 (78,7%)	16.49 (78%)	14.90 (74,6%)
AA	A_1 B_138 B_139 B_140 B_141 B_142 B_143 B_161 B_162 B_163 B_164 B_185 B_186 B_187 B_25 B_26 B_27 B_39 B_41 B_48	A_125 A_135 A_277 A_278 A_279 A_281 A_282 A_283 A_284 A_285 A_3 A_4 A_5 B_125 B_135 B_277 B_278 B_279 B_281 B_282 B_283 B_284 B_285 B_3 B_4 B_5	A_138 A_139 A_140 A_141 A_143 A_161 A_162 A_163 A_164 A_185 A_186 A_187 A_25 A_26 A_41 A_44 A_47 A_48 A_50
1P9U	12.32 (47,5%)	7.34 (22,9%)	
AA	A_136 A_278 A_279 A_280 A_282 A_283 A_284 A_285 A_286 A_287 A_3 A_4 A_5 B_136 B_198 B_278 B_279 B_280 B_282 B_283 B_284 B_285 B_286 B_287 B_3 B_4 B_5	A_141 A_142 A_144 A_163 A_164 A_165 A_186 A_187 A_188 A_25 A_26 A_27 A_41 A_47 A_51	

AA: resíduos de aminoácidos; A: cadeia A; B: cadeia B; 7D1M: SARS-CoV-2; 6W2A: SARS-CoV; 6VH3: MERS-CoV; 2ZU2: HCoV-229E; 3D23: HCoV-HKU1; 5GWY: HCoV-NL63; 2YNB: HKU4 (morcego); 6JIJ: MHV (camundongo); e 1P9U: TGEV (porco).

Nelakuditi e Shrivastava (2021), empregaram a mesma técnica no planejamento de fármacos de reposicionamento contra M^{pro} de SARS-CoV-2, obtendo no sítio de ligação predito mais bem ranqueado os resíduos Thr25, Thr26, Leu27, His41, Met49, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Asp187, Arg188 e Gln189 (NELAKUDITI; SHRIVASTAVA, 2021). Liang et al., (2020), também já haviam demonstrado a consistência de PrankWeb quando comparado ao SiteMap (HALGREN, 2007), outro método de predição. Na análise de três estruturas cristalográficas da M^{pro} de SARS-CoV-2 foram evidenciados, por exemplo, His41, Gly143, Cys145, His163 e Gln189 (LIANG et al., 2020b), também corroborando aos dados obtidos no presente trabalho.

Embora comparado entre apenas seis estruturas de diferentes coronavírus, um sítio principal também foi evidenciado para as PL^{pro} (**Figura 17**), incluindo de TGEV. Os resíduos preditos, por sua vez, não foram compilados nesta seção devido a uma predição inconsistente do número total de sítios entre as proteases dos diferentes coronavírus, sendo: oito preditos para PDCOV, seis para IBV, três para SARS-CoV-2, dois para SARS-CoV e MERS-CoV, e somente um para TGEV.

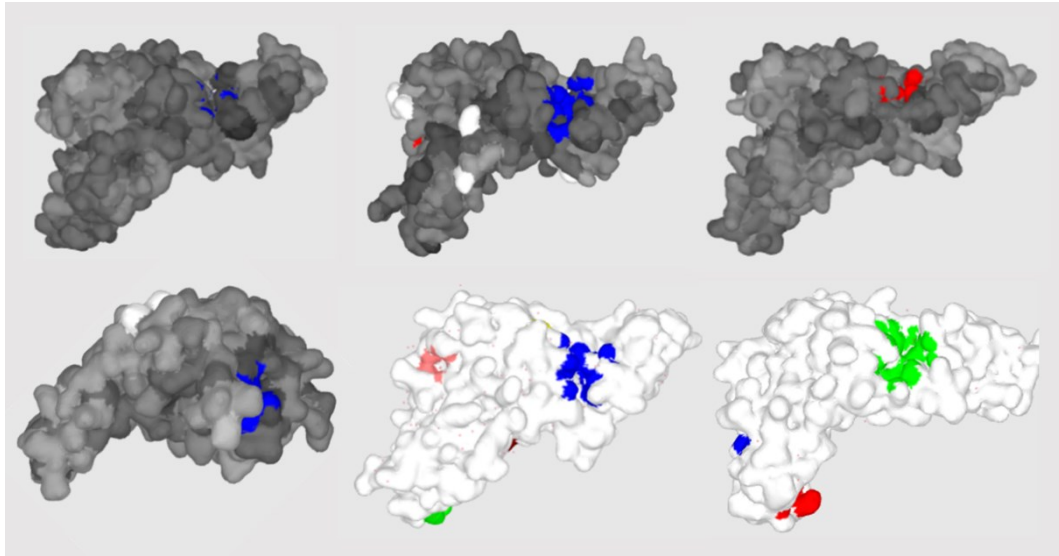


Figura 17. Predições de potenciais sítios de ligação com PrankWeb nas PL^{pro} de diferentes coronavírus. As predições empregando PrankWeb resultaram em potenciais sítios nas PL^{pro} de (esquerda para direita): SARS-CoV, MERS-CoV, SARS-CoV-2, TGEV (porco), IBV (galinha) e PDCOV (porco). Sítio principal evidenciado na porção direita em azul, vermelho ou verde. Demais sítios preditos foram destacados (vermelho, salmão, verde, branco e azul). Imagens foram geradas diretamente do servidor PrankWeb.

Ademais, foi observada uma inconsistência entre os sítios e resíduos das PL^{pro}, inconsistentes entre os alvos avaliados. Adicionalmente, os três resíduos da tríade catalítica (Cys111, His272 e Asp286) foram preditos somente para PL^{pro} de SARS-CoV-2 (**Tabela 11**).

Tabela 11. Potenciais resíduos de interação preditos com PrankWeb na PL^{pro} de SARS-CoV-2

Alvo	Sítio #1 Score (confiabilidade, %)	Sítio #2 Score (confiabilidade, %)	Sítio #3 Score (confiabilidade, %)
6WX4	6.72 (37,3%)	3.09 (10,5%)	2.85 (9,0%)
AA	D_106 D_109 D_111 D_162 D_266 D_270 D_271 D_272 D_274 D_286	D_104 D_106 D_111 D_114 D_115 D_272 D_273 D_285 D_286 D_287 D_289	D_12 D_13 D_131 D_134 D_135 D_138 D_146 D_37 D_56 D_71 D_87

AA: resíduos de aminoácidos; D: cadeia principal.

Embora uma localização equivalente ou similar entre algumas PL^{pro} possa ter sido observada para o sítio que compreenderia a região da tríade catalítica, a inconsistência de predições entre estas proteases de diferentes coronavírus chama atenção para este alvo, diferente dos resultados para M^{pro}. Esse perfil também foi observado por Pitsillou et al., (2020), anteriormente, avaliando PrankWeb nas PL^{pro} de MERS-CoV, SARS-CoV e SARS-CoV-2, com apenas Asp164 e/ou Asp302 concomitantes entre os três alvos, sem identificar resíduos catalíticos em nenhuma das enzimas (PITSILLOU et al., 2020).

Os seis resíduos representados, His41, Phe140, Gly143, Cys145, His163 e Glu166, juntamente de Thr25, Thr26, His164 e Gln189, foram observados em todas as predições de PrankWeb para as M^{pro} de SARS-CoV-2 e de outros coronavírus. Contudo, no alinhamento de sequências, apenas os seis primeiros são conservados dentre todas as estruturas avaliadas. Ainda que ocorra uma substituição de aminoácidos, caso mantida a mesma propriedade entre resíduos na mesma posição (p.ex.: hidrofóbicos), uma interação com ligantes ainda poderia ser favorável, apresentando pouco ou nenhum efeito em um sítio de ligação, como demonstrado por Cho et al., (2021) para as M^{pro} de SARS-CoV e SARS-CoV-2 (CHO et al., 2021).

Nesse sentido, um alinhamento 3D por sobreposição das estruturas também foi feito no programa PyMOL (v0.99c), a fim de se evidenciar o estado conformacional dos resíduos conservados e/ou substituídos nas estruturas das M^{pro}. O alinhamento dos seis resíduos supracitados também apresentou posição e orientação espacial conservadas na estrutura das enzimas. Contudo, dentre os demais resíduos em potencial, como His164 e Gln189, as substituições para Gln164 e Pro189 nas M^{pro} de HCoV-229E e HCoV-NL63, por exemplo, conferindo diferenças conformacionais que podem alterar as interações com potenciais ligantes (**Figura 19**). As demais representações 3D das PL^{pro} foram desconsideradas, uma vez que não há correspondência entre os sítios ativos, representando apenas o alinhamento de suas sequências (**Apêndice C**).

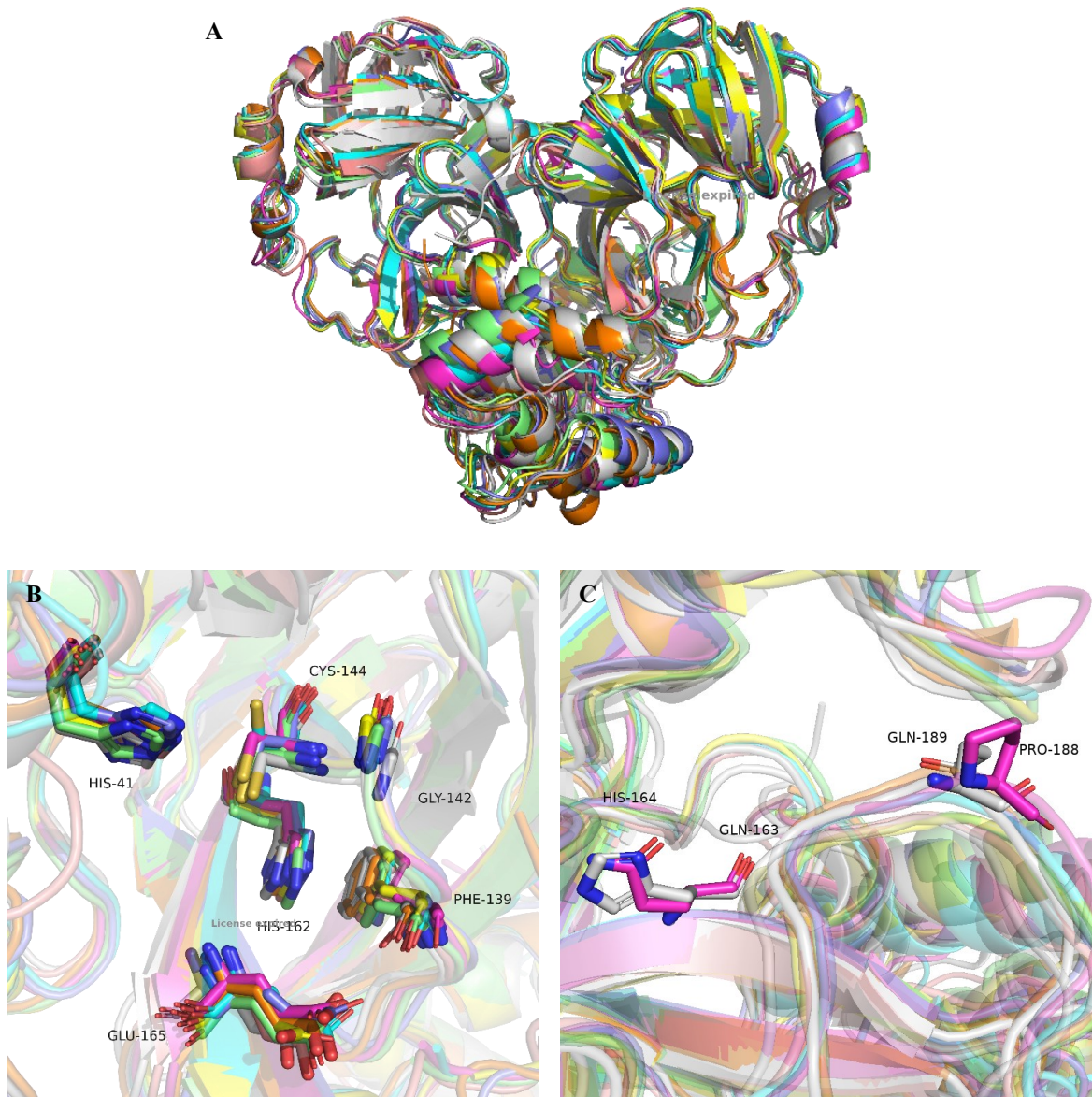


Figura 19. Alinhamento 3D de estruturas das M^{pro} de diferentes coronavírus. (A) O alinhamento 3D de estruturas das M^{pro} no programa PyMOL demonstrou a orientação conservada das conformações dos (B) resíduos His41, Phe140, Gly143, Cys145, His163 e Glu166, bem como da influência na orientação das conformações de (C) resíduos não-conservados referente às posições 164 e 189 (substituições H164Q e Q189P; magenta) de HCoV-229E (posições 163 e 188, em razão do alinhamento das sequências). Imagens foram geradas com o programa PyMOL (v0.99c).

5.2.3 Análises de interações entre ligantes e alvos moleculares de SARS-CoV-2 e diferentes coronavírus

Sabendo que simulações de interações moleculares entre alvos e ligantes são importantes para predição do modo de inibição de moléculas em diferentes alvos moleculares (AZEVEDO et al., 2022), o *redocking* dos ligantes co-cristalizados das proteases também foi

feito a fim de se prever potenciais resíduos de interação. Especificamente para M^{PRO}, uma vez que resíduos de interesse para interação com ligantes foram preditos de forma consistente (His41, Phe140, Gly143, Cys145, His163 e Glu166), evidenciá-los no *redocking* seria de interesse para etapas posteriores, como a construção do farmacóforo.

Nesta etapa, o valor do desvio quadrático médio (do inglês, *Root Mean Score Deviation*; RMSD) entre a pose do ligante co-cristalizado e a pose do ligante reproduzido na estrutura (*redocking*), bem como as pontuações de *docking*, devem ser considerados nas análises. Nesse sentido, um valor de RMSD baixo (≤ 1.0 Å) é considerado suficiente e representativo para comparação de análises de acoplamento com ligantes conhecidos, juntamente da pose de maior score, de acordo com cada algoritmo genético dentre os programas de *docking* (CASTRO-ALVAREZ; COSTA; VILARRASA, 2017; AZEVEDO et al., 2022). Não obstante, valores de RMSD próximos a 2.0 - 2.5 Å também podem ser considerados, a depender do tamanho e da flexibilidade (ligações rotacionáveis) da molécula atuando como ligante em uma estrutura. Essa consideração também se faz importante dado aos diferentes programas, suas respectivas faixas de pontuação e acurácia preditiva (WANG et al., 2016; AZEVEDO et al., 2022), bem como de possíveis especificidades por determinados alvos moleculares ou famílias de proteínas (CROSS et al., 2009), além da resolução (Å) dos cristais de alvos em estudo (AGRAWAL et al., 2019).

O *redocking* do ligante co-cristalizado (GC376) da M^{PRO} de SARS-CoV-2 (PDB ID: 7D1M) foi feito no programa GOLD, e os parâmetros empregados foram: (i) ligações rotacionáveis da proteína mantidas como fixas; (ii) interação com o ligante co-cristalizado e átomos dentro um espaço de até 6 Å; (iii) chemscore_kinase como modelo de interação; (iii) um total de 200 reproduções com o algoritmo genético (*GA runs*); (iv) CHEMPLP como função de pontuação, sem permitir encerramentos prematuros (RMSD < 0,5 Å); e (v) busca do algoritmo genético selecionada como lenta (maior acurácia). Foram identificados interagindo com o ligante co-cristalizado: (i) Thr26, Gly143, Cys145, Met165, Glu166, Pro168 e Gln189 na cadeia A (RMSD = 3,63 Å); e (ii) Phe140, Gly143, Met165, Glu166 e Gln189 na cadeia B (**Figura 20**) (RMSD = 2,19 Å).

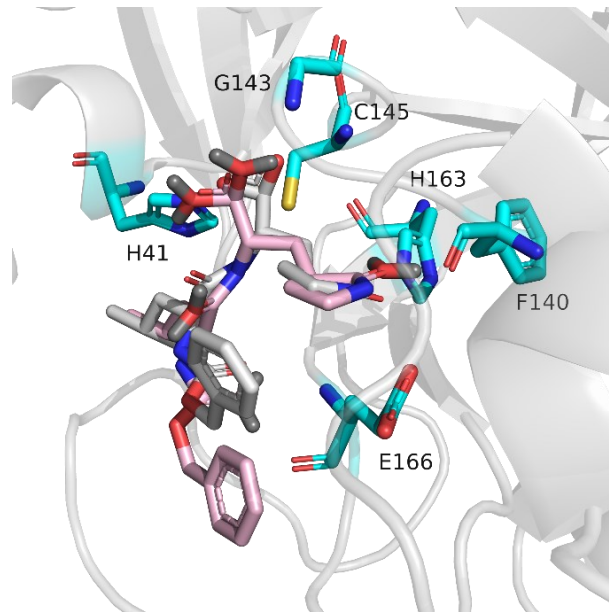


Figura 20. Redocking de GC376 na cadeia B da M^{pro} de SARS-CoV-2, utilizando programa GOLD. Representação da pose (rosa) da molécula de GC376 no *redocking* (RMSD de 2.19 Å) na M^{pro} (PDB ID: 7D1M) de SARS-CoV-2 (cinza), comparado ao ligante co-cristalizado (branco) obtido experimentalmente na estrutura cristalográfica. Os resíduos de interesse His41, Phe140, Gly143, Cys145, His163 e Glu166 (ciano) foram destacados. Imagens foram geradas com o programa PyMOL (v0.99c).

Considerando que os mesmos resíduos preditos nas análises estruturais da M^{pro} de SARS-CoV-2 (PDB ID: 7D1M) no presente trabalho corroboram parcialmente aos dados de *docking* do estudo de origem (FU et al., 2020), que demonstrou uma ligação covalente do inibidor GC376 com Cys145, além de interações com Phe140, Gly143, His163, His164 e Glu166 utilizando o programa GLIDE. Sendo Phe140, Gly143, His163 e Glu166 concomitantes às predições de sítios de ligação, as análises de *docking* também foram feitas com outras estruturas cristalográficas da M^{pro} de SARS-CoV-2, a fim de evidenciar se as variações de interação entre as cadeias A e B com outros ligantes também seriam mantidas. Nesse sentido, foram comparadas as estruturas denominadas “estruturas-irmãs” de 7D1M, as estruturas 7C6U (GC376), 7C6S e 7BRP (boceprevir), além de outras estruturas disponíveis utilizando o programa GOLD (Tabela 12).

Tabela 12. Resíduos de interação com ligantes co-cristalizados nas M^{pro} de SARS-CoV-2 preditos em análises com programa GOLD

	PDB	Ligante	PDB	Ligante	PDB	Ligante	PDB	Ligante
Complexo	7D1M	GC376	7C6U	GC376	7C6S	boceprevir	7BRP	boceprevir
Cadeia	A	B	A	B	A	B	A	B
AA	Cys145	Cys145	His164	His164	His164		Asn142	Asn142
	Gly143	Gly143	His163	His163	His41	His41	Cys145	Cys145
	Phe140	Phe140	Glu166	Glu166	Glu166	Glu166	His164	His164
	Glu166	Glu166	Cys145	Cys145	Cys145	Cys145	Glu166	Glu166
	Gln189	Gln189	Gly143	Gly143	Gly143	Gly143	Thr26	Thr26
	His164	His164	Phe140	Phe140	Met49	Met49		Ser144
					Thr26	Thr26		Leu141
	PDB	Ligante	PDB	Ligante	PDB	Ligante	PDB	Ligante
Complexo	6LU7	N-3	7BQY	N-3	6Y2F	Cetoamida	6Y7M	β-carbamato
Cadeia	A	B	A	B	A	B	A	B
AA	Phe140	Phe140	Phe140	Phe140	His41	His41	Phe140	Phe140
	Gly143	Gly143	Gly143	Gly143	Phe140	Phe140	Gly143	Gly143
	Cys145	Cys145	Cys145	Cys145	Gly143	Gly143	Cys145	Cys145
	Glu166	Glu166	Glu166	Glu166	Cys145	Cys145	His164	His164
	Gln189	Gln189	Gln189	Gln189	His164	His164	Glu166	Glu166
	Thr190	Thr190	Thr190	Thr190	Glu166	Glu166		
			Met49	Met49				

AA: resíduos de aminoácidos; A: cadeia A; B: cadeia B; Códigos PDB das M^{pro} de SARS-CoV-2: 7D1M, 7C6U, 7C6S, 7BRP, 6LU7, 7BQY, 6Y2F e 6Y7M.

Esses dados corroboram as análises de *docking* já observadas na literatura de diversos estudos avaliando o inibidor da M^{pro} de *Feline infectious peritonitis virus* (FIPV), um coronavírus felino, GC376 (KIM et al., 2012), com a M^{pro} de SARS-CoV-2 (FU et al., 2020; LUAN et al., 2023), bem como das M^{pro} de SARS-CoV e MERS-CoV (HU et al., 2020). Nesse sentido e, pensando na estratégia de um inibidor pancoronavírus, a reprodução dos ligantes co-cristalizados das M^{pro} dos demais coronavírus também foram comparadas com programa GOLD (Tabela 13), a fim de se avaliar a consistência de interações ocorrendo com os resíduos preditos anteriormente nas análises estruturais (His41, Phe140, Gly143, Cys145, His163 e Glu166).

Tabela 13. Resíduos de interação com ligantes co-cristalizados nas M^{pro} de outros coronavírus preditos em análises com programa GOLD

Coronavírus	229E	Humano	NL63	Humano	HKU1	Humano	HKU4	Morcego
	PDB	Ligante	PDB	Ligante	PDB	Ligante	PDB	Ligante
Complexo	2ZU2	MPD	5GWY	N-3	3D23	N-3	2YNB	G85
Cadeia	A*	B*	A	B	A	B	A***#	B***
AA	Phe139		Phe139	Glu166	Phe140		His41	
	Asn141		Gly142	Ser190	Gly143		Phe143	
	His162		Cys144		His163		Cys148	
	Glu165		His163		Glu166		His166	
			Gln164		Glu189		Gln167	
							Glu169	
							Gln192	
Coronavírus	SARS	Humano	MERS	Humano	MHV	Camundongo	TGEV	Porco
	PDB	Ligante	PDB	Ligante	PDB	Ligante	PDB	Ligante
Complexo	6W2A	sulfonic	6VH3	sulfonic	6JIJ	PJE-010	1P9U	MRD
Cadeia	A	B	A***	B***	A**	B**	A*	B*
AA	His41	His41	His41	His41	His41	Phe138	His41	
	Cys145	Cys145	Leu49	Leu49	Phe138	Gly141	Phe139	
	Cys148	His163	Cys148	Cys148	Gly141	Cys143	His162	
	His163	His164	Gln167	His166	His161	Gln162	His163	
	His164	His166	His166	Gln167	Glu164	Glu164	Leu164	
	Glu166	Pro168	Glu169	Met168	Gln187	Asp185	Leu166	
	Pro168	Asp187	Ala171	Glu169	Val188	Gln187	Ser189	
	Gln189	Gln189		Ala171		Val188		
	Ala191	Ala191		His194				

AA: resíduos de aminoácidos; A: cadeia A; B: cadeia B; 2ZU2: HCoV-229E; 3D23: HCoV-HKU1; 5GWY: HCoV-NL63; 2YNB: HKU4 (morcego); 6W2A: SARS-CoV; 6VH3: MERS-CoV; 6JIJ: MHV (camundongo); e 1P9U: TGEV (porco). *corresponde a n-1 na sequência de resíduos a partir de Phe140 (Phe139); **n-2: Phe140 (Phe138); e ***n+3: Phe140 (Phe143). #Resíduos Leu49, Leu170, Phe188 e Gln195 foram omitidos da tabela.

Pelo menos dois dos seis resíduos anteriormente preditos foram observados nas interações, o que corrobora a importância desses aminoácidos como potenciais sítios de interação. Estudos da literatura também já demonstraram a importância destes resíduos nas interações com inibidores em análises de *docking*, como His41 e Glu166 nas M^{pro} de SARS-CoV-2, SARS-CoV, MERS-CoV, 229E, NL-63, HKU1, HKU4 e MHV (JIN et al., 2020a). Adicionalmente, His163 e Glu166 também foram demonstrados nas interações com nirmatrelvir (OWEN et al., 2021), além de Cys145 nas diferenças de interação entre nirmatrelvir e ritonavir (Paxlovid) na M^{pro} de SARS-CoV-2 (DAWOOD, 2022).

Subsequente a essas análises, o *redocking* da PL^{pro} de SARS-CoV-2 (PDB ID: 6WX4) também foi feito, identificando os resíduos interagindo com o ligante co-cristalizado, VIR251, um inibidor planejado para esta protease (RUT et al., 2020). Trp106, Cys111, Gly163, Asp164, Met208, Tyr264, Tyr268, Gly271 e His272 foram identificados (RMSD = 2.45 Å), incluindo dois dos três resíduos catalíticos (**Figura 21**).

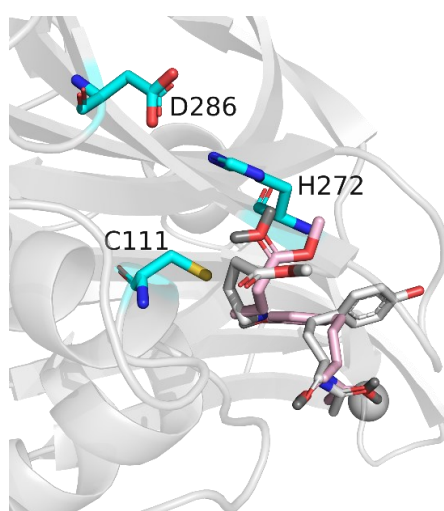


Figura 21. Redocking de VIR251 na PL^{pro} de SARS-CoV-2, utilizando programa GOLD. Representação da pose da molécula de VIR251 (rosa) no *redocking* (RMSD = 2.45 Å) da PL^{pro} (PDB ID: 6WX4), destacando os resíduos da tríade catalítica Cys111, His272 e Asp286 (ciano), íon zinco (cinza) e ligante co-cristalizado (branco) obtido experimentalmente. Imagens foram geradas com o programa PyMOL (v0.99c).

Tomados em conjunto, esses resultados corroboram os resíduos de interesse nas proteases de SARS-CoV-2, bem como das M^{pro} de outros coronavírus, ambos de interesse, por exemplo, na construção dos modelos farmacofóricos de cada enzima. Nesse sentido, sabendo quais resíduos interagem com determinados ligantes, também pode ser avaliada a estabilidade e frequência de interações em simulações ao longo do tempo, empregando simulações de dinâmica molecular, a fim de validar tais aminoácidos para a construção dos modelos farmacofóricos a serem empregados nas triagens virtuais.

5.2.4 Simulações de dinâmica molecular: M^{pro} e PL^{pro} de SARS-CoV-2

As simulações de dinâmica molecular (DM) das proteases foram conduzidas, inicialmente, com ligantes co-cristalizados das duas proteases. Buscando simular os movimentos de átomos a nível molecular, as DM permitiram avaliar mudanças conformacionais dos ligantes e a frequência de interações entre os complexos alvo-ligante ao longo do tempo, corroborando pontos de interação preditos (PrankWeb e *redocking*).

Especificamente para PL^{pro} de SARS-CoV-2, os dados considerados se basearam no estudo de Ferreira et al., (2022) (FERREIRA et al., 2022) e foram apresentados nesta seção.

Inicialmente, simulações de 1 μ s (5 x 200 ns) foram feitas para a estrutura alvo da M^{pro} de SARS-CoV-2 (PDB ID: 7D1M), predizendo contatos (interações) de diferentes grupamentos químicos do inibidor GC376 com resíduos Phe140 (48%), Cys145 (93%), His163 (90%), His164 (75%), Glu166 (96%) e Gln189 (74%), além de Thr26 (30%) e Gly143 (40%) mediados por moléculas de água (**Figura 22**). O resíduo da díade catalítica Cys145 interage com um grupo hidroxila do ligante, bem como Thr26 e Gly143, mediados por uma molécula de água. Os resíduos Phe140, His163, His164, Glu166 e Gln189, por sua vez, se orientam a grupos amida de GC376.

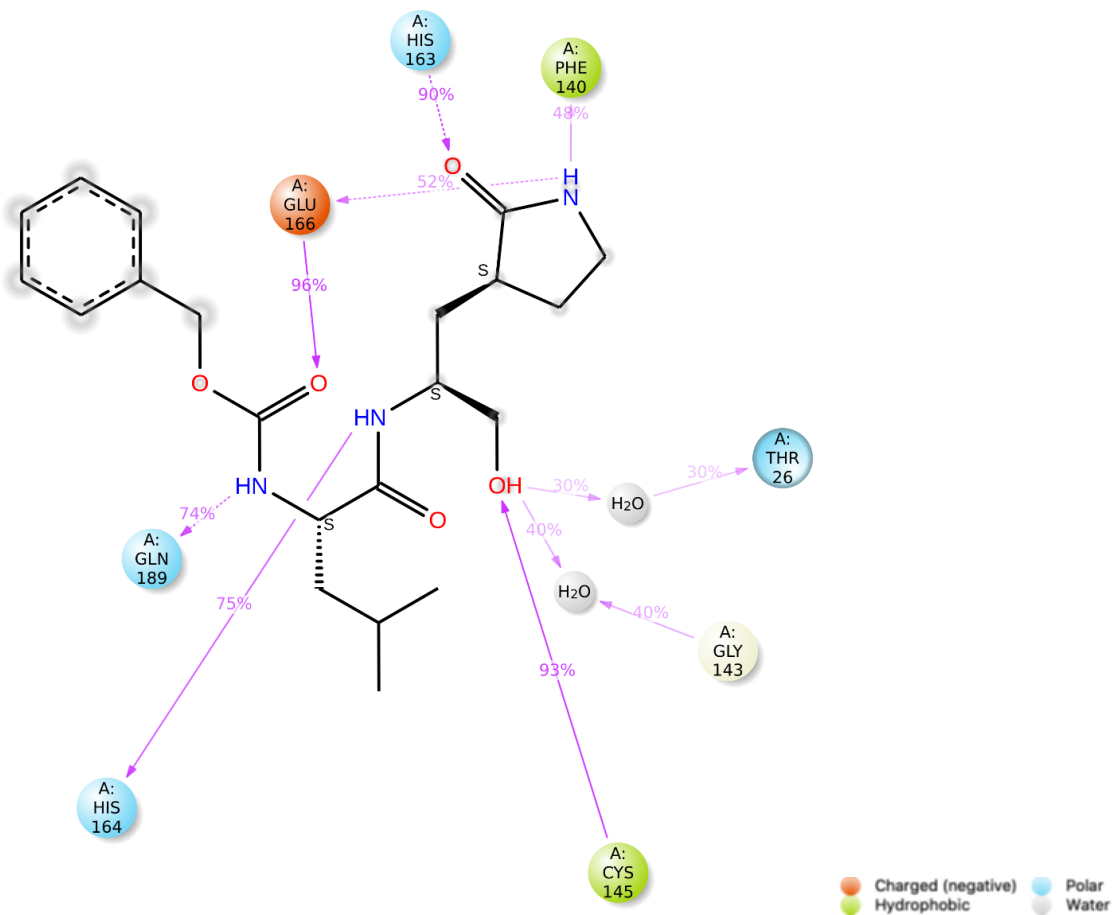


Figura 22. Simulações de DM do ligante co-cristalizado (GC376) da M^{pro} de SARS-CoV-2 (PDB ID: 7D1M). Simulações de 200 ns foram feitas em cinco replicatas. Esquema de contatos de átomos do inibidor GC376 nos resíduos da M^{pro} de SARS-CoV-2. Frequências de interação com grupos amida são observadas para Phe140, His163, His164, Glu166 e Gln189, além de interações de Thr26, Gly143 e Cys145 com um grupo hidroxila. Tipos e linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho), hidrofóbica (verde) e mediadas por água (branco). Imagens foram geradas com o programa Schrodinger.

Simulações na faixa de microsegundos auxiliam, por exemplo, na predição de reorganizações e mudanças estruturais de proteínas em interações com ligantes, enquanto múltiplas replicatas ou escalas de tempo também podem auxiliar na avaliação das diferentes formas de vibração das proteínas durante o relaxamento e recuperação do estado inicial (BRUST et al., 2013). Adicionalmente, 1 μ s dos ligantes boceprevir e GC376, das estruturas PDB ID: 7C6S (**Figura 23**) e 7C6U (**Figura 24**), respectivamente, também foram conduzidas, a fim de comparar os resultados da estrutura primária da M^{pro} de SARS-CoV-2 (7D1M) com suas “estruturas-irmãs”. Na primeira estrutura cristalográfica simulada (7C6S) foram identificados os resíduos His41 (88%), Gly143 (91%), Cys145 (95%), His164 (68%), Glu166 (89, 91 e 92%) e Gln189 (28%). O resíduo da díade catalítica His41 interage com um grupo hidroxila do ligante, enquanto Gly143, Cys145, His164, Glu166 e Gln189 se orientam a grupos amida. Aminoácidos também podem apresentar interações superiores a 100%, em razão de contatos múltiplos de um mesmo resíduo com diferentes átomos do ligante, a exemplo de Glu166 (**Figura 23**).

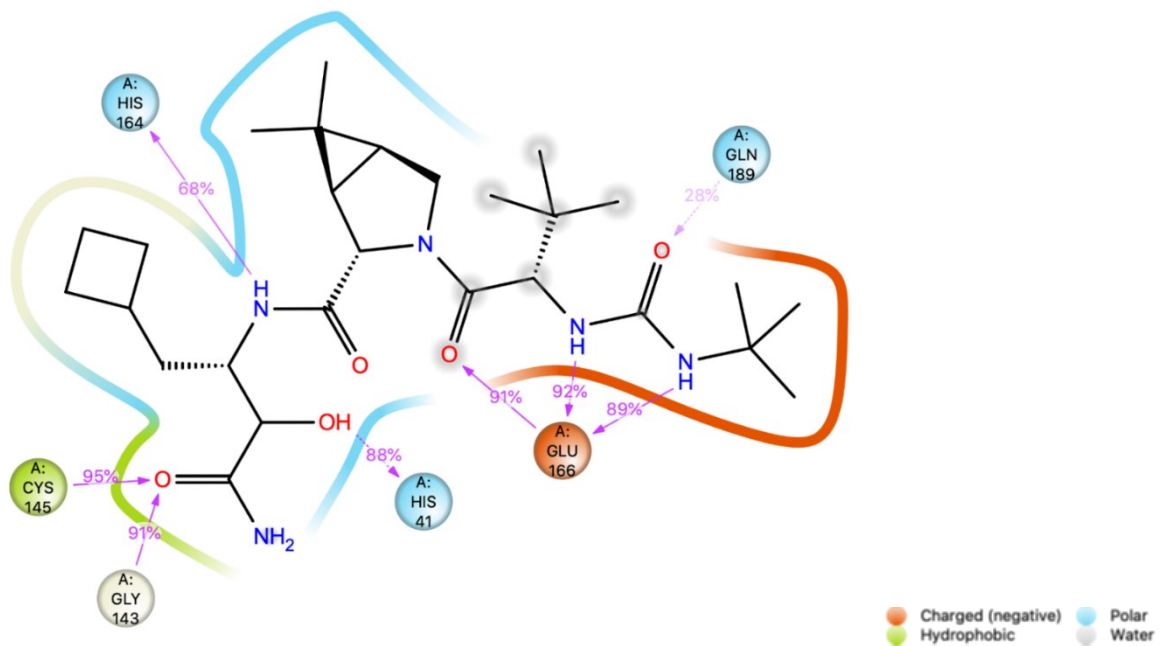


Figura 23. Simulações de DM do ligante co-cristalizado (boceprevir) da M^{pro} de SARS-CoV-2 (PDB ID: 7C6S). Simulações de 200 ns foram feitas em cinco replicatas. Esquema de contatos de átomos do ligante boceprevir nos resíduos da M^{pro} de SARS-CoV-2. Frequências de interação com grupos amida são observadas para Gly143, Cys145, His164, Glu166 e Gln189, além de interações His41 com um grupo hidroxila. Tipos e linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho), hidrofóbica (verde) e mediadas por água (branco). Imagens foram geradas com o programa Schrodinger.

Ao longo do tempo da simulação se corroboram os mesmos resíduos anteriormente preditos como potenciais sítios de interação, além de resíduos concomitantes à primeira simulação de DM, mesmo com um ligante diferente. Na última estrutura cristalográfica avaliada da M^{pro} de SARS-CoV-2 (7C6U), com inibidor GC376, resultados similares (**Figura 24**) àqueles da estrutura alvo primária (7D1M) foram obtidos. Contatos de Cys145 (93%), com um grupo hidroxila do ligante, bem como Thr26 (28%) e Gly143 (27%) mediados por uma molécula de água foram novamente observados. Os resíduos His164 (27%), Glu166 (63%) e Gln189 (42%) também se orientam a grupos amida, embora com menores frequências de interação quando comparados à estrutura primária.

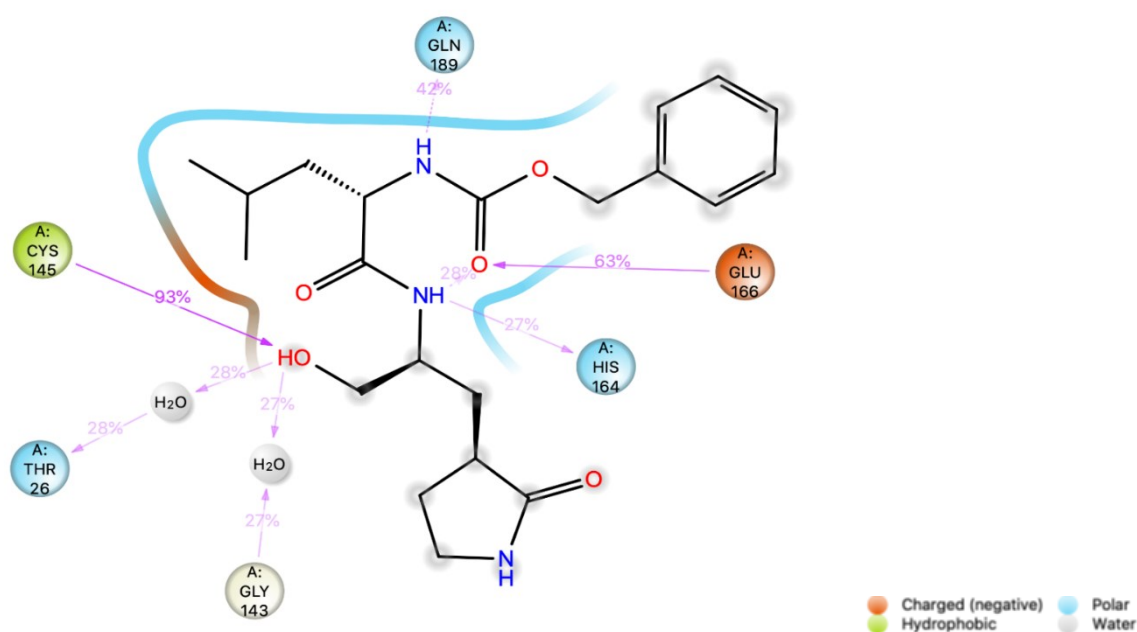


Figura 24. Simulações de DM do ligante co-cristalizado (GC376) da M^{pro} de SARS-CoV-2 (PDB ID: 7C6U). Simulações de 200 ns foram feitas em cinco replicatas. Esquema de contatos de átomos do inibidor GC376 nos resíduos da M^{pro} de SARS-CoV-2. Frequências de interação com grupos amida são observadas para Phe140, His163, His164, Glu166 e Gln189, além de interações de Thr26, Gly143 e Cys145 com um grupo hidroxila. Tipos e linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho), hidrofóbica (verde) e mediadas por água (branco). Imagens foram geradas com o programa Schrodinger.

Os resíduos Gly143, Cys145 e Glu166 foram consistentes nas interações com os ligantes co-cristalizados dentre todas as replicatas, corroborando àqueles preditos nas análises anteriores. Estes dados corroboram achados da literatura de interações com a M^{pro} de SARS-CoV-2, em que resíduos His41, Gly143, His163, Glu166 e Gln189 são frequentemente encontrados interagindo com ligantes em simulações de MD (KUMAR et al., 2020; KAPLA

et al., 2021), aquém de Cys145 em ligações covalentes, e em MD de até 1 μ s, juntamente de Phe140, Gly143, His163, His164 e Glu166 (HAYEK-ORDUZ et al., 2022), que usualmente são sugeridos em influenciar na atividade inibitória de moléculas bioativas. Especificamente Gly143 e His163 também já apresentaram maiores frequências de interações em simulações na faixa de microssegundos, entre 40 e até 70% do tempo das simulações (FERREIRA et al., 2021), similares aos dados obtidos no presente trabalho.

As análises da PL^{pro} de SARS-CoV-2, por sua vez, se basearam nos estudos de Ferreira et al., (2022), que empregaram uma metodologia similar àquela do presente trabalho, comparando quatro inibidores conhecidos desta protease, como GRL-0617, em simulações de até 10 microssegundos (FERREIRA et al., 2022). As análises demonstraram que, além dos resíduos da tríade catalítica, Cys111, His272 e Asp286, Tyr268 e Gln269 apresentaram maiores frequências de interação, de até 96% para determinados inibidores. Junto destes dados, considerando todas as análises estruturais e de interação ligante-alvo, resíduos foram selecionados como pontos de referência na construção de modelos farmacofóricos para triagem virtual de potenciais moléculas bioativas contra as duas proteases nas etapas a seguir.

5.2.5 Construção de modelos farmacofóricos da M^{pro} e PL^{pro} de SARS-CoV-2

Buscando filtrar e obter potenciais moléculas bioativas contra as duas proteases, as informações obtidas até aqui (FTSite, FTMap, PrankWeb, *redocking* e DM) foram utilizadas para construção de modelos farmacofóricos baseados nas estruturas cristalográficas da M^{pro} (PDB ID: 7D1M) e da PL^{pro} (PDB ID: 6WX4) de SARS-CoV-2. Centroides de interação molecular, ou seja, regiões ou pontos passíveis de interagir com estruturas de diferentes moléculas a serem triadas foram construídos sobre resíduos de interesse, utilizando a plataforma UNITY no programa SYBYL-X 2.1.

O modelo farmacofórico da M^{pro} considerou os dois resíduos da díade catalítica (His41 e Cys145), além de Phe140, Gly143, His163 e Glu166, dois caracterizados como doadores de ligações de hidrogênio (Gly143 e Cys145) e quatro como aceptores (His41, Phe140, His163 e Glu166) (**Figura 25**). Já os resíduos Thr26, Asn142, His164 e Gln189, além de Met49 e Met165 (potenciais hidrofóbicos) foram desconsiderados, haja visto uma certa inconsistência de interações e frequências de interação apresentadas para um determinado ligante ou estrutura nos diferentes alvos avaliados de SARS-CoV-2 e dos demais coronavírus, por exemplo, as substituições H164Q.

Arun et al., (2020), observaram resíduos Gly143 e Glu166 interagindo com ligantes durante a construção de um modelo farmacofórico da M^{pro} de SARS-CoV-2 (ARUN et al., 2020). Hayek-Orduz et al., (2022), também observaram Phe140, Gly143, Cys145, His163, Glu166 (selecionados no presente trabalho), além de His164, em interações frequentes com ligantes entre até 15 estruturas cristalográficas da protease, na construção de modelos farmacofóricos baseados em *docking* e simulações de DM (HAYEK-ORDUZ et al., 2022).

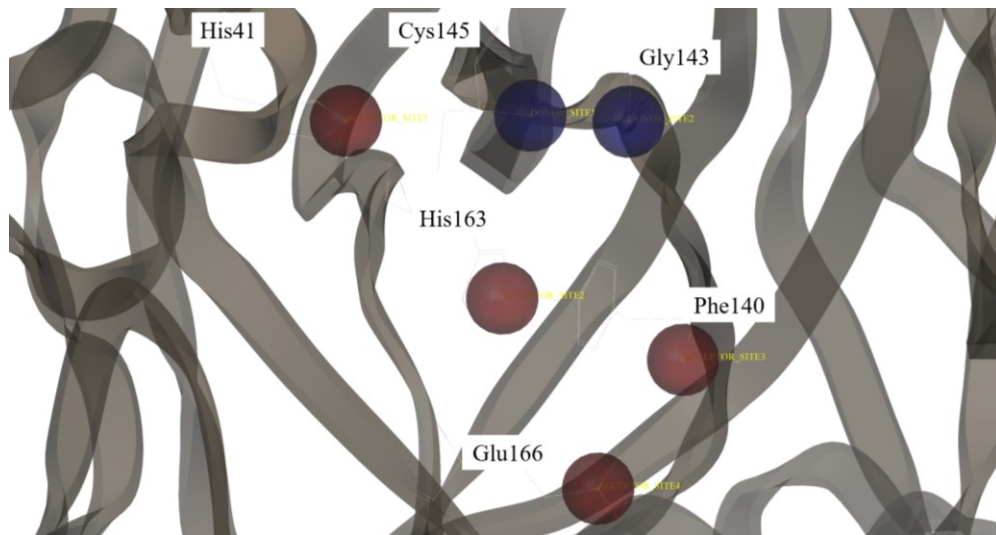


Figura 25. Modelo farmacofórico construído na M^{pro} de SARS-CoV-2 (PDB ID: 7D1M). O modelo apresenta seis centroides de interação sendo dois doadores (azul) e quatro aceptores (vermelho) de ligações de hidrogênio, assinalados aos resíduos de Gly143 e Cys145, e His41, Phe140, His164 e Glu166, respectivamente. Estrutura em modelo de *cartoon* (cinza). Imagens foram geradas com o programa SYBYL-X 2.1.

Na construção do modelo farmacofórico na PL^{pro} de SARS-CoV-2 (**Figura 26**), por sua vez, os resíduos da tríade catalítica (Cys111, His272 e Asp286) foram selecionados juntamente de Asp164, Tyr264, Gly266 e Gly271, idênticos ou similares (Tyr264) dentre todas as sequências das PL^{pro} dos diferentes coronavírus, com exceção de PDCOV (substituições D164C, G268N e G271N). Os resíduos Trp106, Gly163 e Tyr268, embora de interesse nas interações da PL^{pro} de SARS-CoV-2 (RUT et al., 2020; OSIPIUK et al., 2021), foram desconsiderados, uma vez que não foram preditos nas análises com PrankWeb (Gly163 e Tyr268) e não são conservados dentre as PL^{pro} dos demais coronavírus (Trp106). O resíduo Pro248, por sua vez, embora considerado de suporte nas interações com inibidores (FERREIRA et al., 2022), também sofre uma substituição em TGEV, sendo desconsiderado nesse modelo.

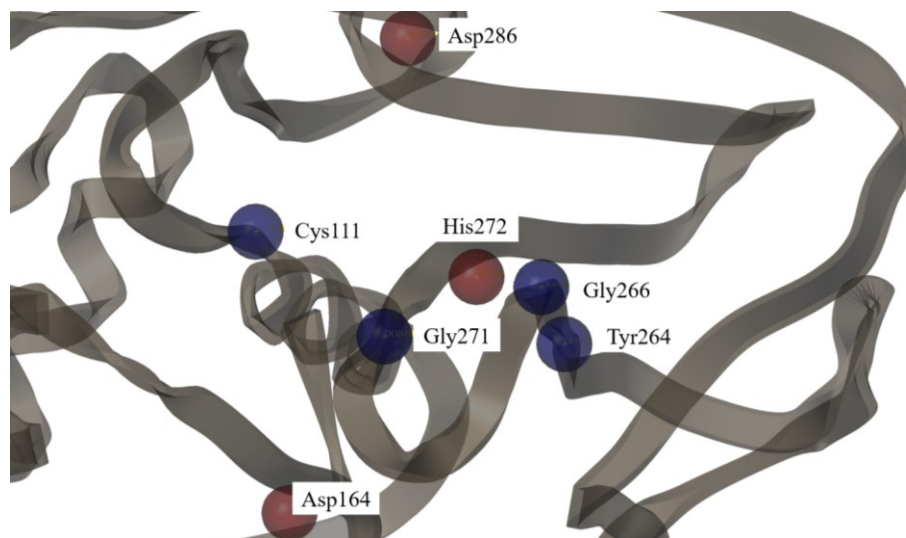


Figura 26. Modelo farmacofórico construído na PL^{pro} de SARS-CoV-2 (PDB ID: 6WX4). O modelo apresenta sete centroides de interação sendo quatro doadores (azul) e três aceptores (vermelho) de ligações de hidrogênio, sendo um destes considerado como *partial match*. Centroides foram assinalados sobre os resíduos catalíticos Cys111, His272 e Asp286, juntamente de Asp164 (*partial match*), Tyr264, Gly266 e Gly271. Estrutura em modelo de *cartoon* (cinza). Imagens foram geradas com o programa SYBYL-X 2.1.

O resíduo Asp164 foi incluído como *partial match*, ou seja, podendo ou não ser considerado em interações no modelo farmacofórico e, uma vez que interações com esse resíduo estão usualmente associadas a inibidores mais potentes (FERREIRA et al., 2022), torna-las obrigatórias poderia descartar demais inibidores na triagem virtual. Conforme mencionado (SASIDHARAN et al., 2020), triagens virtuais permitem selecionar apenas os potenciais inibidores com melhor encaixe ou interações em alvos moleculares, especificamente em uma cavidade ou sítio específico. Pensando nisso, nos modelos das proteases de SARS-CoV-2 aqui apresentados, os centroides foram restritos com uma margem de tolerância de até 1,0 Å cada, evitando descartar potenciais moléculas bioativas a serem avaliadas nas etapas posteriores.

5.2.6 Construção dos conjuntos de dados e triagens virtuais de moléculas bioativas

Construídos os modelos farmacofóricos, as estruturas 3D (construídas ou importadas) das moléculas de cada conjunto de dados (BraCoLi, DrugBank, GMR, NuBBE e ZINC15) foram submetidas à mesma preparação do HQSAR. Foram conferidas estereoquímica e geradas as conformações de maior estabilidade (estado energético mais favorável) no programa OMEGA 4.1.1. Cada molécula também teve seu estado de ionização corrigido de acordo com pH fisiológico a aplicação `fixpka` no pacote computacional QUACPAC 2.1.2.

Posteriormente, os filtros de propriedades físico-químicas com caráter “*druglike*” ou fármaco-similar foram aplicados somente ao conjunto ZINC15, buscando otimizar a triagem virtual a ser feita, reduzindo o número total de moléculas (6,8 milhões). Inicialmente, modificações dos parâmetros baseados nas regras de Lipinski (LIPINSKI et al., 1997) e Veber (VEBER et al., 2002) foram empregados (**Tabela 2; tópico 4.1.5**) buscando aprimorar o filtro inicial, por exemplo, uma pequena margem de tolerância de 10% ajustada para os parâmetros de massa molecular (550 Da) e de log P (5,5), comparado com a regra de Lipinski (até 500 Da e 5,0, respectivamente).

Já comparado às regras de Veber, pouco mais do dobro da superfície polar pode ser considerado (300 \AA^2), em razão do possível alto caráter polar de moléculas a serem triadas. Os átomos permitidos, por sua vez, foram limitados a carbono, oxigênio, nitrogênio, hidrogênio, fósforo, enxofre, flúor, cloro, bromo e iodo, em razão do campo de força do ligante no *docking*, e considerando maximizar a viabilidade sintética de potenciais candidatos obtidos. Dois centros quirais também foram delimitados, buscando reduzir a possibilidade de enantiômeros e facilitar uma eventual viabilidade de síntese de moléculas. Já a carga total (-2 a +2) se baseou em uma maior disponibilidade de cargas positivas para interagir com átomos no alvo, permitindo, por exemplo, interações iônicas e de hidrogênio. Contudo, o número não pode ser elevado o bastante, haja visto que o estado de ionização influencia diretamente na solubilidade e, conseqüentemente, na permeabilidade em membranas celulares (TEN BRINK; EXNER, 2009; H. BROOKS; C. GUIDA; G. DANIEL, 2011).

Finalmente, o filtro de solubilidade mínima das moléculas foi delimitado como baixo, limitado a DMSO) até completamente solúveis, em razão da capacidade deste solvente em dissolver moléculas com diferentes propriedades físico-químicas, apresentar uma baixa volatilidade comparada a demais solventes, miscibilidade com água e outros solventes, além da possibilidade de congelamento em nitrogênio líquido (-196 °C), sendo uma opção mais favorável à manutenção da estabilidade de moléculas em uma solução (WAYBRIGHT; BRITT; MCCLOUD, 2009).

Dentre as mais de 6,8 milhões de moléculas triadas dos cinco conjuntos de dados, 397 foram obtidas do ZINC15 (Molport), 35 do GMR e três do DrugBank para M^{pro}, além de 107 do GMR para PL^{pro} (**Figura 27**). Nenhuma molécula nas quimiotecas NuBBE e BraCoLi foram obtidas ao final das triagens pelos farmacóforos. Devido à indisponibilidade de recursos para aquisição de moléculas e/ou fármacos, bem como da inviabilidade de síntese

dos *hits* do MolPort e DrugBank para M^{pro}, a triagem destes bancos de dados foram interrompidas nesta etapa e, conseqüentemente, não foram executadas para PL^{pro}.

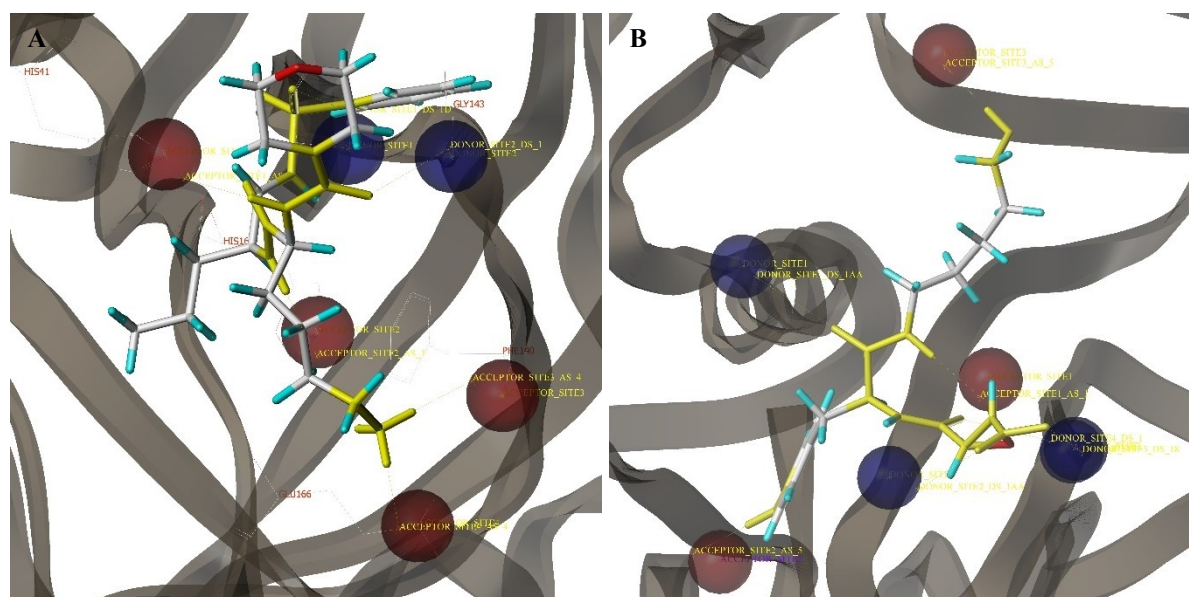


Figura 27. *Hits* das triagens virtuais contra as proteases de SARS-CoV-2. Exemplos de moléculas (*hits*) obtidos nas triagens pelos modelos farmacofóricos (cinza) das (A) M^{pro} e (B) PL^{pro} de SARS-CoV-2. Poses apresentando interações previstas com os centroides de cada farmacóforo (amarelo). Centroides do tipo aceptores de ligação de hidrogênio (vermelho) e doadores de hidrogênio (azul) são destacados.

Selecionados os *hits* dos modelos com as duas proteases, o *docking* das moléculas com Surflex-Dock foi feito sob as mesmas condições anteriores (*redocking*) e, aqueles mais bem ranqueados pela pontuação score ($> 8,5$), foram selecionadas para uma nova análise de *docking* no programa GOLD, também sob os mesmos parâmetros anteriores. Foi feita uma inspeção visual dos resultados deste *crossdocking*, ou seja, análises de interação de acoplamento em dois ou mais programas de *docking* distintos, em que moléculas com poses similares foram selecionadas para análises posteriores (JIMÉNEZ-AVALOS et al., 2021), como simulações de DM e avaliações da atividade inibitória enzimática *in vitro*.

Finalmente, 18 moléculas da quimioteca GMR, todas disponíveis no campus da UCSD, foram selecionadas como potenciais inibidores para M^{pro} (*scores* entre 8,891 e 10,91) e 25 para PL^{pro} (*scores* entre 8,56 e 10,55), totalizando 43 estruturas. Algumas estruturas foram omitidas em razão do termo de confidencialidade da parceria entre UCSD e UFMG, até a publicação dos manuscritos vinculados ao presente trabalho (em preparação). Contudo, dentre as moléculas selecionadas, apenas 14 (32,56%) se encontravam fisicamente disponíveis para as simulações de DM e os ensaios de inibição da atividade enzimática subsequentes (**Tabela 14**).

Tabela 14. Moléculas selecionadas após o protocolo de triagem virtual contra as proteases

#VS	Cód.	MM	Alvo	Situação	#VS	Cód.	MM	Alvo	Situação
338	6429	330,38	PL ^{pro}	Ausente	255	6512	483,57	M ^{pro}	WRR-305
292	6475	350,37	PL ^{pro}	Vazio	259	6508	498,58	PL ^{pro}	WRR-301
286	6481	366,37	M ^{pro}	Vazio	258	6509	489,53	PL ^{pro}	WRR-302
341	6426	358,44	PL ^{pro}	Ausente	1537	0001	545,63	M ^{pro}	Ausente
295	6472	364,4	PL ^{pro}	Ausente	145	6620	511,62	M ^{pro}	WRR-478
362	6405	363,41	M ^{pro}	WRR-193	136	6629	564,71	M ^{pro}	Ausente
335	6432	372,46	PL ^{pro}	Ausente	135	6630	556,72	PL ^{pro}	WRR-500
287	6480	394,42	PL ^{pro}	Ausente	165	6600	556,64	PL ^{pro}	Vazio
360	6407	377,44	M ^{pro}	WRR-195	1244	4035	532,63	PL ^{pro}	Ausente
423	6344	392,45	PL ^{pro}	Ausente	164	6601	563,67	PL ^{pro}	WRR-453
285	6482	408,45	M ^{pro}	Ausente	1256	3927	555,63	PL ^{pro}	Ausente
363	6404	391,47	M ^{pro}	Ausente	147	6618	539,67	M ^{pro}	Vazio
289	6478	406,48	PL ^{pro}	Ausente	1344	3237	560,71	PL ^{pro}	Vazio
361	6406	405,5	M ^{pro}	Ausente	212	6554	576,71	M ^{pro}	WRR-390
1217	4069	836,47	M ^{pro}	Ausente	226	6541	609,71	PL ^{pro}	WRR-366
310	6457	420,51	PL ^{pro}	Ausente	274	6493	590,74	PL ^{pro}	Ausente
419	6348	447,53	PL ^{pro}	Ausente	221	6545	577,74	PL ^{pro}	WRR-377
125	6640	480,62	M ^{pro}	Ausente	163	6602	584,69	PL ^{pro}	WRR-454
1254	3929	507,59	M ^{pro}	Ausente	250	6517	620,77	M ^{pro}	WRR-310
263	6504	468,55	M ^{pro}	WRR-297	162	6603	633,76	PL ^{pro}	Vazio
1245	4034	465,51	M ^{pro}	Ausente	324	6443	681,85	PL ^{pro}	Ausente
-	-	-	-	-	293	6474	378,43	PL ^{pro}	Ausente

#VS: número de origem na triagem virtual (VS); Cód.: código; MM: Massa Molecular.

Portanto, um conjunto final (**Figura 28**) compreendendo 14 moléculas, sendo sete contra M^{pro} e sete contra PL^{pro}, todas oriundas da série denominada “WRR” (William R. Roush), foi selecionado para as análises *in vitro*. O reduzido número de moléculas selecionadas fez por tornar não prioritárias as análises comparativas entre os programas Surflex-Dock e GOLD, bem como das simulações de DM, poupando tempo ao longo dessa etapa da triagem.

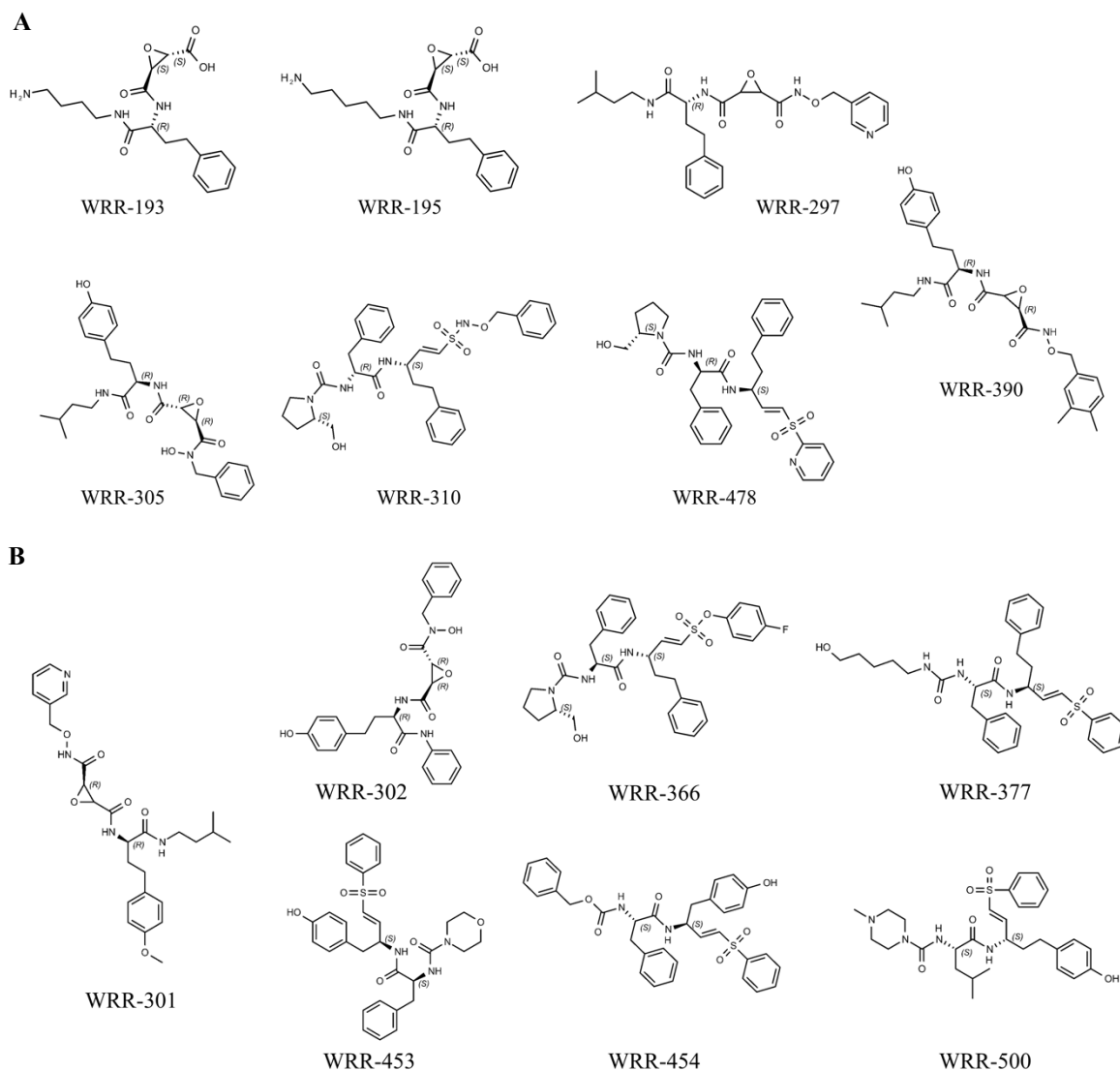


Figura 28. Conjunto de moléculas selecionadas contra as proteases de SARS-CoV-2. (A) Hits selecionados na triagem virtual contra M^{P₁}: WRR-193, WRR-195, WRR-297, WRR-305, WRR-310, WRR-390 e WRR-478. (B) Hits selecionados na triagem virtual contra PL^{P₁}: WRR-301, WRR-302, WRR-366, WRR-377, WRR-453, WRR-454 e WRR-500. Estruturas representadas em 2D. Imagens foram geradas no programa ChemDraw.

Finalmente, visando aumentar o número de moléculas a serem avaliadas, 127 análogos, intermediários e derivados de síntese dos *hits* obtidos da série WRR foram incluídos a fim de expandir o conjunto a ser testado nos ensaios contra as proteases. Assim, inferências sobre a relação entre as estruturas químicas e a atividade biológica também poderiam ser feitas (p.ex.: HQSAR) nesta estratégia.

5.3 Ensaios biológicos

5.3.1 Avaliação da inibição da atividade enzimática

Dentre as 14 moléculas selecionadas das triagens virtuais, no entanto, nenhuma daquelas testadas apresentou inibição superior a 50% quando avaliadas a 10 μM . Contudo, dentre os 127 análogos avaliados, cinco moléculas (WRR-180, 181, 317, 385 e 438) inibiram pelo menos 50% da atividade enzimática da M^{pro} de SARS-CoV-2 a 10 μM (entre $51,55 \pm 1,38 \%$ e $100,00 \pm 0,00 \%$). Um segundo ensaio também foi feito a 5 μM , em que apenas WRR-180 ($98,70 \pm 1,04 \%$) e WRR-181 ($100,00 \pm 0,0 \%$) mantiveram a atividade inibitória inicial acima de 50% (**Apêndice A**). Curiosamente, a molécula WRR-177 apresentou um aumento da atividade inibitória a 5 μM (de $7,45 \pm 0,6 \%$ para $34,24 \pm 2,9 \%$). Nos ensaios com a PL^{pro} , por sua vez, com exceção do controle GRL-0617 a 10 μM ($72,3 \pm 1,03 \%$), nenhuma das moléculas inibiu ao menos 50% da atividade enzimática (dados não apresentados). Nesta avaliação, somente WRR-315 ($38,07 \pm 1,49 \%$) apresentou alguma atividade superior a 20% contra PL^{pro} de SARS-CoV-2.

As cinco moléculas com atividade inibitória a 10 μM ($\geq 50\%$ de inibição) foram então avaliadas em ensaios para determinação da curva de dose-resposta, obtendo os valores da concentração inibitória de 50% (IC_{50}). Dentre as moléculas, WRR-180 e WRR-181 apresentaram valores de IC_{50} na faixa de nanomolar, $0,415 \pm 0,018$ e $0,086 \pm 0,0015 \mu\text{M}$, respectivamente. Adicionalmente, WRR-317, WRR-385 e WRR-438 apresentaram valores de IC_{50} na faixa de micromolar, sendo $8,12 \pm 0,043$, $16,42 \pm 0,095$ e $8,72 \pm 0,54 \mu\text{M}$, respectivamente (**Tabela 15**).

A adição de um grupamento éster após o substituinte eteno na cadeia terminal de WRR-180 aumentou cerca de cinco vezes os valores de IC_{50} quando comparado ao seu análogo, WRR-181, embora ainda mantendo de atividade inibitória da M^{pro} de SARS-CoV-2 na faixa de nanomolar. Similarmente, a adição de um éster e isopentano, bem como a substituição do éter em WRR-385 aumentou em quase duas vezes os valores da IC_{50} comparado a seu análogo WRR-438, que também apresentou valores de IC_{50} similares a seu análogo WRR-317 com a substituição de uma hidroxietil-ciclopropanona para um sulfonilbenzeno metano. O controle nirmatrelvir foi avaliado em uma concentração fixa de 0,1 μM , inibindo a atividade enzimática em pelo menos 90% em todos os ensaios, o que corrobora valores de IC_{50} na literatura, entre 3 e 4 nM (ANTONOPOULOU et al., 2022; DUVEAU; THOMAS, 2022).

Embora a atividade de nirmatrelvir contra as diferentes M^{pro} dos três coronavírus já tenha sido descrita, inclusive para os demais coronavírus que infectam seres humanos (CHEN et al., 2022), a possibilidade resistência a este fármaco também já foi evidenciada em diferentes estudos (ABDELNABI et al., 2023; JOCHMANS et al., 2023), o que denota a importância da contínua busca de novos inibidores com atividade contra essas proteases (IKETANI et al., 2022). A resistência está associada, por exemplo, a mutações de resíduos usualmente associados a sítios de ligação ou ao sítio catalítico, como Phe140 e His172 (DE OLIVEIRA et al., 2022), Ser144, Met165, Glu166 e His172 (HU et al., 2022), Leu50, Glu166 e Leu167 (JOCHMANS et al., 2023), bem como Leu50 e Glu166, com validações em ensaios celulares com células Vero e A549-ACE2 (linhagem pulmonar humana expressando ACE2) (ZHOU et al., 2022), e somente em Glu166, aumentando em até 100 vezes a IC₅₀ contra SARS-CoV-2 recombinante em células Huh7 (linhagem de carcinoma hepático humano) (IKETANI et al., 2022).

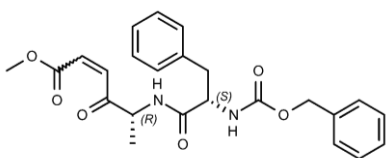
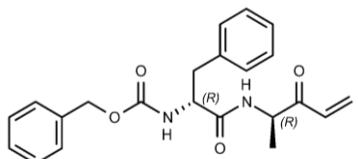
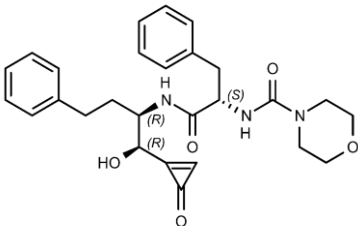
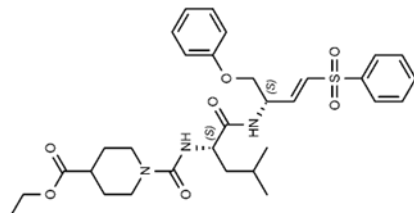
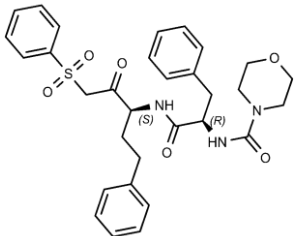
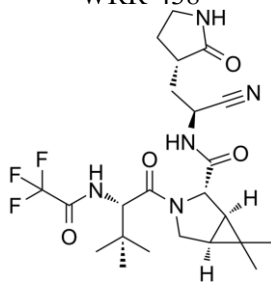
Considerando a consonância dos resíduos (p.ex.: Glu166) entre as M^{pro} de diferentes coronavírus, as proteases de SARS-CoV e MERS-CoV também foram incluídas nos ensaios de inibição da atividade enzimática. Dentre as cinco moléculas com atividade contra M^{pro} de SARS-CoV-2, duas apresentaram atividade inibitória a 10 µM ($\geq 50\%$ de inibição) contra protease de SARS-CoV e tiveram sua IC₅₀ determinada para WRR-180 e WRR-181, ambas também na faixa de nanomolar, $0,7 \pm 0,0125$ e $0,11 \pm 0,0175$ µM, respectivamente. Igualmente, também foi observada atividade inibitória contra M^{pro} de MERS-CoV, com valores de IC₅₀ de $3,786 \pm 0,64$ e $0,83 \pm 0,0535$ µM determinados para WRR-180 e WRR-181, respectivamente. Assim como observado para M^{pro} de SARS-CoV-2, a adição de um grupamento éster após o substituinte eteno na cadeia terminal de WRR-180 também aumentou os valores de IC₅₀ em pelo menos quatro vezes comparado ao seu análogo WRR-181, mantendo a faixa de micromolar. Ademais, WRR-317, WRR-385 e WRR-438 não foram capazes de inibir pelo menos 50% a atividade das enzimas a 10 µM (**Tabela 15**).

Ghosh et al., (2021), também demonstraram o efeito de substituintes éster em pequenas moléculas na manutenção da atividade inibitória contra M^{pro} de SARS-CoV-2, com valores de IC₅₀ de 0,12 µM para molécula denominada **7b**, mas reduzindo a atividade antiviral contra SARS-CoV-2 em ensaios celulares com células Vero, com valores de EC₅₀ de 8 µM. Adicionalmente, os autores também avaliaram a molécula contra M^{pro} SARS-CoV e, assim como evidenciado no presente trabalho, valores mais altos de IC₅₀ (0,9 µM) também foram

obtidos comparados a SARS-CoV-2. Neste estudo, análises de *docking* também foram feitas para **7b**, corroborando interações, por exemplo, com His41 e Cys145 (GHOSH et al., 2021).

Pillaiyar et al., (2022), também empregaram uma abordagem que compreendia a combinação de *docking*, MD, triagem virtual e ensaios enzimáticos para o planejamento e obtenção de inibidores contra as M^{pro} de MERS-CoV, SARS-CoV e SARS-CoV-2. Os autores obtiveram, após etapas de triagem virtual e modificações sintéticas, inibidores contra M^{pro} de SARS-CoV-2 na faixa de nanomolar, a exemplo das moléculas denominadas **3w** e **3x**, com valores de IC₅₀ de 0.0114 ± 0.0028 e 0.0876 ± 0.0109 μM , respectivamente. As mesmas moléculas também apresentaram atividade contra as M^{pro} de SARS-CoV e MERS-CoV, com valores de IC₅₀ de 0.06 ± 0.00 e 0.303 ± 0.059 μM para **3w**, respectivamente, e de 1.52 ± 0.01 e 1.31 ± 0.28 μM para **3x** (PILLAIYAR et al., 2022). Os autores planejaram e avaliaram pequenas moléculas análogos, derivados e intermediários de síntese de tioésteres, estruturas com menor massa molecular quando comparado às moléculas obtidas e avaliadas no presente trabalho (**Tabela 15**).

Tabela 15. Atividade inibitória (μM) contra as M^{pro} de SARS-CoV-2, SARS-CoV e MERS-CoV

Molécula	MM (g/mol)	SARS-CoV-2	SARS-CoV	MERS-CoV
 WRR-180	438,5	0,415 \pm 0,018	0,7 \pm 0,0125	3,786 \pm 0,64
 WRR-181	380,4	0,086 \pm 0,0015	0,11 \pm 0,0175	0,83 \pm 0,0535
 WRR-317	477,6	8,12 \pm 0,043	20,46 \pm 0,69 ^a (10 μM)	0,0 \pm 0,0 ^a (10 μM)
 WRR-385	599,7	16,42 \pm 0,095	0,00 \pm 0,00 ^a (10 μM)	16,05 \pm 2,53 ^a (10 μM)
 WRR-438	577,7	8,72 \pm 0,54	21,28 \pm 0,62 ^a (10 μM)	0,00 \pm 0,00 ^a (10 μM)
 Nirmatrelvir	499,53	98,08 \pm 1,21 ^a (0,1 μM)	95,37 \pm 4,63 ^a (0,1 μM)	53,11 \pm 4,65 ^a (0,1 μM)

^a O percentual de inibição é representado como a média e o erro do desvio-padrão calculados entre pelo menos dois ensaios independentes em triplicata. O erro é apresentado como a razão entre o desvio-padrão e a raiz quadrada do número de replicatas/leituras ($n \geq 6$).

Goetz et al., (2007), avaliaram moléculas da série WRR sintetizadas contra a M^{pro} de SARS-CoV (GOETZ et al., 2007), presentes em uma quimioteca com cerca de 2.000 de inibidores de cisteína proteases da University of California San Francisco (UCSF) do *Sandler Center for Basic Research in Parasitic Diseases* (dados não publicados), não relacionadas àquelas avaliada no presente trabalho (GMR), bem como dos análogos testados (WRR), com exceção de WRR-182 e WRR-183. Um inibidor irreversível (α,β -epoxicetona) foi obtido, corroborando a preferência por resíduos Gln na posição P1 de ligantes e substratos dessa protease, sendo direcionado para síntese de 12 análogos, resultando em moléculas da série WRR (WRR-182, 183, 210, 211, 485-488, 492, 493, 495 e 496) (GOETZ et al., 2007).

As 12 moléculas foram avaliadas em ensaios de inibição da atividade enzimática da M^{pro} de SARS-CoV, sendo oito ativas (WRR-182, 183, 210, 211, 486, 492, 495 e 496), com valores da razão entre a taxa de inibição (k_3) e a constante de inibição (k_i) entre 0,0001 e 0,5 $\mu\text{M/s}$. As moléculas WRR-183 e WRR-495 também foram avaliadas em ensaios de citotoxicidade e atividade antiviral por uma técnica de luminescência, obtendo valores de EC₅₀ de 12 e 50 μM , respectivamente, com IS de 5,04 e 0,4 (GOETZ et al., 2007).

As moléculas WRR-182, 183, 210, 211 e 496, todas ativas contra M^{pro} de SARS-CoV, apresentam similaridade estrutural com WRR-180 e WRR-181, diferenciando-se apenas na configuração dos respectivos estereocentros, ou seja, a orientação R ou S dos grupos substituintes e, especificamente, na presença de um substituinte epóxi no lugar dos substituintes eteno de WRR-180 e WRR-181. Ademais, WRR-182 e WRR-183 não apresentaram atividade inibitória (> 50%) frente à M^{pro} de SARS-CoV-2 até 5 μM . Essas diferenças podem estar associadas ao perfil de interações observado com os resíduos Met49, Phe140, Leu144, His163, Glu166, Met165 e His172 em análises de *docking* para WRR-183 (GOETZ et al., 2007), diferentemente daqueles observados para WRR-180 e WRR-181 discutidos a seguir no presente trabalho (**Figura 31**), interagindo principalmente com Asn142, Gly143 e Cys145 (WRR-180) e His41 (WRR-181), ao longo de simulações de DM.

As curvas de dose-reposta das moléculas, por sua vez, apresentaram um perfil similar dentre as atividades inibitórias avaliadas, especificamente entre WRR-180 e WRR-181 (**Figura 29**). Especificamente para o controle nirmatrelvir, a curva não foi feita, em razão dos resultados consolidados e corroborados na literatura, por exemplo, naquelas com valores de IC₅₀ de 4 nM (DUVEAU; THOMAS, 2022) e de 3,11 nM (ANTONOPOULOU et al., 2022).

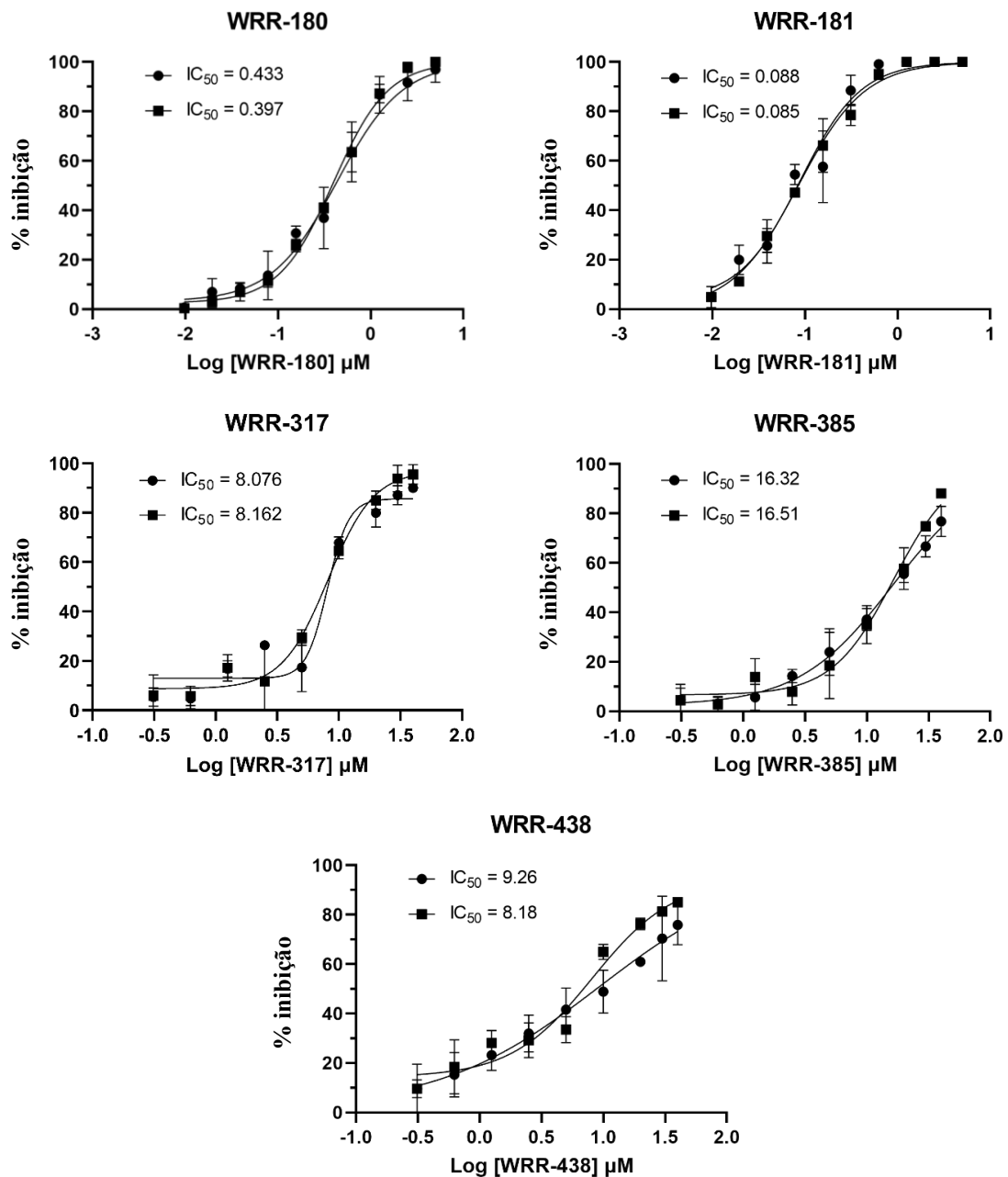


Figura 29. Curvas de dose-resposta dos inibidores contra M^{pro} de SARS-CoV-2. Valores de IC_{50} de cada replicata foram calculados a partir da regressão não linear e do desvio-padrão de cada ponto na determinação das curvas, sendo considerados os menores que 10% ($r^2 > 0.9$). Ao menos nove concentrações foram avaliadas entre 0,01 e 40 μM . Os dados foram analisados com GraphPad Prism 9.0 (GraphPad Software, San Diego, Califórnia, EUA). Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

Igualmente, curvas de dose-reposta das moléculas WRR-180 e WRR-181 determinadas para as M^{pro} de SARS-CoV e MERS-CoV, por sua vez, também apresentaram um perfil similar às curvas de nirmatrelvir (**Figura 30**). Esse perfil poderia sugerir o mesmo comportamento reversível de nirmatrelvir nestas moléculas, especificamente considerando os substituintes eteno atuando como um grupo funcional eletrofílico (HUANG et al., 2022).

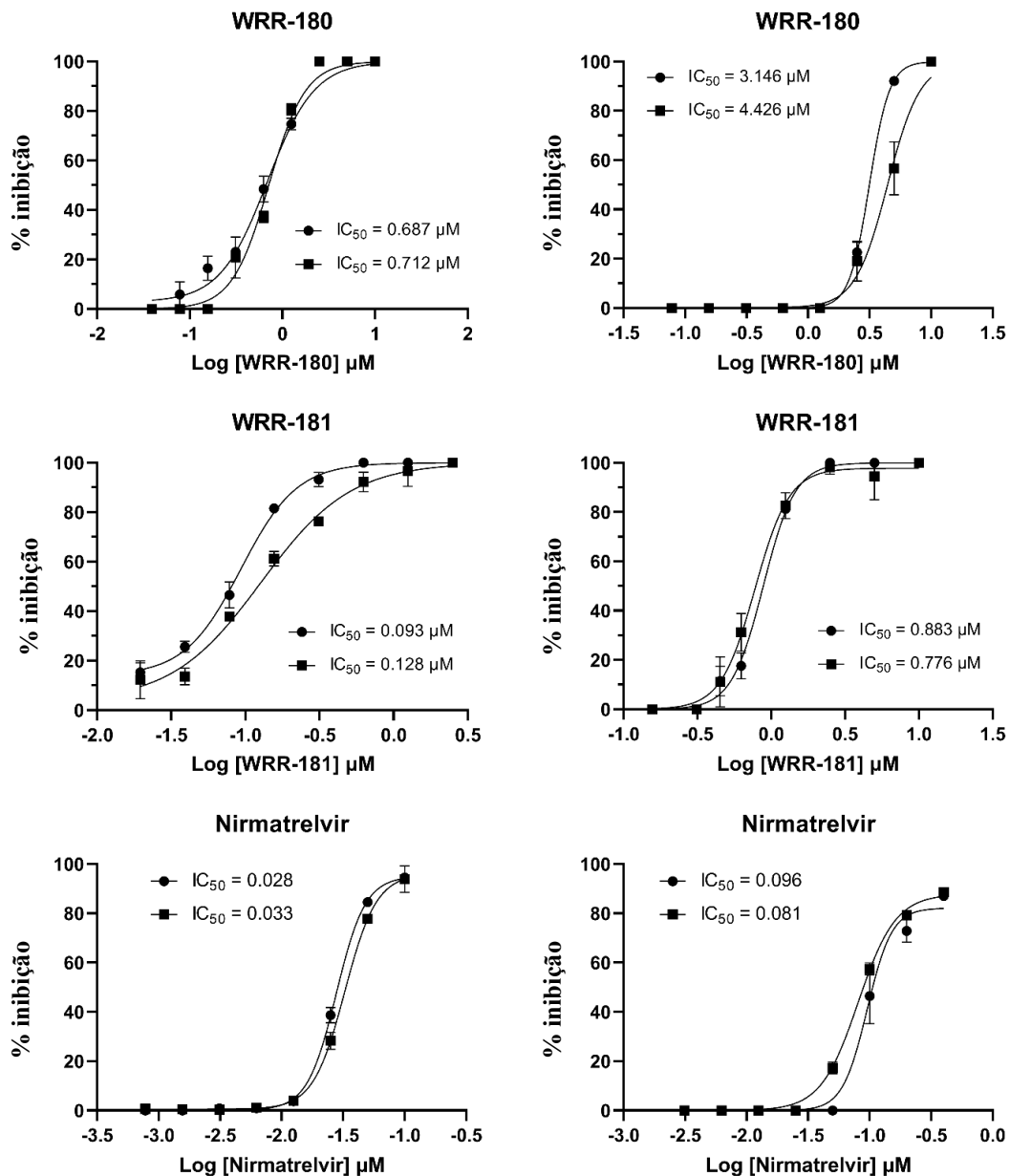


Figura 30. Curvas de dose-resposta de WRR-180 e WRR-181 contra as M^{pro} de SARS-CoV e MERS-CoV. Valores de IC_{50} de cada replicata para SARS-CoV (coluna esquerda) e MERS-CoV (coluna direita) foram calculados a partir da regressão não linear e do desvio-padrão de cada ponto nas curvas, sendo considerados os menores que 10% ($r^2 > 0.9$). Ao menos oito concentrações foram avaliadas entre 0,02 e 10 µM. Os dados foram analisados com GraphPad Prism 9.0 (GraphPad Software, San Diego, Califórnia, EUA). Todas as condições foram avaliadas em triplicata e em pelo menos dois ensaios independentes ($n \geq 6$).

Naturalmente, a natureza da inibição (reversível ou irreversível) poderia ser avaliada mediante ensaios e análises do mecanismo de ação de cada inibidor, especificamente quando a taxa de formação de uma ligação covalente é relativamente lenta e a constante de equilíbrio da reação “k” sofre influência da velocidade da reação, ou seja, considerando a dissociação do complexo inibidor-proteína (HUANG et al., 2022). Mellot et al., (2021), por exemplo,

empregaram ensaios de diluições por salto na caracterização do inibidor **K777** (candidato avaliado em ensaios clínicos), um inibidor irreversível da catepsina L de SARS-CoV-2 (MELLOTT et al., 2021).

Pillaiyar et al., (2022), também sugerem que especificamente para pequenas moléculas atuando como inibidores irreversíveis, a eficácia é frequentemente determinada pela força do grupo funcional eletrofílico, independente da ocupação dos demais sítios nas interações ligante-alvo (PILLAIYAR et al., 2022). Esse perfil pode corresponder às diferenças de potência entre o inibidor obtido no presente trabalho, WRR-181 ($IC_{50} = 0,086 \pm 0,0015 \mu M$) e nirmatrelvir ($IC_{50} = 0,003 \mu M$), dado às interações covalentes reversíveis com Cys145 dos respectivos grupos eteno e cianeto (nitrila), respectivamente. Além das diferenças associadas aos grupos eletrofílicos, as variações nas frequências de interações (p.ex.: simulações de DM) de determinados resíduos também podem corroborar a respeito da potência de determinados inibidores da M^{pro} de SARS-CoV-2, a exemplo dos resíduos His41, Cys145, His163, Glu166 e Gln189, como sugerido anteriormente (FERREIRA et al., 2021).

Nesse sentido, também foram feitas simulações de DM de 1 μs (5 x 200 ns) para avaliação do perfil de interações de WRR-180 e WRR-181 na M^{pro} de SARS-CoV-2 (**Figura 31**), a fim de se comparar dados existentes de interações com nirmatrelvir, por exemplo. Os resultados corroboram aos achados de simulações previamente discutidos, a exemplo das interações com o ligante co-cristalizado GC376 e de boceprevir nas M^{pro} de SARS-CoV-2 (**Figura 22 e 23**). Nas DM de WRR-180 foram evidenciados contatos de Cys145 (63%) com um grupo cetona do ligante, bem como Asn142 (51%) e Gly143 (57%) se orientando a grupos amidas, além de Glu166 (40%) com um éster. Já para WRR-181, somente contatos com His41 (33 e 41%) foram observados.

Além das análises anteriores, esses dados também são similares a simulações com nirmatrelvir, que demonstraram contatos covalentes com Cys145, além de interações com Glu166 (DE OLIVEIRA et al., 2022), bem como de interações com Gly143, Met165 e Glu166 em simulações de um modelo farmacofórico (HAYEK-ORDUZ et al., 2022). Adicionalmente, em um farmacóforo, a quantidade de centroides e, estes podendo ou não ser considerados em interações (*partial match*), poderiam restringir ou desconsiderar possíveis inibidores que interagiriam somente com um ou dois resíduos no alvo, como predito nas simulações de DM para WRR-180 (**Figura 31**).

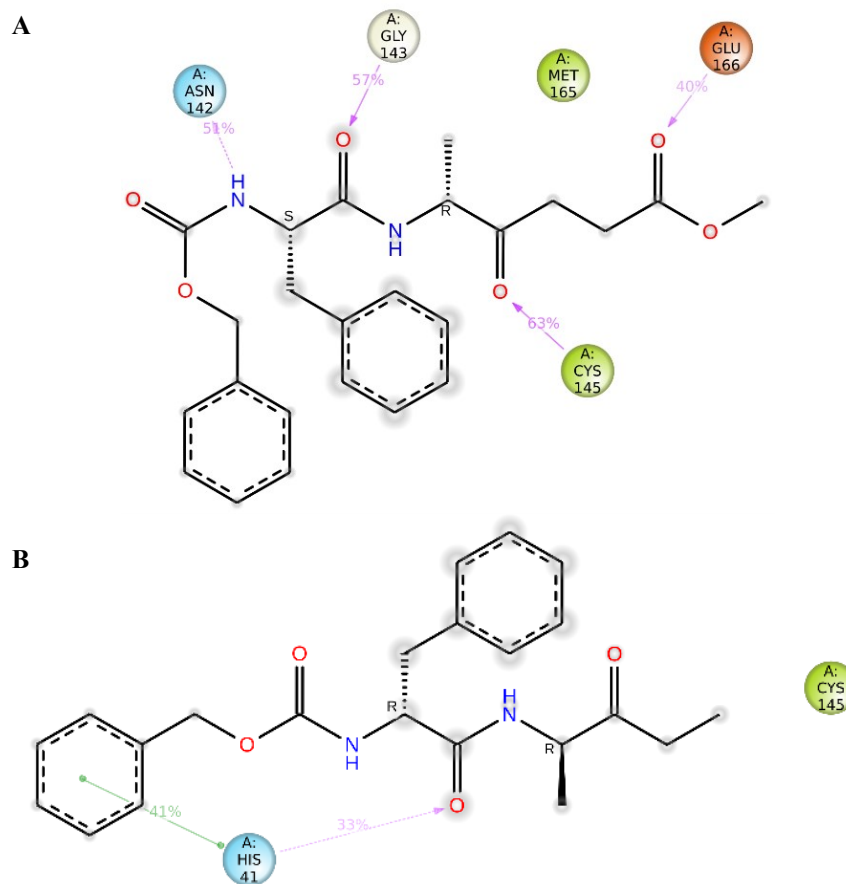


Figura 31. Simulações de DM de WRR-180 e WRR-181 na M^{pro} de SARS-CoV-2 (PDB ID: 7D1M). Simulações de 200 ns foram feitas em cinco replicatas. Esquema de contatos de átomos de (A) WRR-180 e (B) WRR-181 nos resíduos da M^{pro} de SARS-CoV-2. Frequências de interação de WRR-180 com grupos amida observadas para Asn142 e Gly143, além de Cys145 com grupo cetona e Glu166 com éster. Já WRR-181 interage somente com His41 em dois contatos, com amida e benzeno terminal. Tipos e linhas de tendência de interação: ligação de hidrogênio ou polar (azul), iônica ou carga negativa (vermelho), hidrofóbica (verde) e mediadas por água (branco). Imagens foram geradas com o programa Schrodinger.

Esse perfil de interações com determinados resíduos das M^{pro} de SARS-CoV-2, bem como de outros coronavírus, também reforçam a influência de determinados resíduos na atividade de inibidores. Greasley et al., (2022), demonstraram, por exemplo, a conservação de interações de nirmatrelvir com Cys145 e Glu166 entre as M^{pro} de diferentes linhagens de SARS-CoV-2, mantendo também uma equivalência na potência da atividade inibitória nos ensaios de inibição da atividade enzimática (GREASLEY et al., 2022). Nesse sentido, uma alusão ao perfil de interações dos ligantes e inibidores obtidos e avaliados no presente trabalho, a exemplo de WRR-180, também foi similar àquelas observadas para outros coronavírus, como SARS-CoV e MERS-CoV, especificamente sobre ligações covalentes com Cys145, além de interações com Glu166 (LI et al., 2022).

Tomados em conjunto, os dados até aqui apresentados também reforçam a importância de uma abordagem ampla no planejamento de inibidores contra alvos moleculares de diferentes organismos, como foi a busca de moléculas bioativas contra M^{pro} de SARS-CoV-2 e demais coronavírus. A identificação de dois inibidores com atividade pancoronavírus contra as M^{pro} de SARS-CoV-2, SARS-CoV e MERS-CoV abordada no presente trabalho, por sua vez, reflete o sucesso de uma estratégia que combinou diferentes métodos computacionais para triar e avaliar ligantes contra os alvos propostos. Estes, no entanto, seguirão para ensaios biológicos na avaliação de citotoxicidade e de atividade antiviral, ambos em ensaios celulares, conforme se almeja para uma validação, por exemplo, de campanhas computacionais que obtiveram êxito em análises e ensaios preliminares (AZEVEDO et al., 2022).

Pensando nessa abordagem e, sob aguardo do envio das duas moléculas apresentadas no presente trabalho (WRR-180 e WRR-181), uma abordagem experimental *in vitro* complementar foi feita, a fim de se identificar potenciais moléculas bioativas contra um modelo CR-2 de coronavírus murino (MHV-3). Essa triagem biológica, apresentada no tópico a seguir, remete às etapas subsequentes que serão feitas na parceria estabelecida entre UCSD e UFMG, avaliando os dois candidatos (WRR-180 e WRR-181) com atividade inibitória da M^{pro} de SARS-CoV-2 no NB3 do Instituto de Ciências Biológicas (ICB) da UFMG.

5.3.2 Avaliação da atividade antiviral de moléculas bioativas contra MHV-3

Sabendo que a combinação de estratégias computacionais de sucesso é de interesse para investigações posteriores em modelos *in vitro* e *in vivo*, extratos e moléculas bioativas obtidas de uma triagem de isolados triterpenos com atividade antiviral predita contra M^{pro} de SARS-CoV-2 – dados em preparação – foram selecionadas para avaliação de sua atividade antiviral. Nesse sentido e, mediante restrição de trabalho em nível de biossegurança 2 (NB2), um coronavírus murino (MHV-3), CR-2, foi utilizado em ensaios celulares com L929, empregando a técnica colorimétrica de MTT para determinação das concentrações citotóxica (CC₅₀) e efetiva de 50% (EC₅₀). Todos os extratos e isolados avaliados apresentaram valores de CC₅₀ superiores a 200 µg/mL e 100 µM, respectivamente (**Tabela 16**). A concentração efetiva de 50% também foi avaliada sob as mesmas condições, adicionando uma suspensão viral de MHV-3 a uma MOI de 0,1. Dois triterpenos, AG6 e AG11 (2,91 ± 0,26 e 68,87 ± 5,54 µM) e dois extratos, EAE e EH (166,94 ± 14,61 µg/mL e 185,73 ± 15,6 µg/mL), apresentaram atividade antiviral, resultando em IS entre 1,08 e 34,36.

Tabela 16. Avaliação biológica de extratos e triterpenos contra MHV-3 em células L929

Substância	CC ₅₀ (μM ou μg/mL)	EC ₅₀ (μM ou μg/mL)	IS
EH	> 200 μg/mL	185,73 ± 15,6 μg/mL	> 1,08
EC	> 200 μg/mL	NA	-
EAE	> 200 μg/mL	166,94 ± 14,61 μg/mL	> 1,2
AB5	> 100 μM	NA	-
AG11	> 100 μM	68,87 ± 5,54 μM	> 1,45
MB4	> 100 μM	NA	-
AG6	> 100 μM	2,91 ± 0,26 μM	> 34,36
NY7	> 100 μM	NA	-
EL5	> 100 μM	NA	-
RW5	> 100 μM	NA	-
EL7	> 100 μM	NA	-
Ribavirina	80,8 μM	< 10 μM	> 8,08

NA: não ativo.

Chio et al., (2016), já haviam demonstrado o potencial de extratos da folha de menta de peixe (*Houttuynia cordata*) na atividade antiviral contra MHV em células CCL9 (linhagem de epitélio hepático de camundongo). Extratos da folha, além de isolados de cinanserina foram avaliados contra o vírus empregando MTT, apresentando valores de EC₅₀ entre 0,98 e 62,5 μg/mL, com IS entre 2 e > 4 (CHIOU et al., 2016). A cinanserina, por sua vez, já havia sido demonstrada como inibidor das M^{pro} de SARS-CoV e HCoV-229E, com valores de IC₅₀ de 5 μM, e também confirmadas com atividade antiviral em ensaios de redução do número de placas de lise com células Vero, obtendo valores de EC₅₀ entre 19 e 34 μM (CHEN et al., 2005), sugerindo o potencial dessa e de demais moléculas como inibidores para M^{pro} de SARS-CoV-2.

Nesse âmbito, Yi et al., (2021), também identificaram isolados triterpenos de extratos de alcaçuz com atividade antiviral contra SARS-CoV-2 em células Vero, a partir de uma triagem virtual e análises de *docking* de 125 moléculas de uma quimioteca de ervas medicinais. Saponina e ácido glicirretínico foram identificados inibindo a multiplicação de SARS-CoV-2, com valores de EC₅₀ de 0,075 e 3,17 μM, respectivamente, com IS > 1.333,33 e > 31,54 (YI et al., 2022).

Alhadrami et al., (2021), também identificaram triterpenos isolados (ácidos ursólico, betulínico e maslínico) com atividade inibitória contra M^{pro} de SARS-CoV-2, inicialmente selecionados a partir de análises de *docking* e simulações de DM, obtendo valores de IC₅₀ entre 3,22 e 89,67 μ M (ALHADRAMI et al., 2021). Baev et al., (2022), também identificaram análogos de triterpenos derivados de um triterpenoide com atividade contra M^{pro} de SARS-CoV-2, preditos com análises de *docking*, obtendo valores de IC₅₀ entre $12,5 \pm 13$ e 100 ± 5 μ M nos ensaios de inibição da atividade enzimática (BAEV et al., 2022). Corroborando a breve triagem contra MHV-3 em células L929, esses dados demonstram a possibilidade do emprego de ensaios *in vitro* para validação de resultados de estratégias e abordagens computacionais de êxito (AZEVEDO et al., 2022), a exemplo do planejamento de moléculas bioativas desenvolvido no presente trabalho.

6 CONCLUSÃO

A estratégia com um modelo de HQSAR empregado para triagem da quimioteca BraCoLi identificou 24 potenciais inibidores da M^{pro} de SARS-CoV-2, que não inibiram ao menos 50% da atividade enzimática da protease quando avaliados na concentração de 10 μ M. No entanto, não foi definido o domínio de aplicabilidade nesta análise. Estes dados demonstram que a estratégia de HQSAR, isoladamente, pode apresentar robustez e acurácia preditiva, mas requer, por exemplo, a determinação do espaço químico dos ligantes propostos, bem como de um limiar para atividade biológica bem definido, a fim de aprimorar sua capacidade preditiva e potencial de sucesso.

Na estratégia de SBDD, seis resíduos de interesse foram identificados nas M^{pro} de diferentes coronavírus (His41, Phe140, Gly143, Cys145, His163 e Glu166), além de sete para PL^{pro} de SARS-CoV-2 (Cys111, Asp164, yr264, Gly266, Gly271, His272 e Asp286). Modelos farmacofóricos foram construídos e permitiram filtrar grandes bancos de dados buscando reduzir o número de candidatos a serem avaliados em ensaios biológicos, embora nenhuma das moléculas obtidas e disponíveis após a triagem tenha apresentado atividade inibitória superior a 50% nas concentrações de 10 e 5 μ M avaliadas. Ainda que plausíveis, o número de centroides “pancoronavírus” delimitados podem ter restringido ou desconsideraram possíveis inibidores que interagiriam somente com um ou dois resíduos nos alvos. Igualmente, a ausência de inibidores na quimioteca GMR avaliada (14 de 43 disponíveis), bem como a falta de recursos para aquisição de moléculas comercialmente disponíveis (MolPort), podem ter mitigado o potencial de *hits* positivos.

As moléculas consideradas análogos estruturais disponíveis e avaliadas, cinco tiveram IC₅₀ determinadas para as M^{pro} de SARS-CoV-2 (0.086 ± 0.0015 a 16.42 ± 0.095 μ M), sendo duas (WRR-180 e WRR-181) também ativas contra as M^{pro} de SARS-CoV e MERS-CoV (0.11 ± 0.0175 a 3.786 ± 0.64 μ M). Especificamente nestas duas moléculas, um substituinte eteno pode estar associado a atividade observada nas M^{pro} dos diferentes coronavírus, similarmente ao perfil de inibição de nirmatrelvir. Ademais, esse perfil é sugerido nas interações preditas com His41, Cys145 e Glu166 discutidas na literatura e corroboradas ao longo do presente trabalho.

Finalmente, uma série de extratos e isolados triterpenos previamente preditos contra SARS-CoV-2, especificamente com atividade inibitória contra M^{Pro}, foram avaliados contra um modelo MHV-3, mimetizando ensaios biológicos a serem feitos com SARS-CoV-2 no NB3 da UFMG (perspectivas do presente trabalho). Neste modelo de estudo em NB2, quatro substâncias, sendo dois triterpenos e dois extratos, apresentaram atividade antiviral resultando em IS entre 1,08 e 34,36. Portanto, estratégias computacionais, como as baseadas em alvos moleculares empregando a predição de sítios de ligação, *docking*, simulações de DM e triagens virtuais, a exemplo daquelas conduzidas no presente trabalho, são capazes de obter potenciais inibidores a serem confirmados posteriormente em ensaios *in vitro*, demonstrando o potencial de sucesso no planejamento de moléculas bioativas contra SARS-CoV-2.

7 REFERÊNCIAS BIBLIOGRÁFICAS

ABDELKADER, A. et al. In-Silico targeting of SARS-CoV-2 NSP6 for drug and natural products repurposing. **Virology**, v. 573, p. 96–110, ago. 2022.

ABDELNABI, R. et al. Nirmatrelvir-resistant SARS-CoV-2 is efficiently transmitted in female Syrian hamsters and retains partial susceptibility to treatment. **Nature Communications**, v. 14, n. 1, p. 2124, 14 abr. 2023.

AGRAWAL, P. et al. Benchmarking of different molecular docking methods for protein-peptide docking. **BMC Bioinformatics**, v. 19, n. S13, p. 426, 4 fev. 2019.

AHN, D.-G. et al. Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. 2012.

ALHADRAMI, H. A. et al. Olive-Derived Triterpenes Suppress SARS COV-2 Main Protease: A Promising Scaffold for Future Therapeutics. **Molecules**, v. 26, n. 9, p. 2654, 1 maio 2021.

ALJUHANI, A. et al. In vitro and computational investigations of novel synthetic carboxamide-linked pyridopyrrolopyrimidines with potent activity as SARS-CoV-2-M Pro inhibitors. **RSC Advances**, v. 12, n. 41, p. 26895–26907, 2022.

ALPERN, J. D.; GERTNER, E. Off-Label Therapies for COVID-19-Are We All In This Together? **CLINICAL PHARMACOLOGY & THERAPEUTICS** |, 2020.

ALVES, V. M. et al. QUIMIOINFORMÁTICA: UMA INTRODUÇÃO. **Química Nova**, v. 41, p. 202–212, fev. 2018.

AMANAT, F.; KRAMMER, F. **SARS-CoV-2 Vaccines: Status Report**. ImmunityCell Press, , 14 abr. 2020.

AMIN, M.; SOROUR, M. K.; KASRY, A. Comparing the Binding Interactions in the Receptor Binding Domains of SARS-CoV-2 and SARS-CoV. **The journal of physical chemistry letters**, v. 11, n. 12, p. 4897–4900, 2020.

ANDERSON, A. S.; CAUBEL, P.; RUSNAK, J. M. Nirmatrelvir–Ritonavir and Viral Load Rebound in Covid-19. **New England Journal of Medicine**, v. 387, n. 11, p. 1047–1049, 15 set. 2022.

ANG, K. K. H. et al. Mining a Cathepsin Inhibitor Library for New Antiparasitic Drug Leads. **PLoS Neglected Tropical Diseases**, v. 5, n. 5, p. e1023, 3 maio 2011.

ANTONOPOULOU, I. et al. Inhibition of the main protease of SARS-CoV-2 (Mpro) by repurposing/designing drug-like substances and utilizing nature’s toolbox of bioactive compounds. **Computational and Structural Biotechnology Journal**, v. 20, p. 1306–1344, 1 jan. 2022.

AQUINO, E. M. L. et al. Social distancing measures to control the COVID-19 pandemic: Potential impacts and challenges in Brazil. **Ciencia e Saude Coletiva**, v. 25, p. 2423–2446, 1 jun. 2020.

ARBEL, R. et al. Nirmatrelvir Use and Severe Covid-19 Outcomes during the Omicron Surge. **New England Journal of Medicine**, v. 387, n. 9, p. 790–798, 1 set. 2022.

ARUN, K. G. et al. Drug repurposing against SARS-CoV-2 using E-pharmacophore based virtual screening, molecular docking and molecular dynamics with main protease as the target. **Journal of biomolecular structure & dynamics**, v. 0, n. 0, p. 1–12, 2020.

ASHHURST, A. S. et al. Potent Anti-SARS-CoV-2 Activity by the Natural Product Gallinamide A and Analogues via Inhibition of Cathepsin L. **Journal of Medicinal Chemistry**, p. acs.jmedchem.1c01494, 3 nov. 2021.

ASHKENAZY, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. **Nucleic Acids Research**, v. 44, n. Web Server issue, p. W344–W350, 8 jul. 2016.

AZEVEDO, L. et al. Atherosclerosis fate in the era of tailored functional foods: Evidence-based guidelines elicited from structure- and ligand-based approaches. **Trends in Food Science & Technology**, v. 128, p. 75–89, 1 out. 2022.

BAEV, D. S. et al. Triterpenic Acid Amides as Potential Inhibitors of the SARS-CoV-2 Main Protease. **Molecules (Basel, Switzerland)**, v. 28, n. 1, p. 303, 30 dez. 2022.

BAI, Y. et al. **Presumed Asymptomatic Carrier Transmission of COVID-19. JAMA - Journal of the American Medical Association**American Medical Association, , 14 abr. 2020.

BAI, Z. et al. The SARS-CoV-2 Nucleocapsid Protein and Its Role in Viral Structure, Biological Functions, and a Potential Target for Drug or Vaccine Mitigation. **Viruses**, v. 13, n. 6, p. 1115, 10 jun. 2021.

BAKER, A. N. et al. The SARS-COV-2 Spike Protein Binds Sialic Acids and Enables Rapid Detection in a Lateral Flow Point of Care Diagnostic Device. **ACS Central Science**, v. 6, n. 11, p. 2046–2052, 25 nov. 2020.

BANERJEE, P. et al. ProTox-II: a webserver for the prediction of toxicity of chemicals. **Nucleic Acids Research**, v. 46, n. W1, p. W257–W263, 2 jul. 2018.

BARDA, N. et al. Effectiveness of a third dose of the BNT162b2 mRNA COVID-19 vaccine for preventing severe outcomes in Israel: an observational study. **The Lancet**, v. 398, n. 10316, p. 2093–2100, 4 dez. 2021.

BARTIK, A. W. et al. The impact of COVID-19 on small business outcomes and expectations. **Proceedings of the National Academy of Sciences**, v. 117, n. 30, p. 17656–17666, 28 jul. 2020.

BEIGEL, J. H. et al. Remdesivir for the Treatment of Covid-19 — Final Report. **New England Journal of Medicine**, v. 383, n. 19, p. 1813–1826, 5 nov. 2020.

BERMAN, H. M. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, jan. 2000.

BERTHOLD, M. R. et al. KNIME: The Konstanz Information Miner. Em: PREISACH, C. et al. (Eds.). **Data Analysis, Machine Learning and Applications**. Studies in Classification, Data Analysis, and Knowledge Organization. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. p. 319–326.

BOLCATO, G. et al. Can We Still Trust Docking Results? An Extension of the Applicability of DockBench on PDBbind Database. **International Journal of Molecular Sciences**, v. 20, n. 14, p. 3558, 20 jul. 2019.

BORODINA, Y. V. et al. Assessment of conformational ensemble sizes necessary for specific resolutions of coverage of conformational space. **Journal of Chemical Information and Modeling**, v. 47, n. 4, p. 1428–1437, jul. 2007.

BOSCH, B. J. et al. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. **Journal of Virology**, v. 77, n. 16, p. 8801–8811, 15 ago. 2003.

BOWERS, K. J. et al. **Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters**. ACM/IEEE SC 2006 Conference (SC'06). **Anais...**Tampa, FL: IEEE, nov. 2006. Disponível em: <<https://ieeexplore.ieee.org/document/4090217/>>. Acesso em: 25 mar. 2021

BRAGA, R. C. et al. Pred-hERG: A Novel web-Accessible Computational Tool for Predicting Cardiac Toxicity. **Molecular Informatics**, v. 34, n. 10, p. 698–701, out. 2015.

BRÜMMER, L. E. et al. Accuracy of novel antigen rapid diagnostics for SARS-CoV-2: A living systematic review and meta-analysis. **PLOS Medicine**, v. 18, n. 8, p. e1003735, 12 ago. 2021.

BRUST, R. et al. Proteins in Action: Femtosecond to Millisecond Structural Dynamics of a Photoactive Flavoprotein. **Journal of the American Chemical Society**, v. 135, n. 43, p. 16168–16174, 30 out. 2013.

BUTOWT, R.; BILINSKA, K.; VON BARTHELD, C. S. Olfactory dysfunction in COVID-19: new insights into the underlying mechanisms. **Trends in Neurosciences**, v. 46, n. 1, p. 75–90, jan. 2023.

CAO, Y. et al. Characterization of the SARS-CoV-2 E Protein: Sequence, Structure, Viroporin, and Inhibitors. **Protein Science: A Publication of the Protein Society**, v. 30, n. 6, p. 1114–1130, jun. 2021.

CASSUTO, N. G. et al. Evaluation of a SARS-CoV-2 antigen-detecting rapid diagnostic test as a self-test: Diagnostic performance and usability. **Journal of Medical Virology**, v. 93, n. 12, p. 6686–6692, dez. 2021.

CASTRO, J. T. et al. Promotion of neutralizing antibody-independent immunity to wild-type and SARS-CoV-2 variants of concern using an RBD-Nucleocapsid fusion protein. **Nature Communications**, v. 13, n. 1, p. 4831, 17 ago. 2022.

CASTRO-ALVAREZ, A.; COSTA, A. M.; VILARRASA, J. The Performance of Several Docking Programs at Reproducing Protein-Macrolide-Like Crystal Structures. **Molecules**, v. 22, n. 1, p. 136, 17 jan. 2017.

CDC. **COVID-19 Vaccination**. Disponível em: <<https://www.cdc.gov/coronavirus/2019-nCoV/vaccines/safety/adverse-events.html>>. Acesso em: 22 mar. 2022.

CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature. **Geosci. Model Dev**, v. 7, p. 1247–1250, 2014.

CHANG, C. K. et al. Modular organization of SARS coronavirus nucleocapsid protein. **Journal of Biomedical Science**, v. 13, n. 1, p. 59–72, jan. 2006.

CHEN, J. Pathogenicity and transmissibility of 2019-nCoV—A quick overview and comparison with other emerging viruses. **Microbes and Infection**, v. 22, n. 2, p. 69–71, 1 mar. 2020.

CHEN, L. et al. Cinanserin Is an Inhibitor of the 3C-Like Proteinase of Severe Acute Respiratory Syndrome Coronavirus and Strongly Reduces Virus Replication In Vitro. **Journal of Virology**, v. 79, n. 11, p. 7095–7103, jun. 2005.

CHEN, N. et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. **The Lancet**, v. 395, n. 10223, p. 507–513, 15 fev. 2020.

CHEN, W. et al. Advances and challenges in using nirmatrelvir and its derivatives against SARS-CoV-2 infection. **Journal of Pharmaceutical Analysis**, 2 nov. 2022.

CHEN, Y.; LIU, Q.; GUO, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. **Journal of Medical Virology**, v. 92, n. 4, p. 418–423, 7 abr. 2020.

CHIU, K. H. et al. Evaluation of antiviral activities of *Houttuynia cordata* Thunb. extract, quercetin, quercetrin and cinanserin on murine coronavirus and dengue virus infection. **Asian Pacific Journal of Tropical Medicine**, v. 9, n. 1, p. 1–7, jan. 2016.

CHIRICO, N.; GRAMATICA, P. Real External Predictivity of QSAR Models: How To Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient. **Journal of Chemical Information and Modeling**, v. 51, n. 9, p. 2320–2335, 26 set. 2011.

CHO, E. et al. **Structural dynamics of the β -coronavirus M^{pro} protease ligand binding sites**. [s.l.] Biophysics, 1 abr. 2021. Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2021.03.31.437918>>. Acesso em: 11 jan. 2023.

CIOTTI, M. et al. Antimicrobial Section / Review COVID-19 Outbreak: An Overview Keywords COVID-19 · SARS-CoV-2 · Pandemic · Phylogenesis · Protein modeling · Real-time polymerase chain reaction. **Chemotherapy**, v. 64, p. 215–223, 2019.

CITARELLA, A. et al. SARS-CoV-2 Mpro: A Potential Target for Peptidomimetics and Small-Molecule Inhibitors. **Biomolecules**, v. 11, n. 4, p. 607, abr. 2021.

COX, R. M.; WOLF, J. D.; PLEMPER, R. K. Therapeutically administered ribonucleoside analogue MK-4482/EIDD-2801 blocks SARS-CoV-2 transmission in ferrets. **Nature Microbiology**, v. 6, n. 1, p. 11–18, jan. 2021.

CROSS, J. B. et al. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. **Journal of Chemical Information and Modeling**, v. 49, n. 6, p. 1455–1474, 22 jun. 2009.

CUI, J.; LI, F.; SHI, Z. L. **Origin and evolution of pathogenic coronaviruses.** **Nature Reviews Microbiology** Nature Publishing Group, , 1 mar. 2019.

CUI, T. et al. Role of Porcine Aminopeptidase N and Sialic Acids in Porcine Coronavirus Infections in Primary Porcine Enterocytes. **Viruses**, v. 12, n. 4, p. 402, abr. 2020.

DAWOOD, A. A. The efficacy of Paxlovid against COVID-19 is the result of the tight molecular docking between Mpro and antiviral drugs (nirmatrelvir and ritonavir). **Advances in Medical Sciences**, v. 68, n. 1, p. 1–9, 2 nov. 2022.

DE OLIVEIRA, V. et al. H172Y mutation perturbs the S1 pocket and nirmatrelvir binding of SARS-CoV-2 main protease through a nonnative hydrogen bond. **Research Square**, p. rs.3.rs-1915291, 9 ago. 2022.

DE WIT, E. et al. **SARS and MERS: Recent insights into emerging coronaviruses.** **Nature Reviews Microbiology** Nature Publishing Group, , 1 ago. 2016.

DEDIEGO, M. L. et al. A Severe Acute Respiratory Syndrome Coronavirus That Lacks the E Gene Is Attenuated In Vitro and In Vivo. **Journal of Virology**, v. 81, n. 4, p. 1701–1713, 15 fev. 2007.

DENG, X. et al. Coronaviruses Resistant to a 3C-Like Protease Inhibitor Are Attenuated for Replication and Pathogenesis, Revealing a Low Genetic Barrier but High Fitness Cost of Resistance. **Journal of Virology**, v. 88, n. 20, p. 11886–11898, out. 2014.

DINNES, J. et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. **The Cochrane Database of Systematic Reviews**, v. 3, p. CD013705, 24 mar. 2021.

DOBAÑO, C. et al. Immunogenicity and crossreactivity of antibodies to the nucleocapsid protein of SARS-CoV-2: utility and limitations in seroprevalence and immunity studies. **Translational Research**, v. 232, p. 60–74, jun. 2021.

DONG, S. et al. A guideline for homology modeling of the proteins from newly discovered betacoronavirus, 2019 novel coronavirus (2019-nCoV). **Journal of Medical Virology**, 2020.

DUAN, Y. et al. Disparities in COVID-19 Vaccination among Low-, Middle-, and High-Income Countries: The Mediating Role of Vaccination Policy. **Vaccines**, v. 9, n. 8, p. 905, 14 ago. 2021.

DUAN, Y.; KOLLMAN, P. A. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. **Science**, v. 282, n. 5389, p. 740–744, 23 out. 1998.

DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. **BMC Biology**, v. 9, n. 1, p. 71, 28 dez. 2011.

DUVEAU, D. Y.; THOMAS, C. J. The Remarkable Selectivity of Nirmatrelvir. **ACS Pharmacology & Translational Science**, v. 5, n. 6, p. 445–447, 10 jun. 2022.

EMA. **Paxlovid**. Text. Disponível em: <<https://www.ema.europa.eu/en/medicines/human/EPAR/paxlovid>>. Acesso em: 18 jan. 2023.

EMANUEL, E. J.; OSTERHOLM, M.; GOUNDER, C. R. A National Strategy for the “New Normal” of Life With COVID. **JAMA**, v. 327, n. 3, p. 211–212, 18 jan. 2022.

FEHR, A. R.; PERLMAN, S. Coronaviruses: An overview of their replication and pathogenesis. Em: **Coronaviruses: Methods and Protocols**. [s.l.] Springer New York, 2015. v. 1282p. 1–23.

FERREIRA, G. M. et al. SARS-COV-2 Mpro conformational changes induced by covalently bound ligands. **Journal of Biomolecular Structure & Dynamics**, p. 1–11, 13 set. 2021.

FERREIRA, G. M. et al. Inhibitor induced conformational changes in SARS-COV-2 papain-like protease. **Scientific Reports**, v. 12, p. 11585, 8 jul. 2022.

FERRETTI, L. et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. **Science**, v. 368, n. 6491, 8 maio 2020.

FISCHER, A. et al. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. **Journal of Medicinal Chemistry**, v. 64, n. 5, p. 2489–2500, 11 mar. 2021.

FOLEGATTI, P. M. et al. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. **The Lancet**, v. 396, n. 10249, p. 467–478, 15 ago. 2020.

FORSYTHE, S. S. et al. Twenty Years Of Antiretroviral Therapy For People Living With HIV: Global Costs, Health Achievements, Economic Benefits. **Health Affairs**, v. 38, n. 7, p. 1163–1172, jul. 2019.

FRY, A. M. et al. Efficacy of oseltamivir treatment started within 5 days of symptom onset to reduce influenza illness duration and virus shedding in an urban setting in Bangladesh: a randomised placebo-controlled trial. **The Lancet. Infectious Diseases**, v. 14, n. 2, p. 109–118, fev. 2014.

FU, L. et al. Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. **Nature Communications**, v. 11, n. 1, p. 4417, 4 set. 2020.

GADHAVE, K. et al. Conformational dynamics of 13 amino acids long NSP11 of SARS-CoV-2 under membrane mimetics and different solvent conditions. **Microbial Pathogenesis**, v. 158, p. 105041, set. 2021.

GARCIA, T. S.; HONÓRIO, K. M. Two-dimensional quantitative structure-activity relationship studies on bioactive ligands of Peroxisome proliferator-activated receptor δ . **Journal of the Brazilian Chemical Society**, v. 22, n. 1, p. 65–72, 2011.

GHOSH, A. K. et al. Indole Chloropyridinyl Ester-Derived SARS-CoV-2 3CLpro Inhibitors: Enzyme Inhibition, Antiviral Efficacy, Structure–Activity Relationship, and X-ray Structural Studies. **Journal of Medicinal Chemistry**, p. acs.jmedchem.1c01214, 16 set. 2021.

GIOFRÈ, S. V. et al. Interaction of selected terpenoids with two SARS-CoV-2 key therapeutic targets: An in silico study through molecular docking and dynamics simulations. **Computers in Biology and Medicine**, v. 134, p. 104538, jul. 2021.

GIRARDIN, F. et al. Evaluating the risk of drug-drug interactions with pharmacokinetic boosters: the case of ritonavir-enhanced nirmatrelvir to prevent severe COVID-19. **Clinical Microbiology and Infection**, v. 28, n. 8, p. 1044–1046, 1 ago. 2022.

GOETZ, D. H. et al. Substrate specificity profiling and identification of a new class of inhibitor for the major protease of the SARS coronavirus. **Biochemistry**, v. 46, n. 30, p. 8744–8752, 31 jul. 2007.

GOLBRAIKH, A.; TROPSHA, A. **Beware of q2!** Journal of Molecular Graphics and Modelling. **Anais...**2002.

GORBALENYA, A. E. et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. **Nature Microbiology**, v. 5, n. 4, p. 536–544, 2020.

GORMAN, J. M. et al. Implications of COVID-19 Vaccine Hesitancy: Results of Online Bulletin Board Interviews. **Frontiers in Public Health**, v. 9, 2022.

GOTTLIEB, R. L. et al. Early Remdesivir to Prevent Progression to Severe Covid-19 in Outpatients. **New England Journal of Medicine**, v. 386, n. 4, p. 305–315, 27 jan. 2022.

GRAMATICA, P. Principles of QSAR Modeling. **International Journal of Quantitative Structure-Property Relationships**, v. 5, n. 3, p. 61–97, jul. 2020.

GRAMATICA, P.; SANGION, A. A Historical Excursus on the Statistical Validation Parameters for QSAR Models: A Clarification Concerning Metrics and Terminology. **Journal of Chemical Information and Modeling**, v. 56, n. 6, p. 1127–1131, 27 jun. 2016.

GREASLEY, S. E. et al. Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. **Journal of Biological Chemistry**, v. 298, n. 6, p. 101972, 1 jun. 2022.

GUIDO, R. V. C.; ANDRICOPULO, A. D.; OLIVA, G. Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas. **Estudos Avançados**, v. 24, n. 70, p. 81–98, 2010.

GUO, L. et al. Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19). **Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America**, v. 71, n. 15, p. 778–785, 28 jul. 2020.

H. BROOKS, W.; C. GUIDA, W.; G. DANIEL, K. The Significance of Chirality in Drug Design and Development. **Current Topics in Medicinal Chemistry**, v. 11, n. 7, p. 760–770, 28 mar. 2011.

- HAGE-MELIM, L. I. DA S. et al. Virtual screening, ADME/Tox predictions and the drug repurposing concept for future use of old drugs against the COVID-19. **Life Sciences**, v. 256, 1 set. 2020.
- HALGREN, T. New method for fast and accurate binding-site identification and analysis. **Chemical Biology & Drug Design**, v. 69, n. 2, p. 146–148, fev. 2007.
- HARMON, A. et al. Validation of an At-Home Direct Antigen Rapid Test for COVID-19. **JAMA Network Open**, v. 4, n. 8, p. e2126931, 27 ago. 2021.
- HAUG, N. et al. Ranking the effectiveness of worldwide COVID-19 government interventions. **Nature Human Behaviour**, v. 4, n. 12, p. 1303–1312, dez. 2020.
- HAVERI, A. et al. Serological and molecular findings during SARS-CoV-2 infection: the first case study in Finland, January to February 2020. **Eurosurveillance**, v. 25, n. 11, p. 2000266, 19 mar. 2020.
- HAYEK-ORDUZ, Y. et al. Novel covalent and non-covalent complex-based pharmacophore models of SARS-CoV-2 main protease (Mpro) elucidated by microsecond MD simulations. **Scientific Reports**, v. 12, n. 1, p. 14030, 18 ago. 2022.
- HERITAGE, T. W.; LOWIS, D. R. Molecular Hologram QSAR. Em: [s.l: s.n.]. p. 212–225.
- HESKIN, J. et al. Caution required with use of ritonavir-boosted PF-07321332 in COVID-19 management. **Lancet**, v. 399, n. 10319, p. 21–22, 1 jan. 2022.
- HIPPISLEY-COX, J. et al. Risk of thrombocytopenia and thromboembolism after covid-19 vaccination and SARS-CoV-2 positive testing: self-controlled case series study. **BMJ (Clinical research ed.)**, v. 374, p. n1931, 26 ago. 2021.
- HOFFMANN, M. et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. **Cell**, v. 181, n. 2, p. 271–280.e8, 2020.
- HOGAN, J.; HOWLETT, M.; MURPHY, M. Re-thinking the coronavirus pandemic as a policy punctuation: COVID-19 as a path-clearing policy accelerator. **Policy and Society**, v. 41, n. 1, p. 40–52, 1 jan. 2022.
- HU, Y. et al. Boceprevir, calpain inhibitors II and XII, and GC-376 have broad-spectrum antiviral activity against coronaviruses in cell culture. **bioRxiv**, p. 2020.10.30.362335, 1 nov. 2020.
- HU, Y. et al. Naturally occurring mutations of SARS-CoV-2 main protease confer drug resistance to nirmatrelvir. **bioRxiv**, p. 2022.06.28.497978, 6 set. 2022.
- HUANG, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. **The Lancet**, v. 395, n. 10223, p. 497–506, 15 fev. 2020a.
- HUANG, F. et al. Covalent Warheads Targeting Cysteine Residue: The Promising Approach in Drug Development. **Molecules**, v. 27, n. 22, p. 7728, jan. 2022.

HUANG, Y. et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. **Acta Pharmacologica Sinica**, v. 41, n. 9, p. 1141–1149, set. 2020b.

HUNG, Y.-P. et al. Oral Nirmatrelvir/Ritonavir Therapy for COVID-19: The Dawn in the Dark? **Antibiotics**, v. 11, n. 2, p. 220, 9 fev. 2022.

HURST, K. R.; KOETZNER, C. A.; MASTERS, P. S. Characterization of a Critical Interaction between the Coronavirus Nucleocapsid Protein and Nonstructural Protein 3 of the Viral Replicase-Transcriptase Complex. **Journal of Virology**, v. 87, n. 16, p. 9159–9172, 15 ago. 2013.

ICTV. **Coronaviridae - Positive Sense RNA Viruses**. Disponível em: <https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/222/coronaviridae>. Acesso em: 17 mar. 2021.

IKETANI, S. et al. Multiple pathways for SARS-CoV-2 resistance to nirmatrelvir. **Nature**, p. 1–7, 9 nov. 2022.

IONESCU, M. I. An Overview of the Crystallized Structures of the SARS-CoV-2. **The Protein Journal**, v. 39, n. 6, p. 600–618, dez. 2020.

JACKSON, C. B. et al. Mechanisms of SARS-CoV-2 entry into cells. **Nature Reviews. Molecular Cell Biology**, v. 23, n. 1, p. 3–20, jan. 2022.

JACKSON, L. A. et al. An mRNA Vaccine against SARS-CoV-2 — Preliminary Report. **New England Journal of Medicine**, v. 383, n. 20, p. 1920–1931, 12 nov. 2020.

JACOBSON, M. P. et al. A Hierarchical Approach to All-Atom Protein Loop Prediction. **Proteins: Structure, Function and Genetics**, v. 55, n. 2, p. 351–367, 2004.

JAIN, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. **Journal of Medicinal Chemistry**, v. 46, n. 4, p. 499–511, 13 fev. 2003.

JAYK BERNAL, A. et al. Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. **New England Journal of Medicine**, v. 386, n. 6, p. 509–520, 10 fev. 2022.

JEFFERSON, T. et al. Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. **BMJ**, v. 348, p. g2545, 9 abr. 2014.

JENDELE, L. et al. PrankWeb: a web server for ligand binding site prediction and visualization. **Nucleic Acids Research**, v. 47, n. 1, p. 345–349, 2 jul. 2019.

JIMÉNEZ-AVALOS, G. et al. Comprehensive virtual screening of 4.8 k flavonoids reveals novel insights into allosteric inhibition of SARS-CoV-2 MPRO. **Scientific Reports**, v. 11, n. 1, p. 15452, 29 jul. 2021.

JIN, Z. et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. **Nature**, v. 582, n. 7811, p. 289–293, jun. 2020a.

JIN, Z. et al. Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. **Nature Structural and Molecular Biology**, v. 27, n. 6, p. 529–532, 1 jun. 2020b.

JOCHMANS, D. et al. The Substitutions L50F, E166A, and L167F in SARS-CoV-2 3CLpro Are Selected by a Protease Inhibitor In Vitro and Confer Resistance To Nirmatrelvir. **mBio**, v. 14, n. 1, p. e0281522, 28 fev. 2023.

JOSEFSBERG, J. O.; BUCKLAND, B. Vaccine process technology. **Biotechnology and Bioengineering**, v. 109, n. 6, p. 1443–1460, 2012.

KALIL, A. C. **Treating COVID-19 - Off-Label Drug Use, Compassionate Use, and Randomized Clinical Trials during Pandemics**. **JAMA - Journal of the American Medical Association**American Medical Association, , 19 maio 2020.

KAPLA, J. et al. Can molecular dynamics simulations improve the structural accuracy and virtual screening performance of GPCR models? **PLoS computational biology**, v. 17, n. 5, p. e1008936, maio 2021.

KELLER, M. J. et al. Effect of Systemic Glucocorticoids on Mortality or Mechanical Ventilation in Patients With COVID-19. **Journal of Hospital Medicine**, v. 22, n. 2020- 07-22 ONLINE FIRST, p. 2020, 22 jul. 2020.

KHOO, S. H. et al. Optimal dose and safety of molnupiravir in patients with early SARS-CoV-2: a Phase I, open-label, dose-escalating, randomized controlled study. **The Journal of Antimicrobial Chemotherapy**, v. 76, n. 12, p. 3286–3295, 12 nov. 2021.

KIM, Y. et al. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. **Journal of Virology**, v. 86, n. 21, p. 11754–11762, nov. 2012.

KIRALJ, R.; FERREIRA, M. M. C. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. **Journal of the Brazilian Chemical Society**, v. 20, n. 4, p. 770–787, 2009.

KIRCHDOERFER, R. N.; WARD, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. **Nature Communications**, v. 10, n. 1, p. 1–9, 1 dez. 2019.

KOLA, L. et al. COVID-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. **The Lancet Psychiatry**, v. 8, n. 6, p. 535–550, jun. 2021.

KOO, J. R. et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. **The Lancet Infectious Diseases**, v. 20, p. 678–688, 2020.

KOZAKOV, D. et al. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. **Nature Protocols**, v. 10, n. 5, p. 733–755, 6 maio 2015.

KRONENBERGER, T. et al. On the relationship of anthranilic derivatives structure and the FXR (Farnesoid X receptor) agonist activity. **Journal of Biomolecular Structure & Dynamics**, v. 36, n. 16, p. 4378–4391, dez. 2018.

KRONENBERGER, T. et al. Ligand Accessibility Insights to the Dengue Virus NS3-NS2B Protease Assessed by Long-Timescale Molecular Dynamics Simulations. **ChemMedChem**, v. 16, n. 16, p. 2524–2534, 19 ago. 2021.

KUMAR, S. et al. Discovery of New Hydroxyethylamine Analogs against 3CLpro Protein Target of SARS-CoV-2: Molecular Docking, Molecular Dynamics Simulation, and Structure-Activity Relationship Studies. **Journal of chemical information and modeling**, 2020.

KUMAR, Y.; SINGH, H.; PATEL, C. N. In silico prediction of potential inhibitors for the main protease of SARS-CoV-2 using molecular docking and dynamics simulation based drug-repurposing. **Journal of Infection and Public Health**, v. 13, n. 9, p. 1210–1223, set. 2020.

LANG, Y. et al. Coronavirus hemagglutinin-esterase and spike proteins coevolve for functional balance and optimal virion avidity. **Proceedings of the National Academy of Sciences**, v. 117, n. 41, p. 25759–25770, 13 out. 2020.

LANGE, N. W. et al. Nirmatrelvir/ritonavir use: Managing clinically significant drug-drug interactions with transplant immunosuppressants. **American Journal of Transplantation**, v. 22, n. 7, p. 1925–1926, 1 jul. 2022.

LAU, C. L.; GALEA, I. Risk–benefit analysis of COVID-19 vaccines — a neurological perspective. **Nature Reviews Neurology**, v. 18, n. 2, p. 69–70, fev. 2022.

LAU, S. K. P. et al. Possible Bat Origin of Severe Acute Respiratory Syndrome Coronavirus 2. **Emerging Infectious Diseases**, v. 26, n. 7, p. 1542–1547, jul. 2020.

LEE, C.-C.; HSIEH, C.-C.; KO, W.-C. Molnupiravir-A Novel Oral Anti-SARS-CoV-2 Agent. **Antibiotics (Basel, Switzerland)**, v. 10, n. 11, p. 1294, 23 out. 2021.

LEFKOWITZ, E. J. et al. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). **Nucleic Acids Research**, v. 46, 2018.

LEMAITRE, F. et al. Management of drug-drug interactions with nirmatrelvir/ritonavir in patients treated for Covid-19: Guidelines from the French Society of Pharmacology and Therapeutics (SFPT). **Therapie**, v. 77, n. 5, p. 509–521, 2022.

LEWNARD, J. A.; LO, N. C. **Scientific and ethical basis for social-distancing interventions against COVID-19. The Lancet Infectious Diseases**Lancet Publishing Group, , 2020.

LI, H. et al. **SARS-CoV-2 and viral sepsis: observations and hypotheses. The Lancet**Lancet Publishing Group, , 9 maio 2020.

LI, H.; XUE, Q.; XU, X. **Involvement of the Nervous System in SARS-CoV-2 Infection. Neurotoxicity Research**Springer, , 1 jun. 2020.

LI, J. et al. Structural Basis of the Main Proteases of Coronavirus Bound to Drug Candidate PF-07321332. **Journal of Virology**, v. 96, n. 8, p. e0201321, 27 abr. 2022.

LI, Y. et al. Role of air distribution in SARS transmission during the largest nosocomial outbreak in Hong Kong. **Indoor Air**, v. 15, n. 2, p. 83–95, abr. 2005.

- LIANG, J. et al. Interaction of the prototypical α -ketoamide inhibitor with the SARS-CoV-2 main protease active site in silico: Molecular dynamic simulations highlight the stability of the ligand-protein complex. **Computational Biology and Chemistry**, v. 87, p. 107292, 28 maio 2020a.
- LIANG, J. et al. Site mapping and small molecule blind docking reveal a possible target site on the SARS-CoV-2 main protease dimer interface. **Computational Biology and Chemistry**, v. 89, p. 107372, dez. 2020b.
- LIMA, A. N. et al. Use of machine learning approaches for novel drug discovery. **Expert Opinion on Drug Discovery**, v. 11, n. 3, p. 225–239, 3 mar. 2016.
- LIPINSKI, C.; HOPKINS, A. **Navigating chemical space for biology and medicine.** *Nature*, 16 dez. 2004.
- LIU, S.; ZHENG, Q.; WANG, Z. Potential covalent drugs targeting the main protease of the SARS-CoV-2 coronavirus. **Bioinformatics (Oxford, England)**, v. 36, n. 11, p. 3295–3298, 01 2020.
- LOGUNOV, D. Y. et al. Safety and immunogenicity of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine in two formulations: two open, non-randomised phase 1/2 studies from Russia. **The Lancet**, v. 396, n. 10255, p. 887–897, 26 set. 2020.
- LONG, Q. X. et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. **Nature Medicine**, p. 1–5, 18 jun. 2020.
- LU, C. et al. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. **Journal of Chemical Theory and Computation**, v. 17, n. 7, p. 4291–4300, 13 jul. 2021.
- LU, H.; STRATTON, C. W.; TANG, Y. W. **An Evolving Approach to the Laboratory Assessment of COVID-19.** *Journal of medical virology*NLM (Medline), , 29 abr. 2020.
- LU, X. et al. The Development of Pharmacophore Modeling: Generation and Recent Applications in Drug Discovery. **Current Pharmaceutical Design**, v. 24, n. 29, p. 3424–3439, 8 dez. 2018.
- LUAN, X. et al. Structure basis for inhibition of SARS-CoV-2 by the feline drug GC376. **Acta Pharmacologica Sinica**, v. 44, n. 1, p. 255–257, jan. 2023.
- LUO, D. et al. Computational strategies towards developing novel SARS-CoV-2 Mpro inhibitors against COVID-19. **Journal of Molecular Structure**, v. 1247, p. 131378, 5 jan. 2022.
- MADEIRA, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. **Nucleic Acids Research**, v. 47, n. 1, p. 636–641, 2 jul. 2019.
- MAGGIORA, G. et al. Molecular Similarity in Medicinal Chemistry. **Journal of Medicinal Chemistry**, v. 57, n. 8, p. 3186–3204, 24 abr. 2014.

MAGRO, G. **COVID-19: Review on latest available drugs and therapies against SARS-CoV-2. Coagulation and inflammation cross-talking.** *Virus Research* Elsevier B.V., , 1 set. 2020.

MAHANTA, S. et al. **Potential anti-viral activity of approved repurposed drug against main protease of SARS-CoV-2: an in silico based approach.** *Journal of Biomolecular Structure and Dynamics* Taylor and Francis Ltd., , 2020.

MALONE, B. et al. Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nature Reviews Molecular Cell Biology*, v. 23, n. 1, p. 21–39, jan. 2022.

MALTAROLLO, V. G. et al. Do Go Chasing Waterfalls: Enoyl Reductase (FabI) in Complex with Inhibitors Stabilizes the Tetrameric Structure and Opens Water Channels. *Journal of Chemical Information and Modeling*, v. 62, n. 22, p. 5746–5761, 28 nov. 2022.

MATHEA, M.; KLINGSPOHN, W.; BAUMANN, K. Chemoinformatic Classification Methods and their Applicability Domain. *Molecular Informatics*, v. 35, n. 5, p. 160–180, 2016.

MCCLOSKEY, B. et al. Emerging infectious diseases and pandemic potential: status quo and reducing risk of global spread. *The Lancet Infectious Diseases*, v. 14, n. 10, p. 1001–1010, 1 out. 2014.

MEINHARDT, J. et al. Olfactory transmucosal SARS-CoV-2 invasion as a port of central nervous system entry in individuals with COVID-19. *Nature Neuroscience*, v. 24, n. 2, p. 168–175, fev. 2021.

MELLOTT, D. M. et al. A Clinical-Stage Cysteine Protease Inhibitor blocks SARS-CoV-2 Infection of Human and Monkey Cells. *ACS chemical biology*, v. 16, n. 4, p. 642–650, 16 abr. 2021.

MIELECH, A. M. et al. Nidovirus papain-like proteases: Multifunctional enzymes with protease, deubiquitinating and deISGylating activities. *Virus Research*, v. 194, p. 184–190, 19 dez. 2014.

MOGHADAS, S. M. et al. The Impact of Vaccination on Coronavirus Disease 2019 (COVID-19) Outbreaks in the United States. *Clinical Infectious Diseases*, v. 73, n. 12, p. 2257–2264, 15 dez. 2021.

MORENS, D. M.; FAUCI, A. S. Emerging Pandemic Diseases: How We Got to COVID-19. *Cell*, v. 182, n. 5, p. 1077–1092, 3 set. 2020.

MOSMANN, T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *Journal of Immunological Methods*, v. 65, n. 1–2, p. 55–63, 1983.

MULANGU, S. et al. A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *New England Journal of Medicine*, v. 381, n. 24, p. 2293–2303, 12 dez. 2019.

MULLIGAN, M. J. et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. *Nature*, v. 586, n. 7830, p. 589–593, 2020.

NAL, B. et al. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. **Journal of General Virology**, v. 86, n. 5, p. 1423–1434, maio 2005.

NELAKUDITI, B.; SHRIVASTAVA, D. A. Drug Repurposing 57 well-known drugs for three COVID-19 targets: Mpro, Spike, RdRp. 2021.

NEUMAN, B. W. et al. A structural analysis of M protein in coronavirus assembly and morphology. **Journal of Structural Biology**, v. 174, n. 1, p. 11–22, abr. 2011.

NGAN, C.-H. et al. FTSite: high accuracy detection of ligand binding sites on unbound protein structures. **Bioinformatics**, v. 28, n. 2, p. 286–287, 15 jan. 2012.

NGO, S. T. et al. Computational Determination of Potential Inhibitors of SARS-CoV-2 Main Protease. **Journal of Chemical Information and Modeling**, v. 60, n. 12, p. 5771–5780, 28 dez. 2020.

NIETO-TORRES, J. L. et al. Severe Acute Respiratory Syndrome Coronavirus Envelope Protein Ion Channel Activity Promotes Virus Fitness and Pathogenesis. **PLoS Pathogens**, v. 10, n. 5, 2014.

NJOROGE, F. G. et al. Challenges in modern drug discovery: a case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection. **Accounts of Chemical Research**, v. 41, n. 1, p. 50–59, jan. 2008.

OLIVERA MESA, D. et al. Modelling the impact of vaccine hesitancy in prolonging the need for Non-Pharmaceutical Interventions to control the COVID-19 pandemic. **Communications Medicine**, v. 2, n. 1, p. 1–8, 10 fev. 2022.

ONG, S. W. X. et al. **Air, Surface Environmental, and Personal Protective Equipment Contamination by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) from a Symptomatic Patient.** **JAMA - Journal of the American Medical Association**American Medical Association, , 28 abr. 2020.

ORABY, T. et al. Modeling the effect of lockdown timing as a COVID-19 control measure in countries with differing social contacts. **Scientific Reports**, v. 11, n. 1, p. 3354, 8 fev. 2021.

ORNELL, F. et al. “Pandemic fear” and COVID-19: mental health burden and strategies. **Revista Brasileira De Psiquiatria (Sao Paulo, Brazil: 1999)**, v. 42, n. 3, p. 232–235, 2020.

OSIPIUK, J. et al. Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. **Nature Communications**, v. 12, n. 1, p. 743, 2 fev. 2021.

OTTER, J. A. et al. **Transmission of SARS and MERS coronaviruses and influenza virus in healthcare settings: The possible role of dry surface contamination.** **Journal of Hospital Infection**W.B. Saunders Ltd, , 1 mar. 2016.

OWEN, D. R. et al. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. **Science**, v. 374, n. 6575, p. 1586–1593, 24 dez. 2021.

PADRON-REGALADO, E. Vaccines for SARS-CoV-2: Lessons from Other Coronavirus Strains. **Infectious Diseases and Therapy**, v. 9, n. 2, p. 255–274, 1 jun. 2020.

PAINTER, W. P. et al. Human Safety, Tolerability, and Pharmacokinetics of Molnupiravir, a Novel Broad-Spectrum Oral Antiviral Agent with Activity Against SARS-CoV-2. **Antimicrobial Agents and Chemotherapy**, v. 65, n. 5, p. e02428-20, AAC.02428-20, 1 mar. 2021.

PANT, S. et al. Peptide-like and small-molecule inhibitors against Covid-19. **Journal of biomolecular structure & dynamics**, v. 0, n. 0, p. 1–15, 2020.

PANTALEÃO, S. Q. et al. Recent Advances in the Prediction of Pharmacokinetics Properties in Drug Design Studies: A Review. **ChemMedChem**, v. 17, n. 1, p. e202100542, 5 jan. 2022.

PENG, Y.-C. et al. A Novel Rapid Test to Detect Anti-SARS-CoV-2 N Protein IgG Based on Shear Horizontal Surface Acoustic Wave (SH-SAW). **Diagnostics**, v. 11, n. 10, p. 1838, 5 out. 2021.

PEREIRA, D. G. Importância do metabolismo no planejamento de fármacos. **Química Nova**, v. 30, p. 171–177, fev. 2007.

PEREIRA OLIVEIRA, G.; KROON, E. G. Mouse hepatitis virus: A betacoronavirus model to study the virucidal activity of air disinfection equipment on surface contamination. **Journal of Virological Methods**, v. 297, p. 114274, nov. 2021.

PETERS, P. H. et al. Long-term use of oseltamivir for the prophylaxis of influenza in a vaccinated frail older population. **Journal of the American Geriatrics Society**, v. 49, n. 8, p. 1025–1031, ago. 2001.

PFIZER. **A PHASE 1, RANDOMIZED, DOUBLE-BLIND, SPONSOR-OPEN, PLACEBO CONTROLLED, SINGLE- AND MULTIPLE-DOSE ESCALATION STUDY TO EVALUATE THE SAFETY, TOLERABILITY, AND PHARMACOKINETICS OF PF-07321332 IN HEALTHY ADULT PARTICIPANTS.** [s.l.] clinicaltrials.gov, 18 out. 2021. Disponível em: <<https://clinicaltrials.gov/ct2/show/NCT04756531>>. Acesso em: 22 mar. 2022.

PILLAIYAR, T. et al. An Overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. **Journal of Medicinal Chemistry**, v. 59, n. 14, p. 6595–6628, 28 jul. 2016.

PILLAIYAR, T. et al. Small-Molecule Thioesters as SARS-CoV-2 Main Protease Inhibitors: Enzyme Inhibition, Structure-Activity Relationships, Antiviral Activity, and X-ray Structure Determination. **Journal of Medicinal Chemistry**, v. 65, n. 13, p. 9376–9395, 14 jul. 2022.

PILON, A. C. et al. NuBBEDB: an updated database to uncover chemical and biological information from Brazilian biodiversity. **Scientific Reports**, v. 7, n. 1, p. 7215, 3 ago. 2017.

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

PITSILLOU, E. et al. Identification of Small Molecule Inhibitors of the Deubiquitinating Activity of the SARS-CoV-2 Papain-Like Protease: in silico Molecular Docking Studies and in vitro Enzymatic Activity Assay. **Frontiers in Chemistry**, v. 8, p. 623971, 8 dez. 2020.

- PITSILLOU, E. et al. Interaction of small molecules with the SARS-CoV-2 papain-like protease: In silico studies and in vitro validation of protease activity inhibition using an enzymatic inhibition assay. **Journal of Molecular Graphics & Modelling**, v. 104, p. 107851, maio 2021.
- POLACK, F. P. et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. **The New England Journal of Medicine**, v. 383, n. 27, p. 2603–2615, 31 dez. 2020.
- PUSHPAKOM, S. et al. **Drug repurposing: Progress, challenges and recommendations.** **Nature Reviews Drug Discovery** Nature Publishing Group, , 28 dez. 2018.
- RAMAJAYAM, R.; TAN, K. P.; LIANG, P. H. Recent development of 3C and 3CL protease inhibitors for anti-coronavirus and anti-picornavirus drug discovery. **Biochemical Society Transactions**, v. 39, n. 5, p. 1371–1375, 2011.
- RAMOS-GUZMÁN, C. A.; RUIZ-PERNI, J. J.; TUÑÓN, I. Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale DFT / MM Methods Supplementary Information Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale DFT / MM Methods. n. 2, p. 1–23, 2020.
- RAUCH, S. et al. New Vaccine Technologies to Combat Outbreak Situations. **Frontiers in Immunology**, v. 9, p. 1963, 2018.
- REDONDO, N. et al. SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns. **Frontiers in Immunology**, v. 12, p. 708264, 7 jul. 2021.
- ROCHA, R. E. O. et al. A higher flexibility at the SARS-CoV-2 main protease active site compared to SARS-CoV and its potentialities for new inhibitor virtual screening targeting multi-conformers. **Journal of Biomolecular Structure & Dynamics**, p. 1–21, 10 maio 2021.
- ROSLING, L.; ROSLING, M. Pneumonia causes panic in Guangdong province. **BMJ (Clinical research ed.)**, v. 326, n. 7386, p. 416, 22 fev. 2003.
- RUAN, Z. et al. SARS-CoV-2 and SARS-CoV: Virtual screening of potential inhibitors targeting RNA-dependent RNA polymerase activity (NSP12). **Journal of Medical Virology**, v. 93, n. 1, p. 389–400, 2021.
- RUBIN, R. COVID-19 Vaccines vs Variants—Determining How Much Immunity Is Enough. **JAMA**, v. 325, n. 13, p. 1241–1243, 6 abr. 2021.
- RUIZ-GARCIA, A. et al. Pharmacokinetics in Drug Discovery. **Journal of Pharmaceutical Sciences**, v. 97, n. 2, p. 654–690, 2008.
- RUT, W. et al. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti-COVID-19 drug design. **Science Advances**, v. 6, n. 42, p. eabd4596, out. 2020.
- RUT, W. et al. SARS-CoV-2 Mpro inhibitors and activity-based probes for patient-sample imaging. **Nature Chemical Biology**, v. 17, n. 2, p. 222–228, fev. 2021.
- SALVATORI, G. et al. SARS-CoV-2 SPIKE PROTEIN: an optimal immunological target for vaccines. **Journal of Translational Medicine**, v. 18, n. 1, p. 222, 3 jun. 2020.

SASIDHARAN, S. et al. Bacterial protein azurin and derived peptides as potential anti-SARS-CoV-2 agents: insights from molecular docking and molecular dynamics simulations. **Journal of Biomolecular Structure and Dynamics**, p. 1–16, 3 jul. 2020.

SERAFIM, M. S. M. et al. Knowing and combating the enemy: a brief review on SARS-CoV-2 and computational approaches applied to the discovery of drug candidates. **Bioscience Reports**, v. 41, n. 3, p. BSR20202616, 26 mar. 2021a.

SERAFIM, M. S. M. et al. Machine learning techniques applied to the drug design and discovery of new antivirals: a brief look over the past decade. **Expert Opinion on Drug Discovery**, v. 16, n. 9, p. 961–975, set. 2021b.

SHANG, J. et al. Structural basis of receptor recognition by SARS-CoV-2. **Nature**, v. 581, n. 7807, p. 221–224, maio 2020.

SHEAHAN, T. P. et al. Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. **Nature Communications**, v. 11, n. 1, p. 1–14, 1 dez. 2020.

SHELLEY, J. C. et al. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. **Journal of Computer-Aided Molecular Design**, v. 21, n. 12, p. 681–691, 2007.

SHI, Y. et al. **COVID-19 infection: the perspectives on immune responses**. **Cell Death and Differentiation** Springer Nature, , 1 maio 2020.

SHIN, D. et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. **Nature**, v. 587, n. 7835, p. 657–662, nov. 2020.

SIDDIQI, H. K.; MEHRA, M. R. **COVID-19 illness in native and immunosuppressed states: A clinical–therapeutic staging proposal**. **Journal of Heart and Lung Transplantation** Elsevier USA, , 1 maio 2020.

SILVA, L.; FIGUEIREDO FILHO, D.; FERNANDES, A. The effect of lockdown on the COVID-19 epidemic in Brazil: evidence from an interrupted time series design. **Cadernos De Saude Publica**, v. 36, n. 10, p. e00213920, 2020.

SILVA, R. C. et al. Natural Products-Based Drug Design against SARS-CoV-2 Mpro 3CLpro. **International Journal of Molecular Sciences**, v. 22, n. 21, p. 11739, jan. 2021.

STADER, F.; BATTEGAY, M.; MARZOLINI, C. Physiologically-Based Pharmacokinetic Modeling to Support the Clinical Management of Drug-Drug Interactions With Bictegravir. **Clinical Pharmacology and Therapeutics**, v. 110, n. 5, p. 1231–1239, nov. 2021.

STADLER, K. et al. SARS--beginning to understand a new virus. **Nature Reviews. Microbiology**, v. 1, n. 3, p. 209–218, dez. 2003.

STODDARD, M. et al. Beyond the new normal: Assessing the feasibility of vaccine-based suppression of SARS-CoV-2. **PLOS ONE**, v. 16, n. 7, p. e0254734, 16 jul. 2021.

SU, S.; DU, L.; JIANG, S. Learning from the past: development of safe and effective COVID-19 vaccines. **Nature Reviews. Microbiology**, v. 19, n. 3, p. 211–219, mar. 2021.

TA, G. C. GHS Implementation to Strengthen Global Chemical Hazard Communication: Will We Ever Get There? **ACS Chemical Health & Safety**, v. 28, n. 3, p. 153–158, 24 maio 2021.

TALELE, T. T.; KHEDKAR, S. A.; RIGBY, A. C. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. **Current Topics in Medicinal Chemistry**, v. 10, n. 1, p. 127–141, 2010.

TALIC, S. et al. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. **BMJ**, v. 375, p. e068302, 18 nov. 2021.

TAN, C.-C. **SARS in Singapore-Key Lessons from an Epidemic**. [s.l: s.n.].

TAY, M. Z. et al. **The trinity of COVID-19: immunity, inflammation and intervention**. **Nature Reviews Immunology** Nature Research, , 1 jun. 2020.

TEN BRINK, T.; EXNER, T. E. Influence of Protonation, Tautomeric, and Stereoisomeric States on Protein–Ligand Docking Results. **Journal of Chemical Information and Modeling**, v. 49, n. 6, p. 1535–1546, 22 jun. 2009.

THOMPSON, M. G. Effectiveness of a Third Dose of mRNA Vaccines Against COVID-19–Associated Emergency Department and Urgent Care Encounters and Hospitalizations Among Adults During Periods of Delta and Omicron Variant Predominance — VISION Network, 10 States, August 2021–January 2022. **MMWR. Morbidity and Mortality Weekly Report**, v. 71, 2022.

TIMIRI, A. K.; SINHA, B. N.; JAYAPRAKASH, V. Progress and prospects on DENV protease inhibitors. **European Journal of Medicinal Chemistry**, v. 117, p. 125–143, jul. 2016.

TO, K. K. W. et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. **The Lancet Infectious Diseases**, v. 20, n. 5, p. 565–574, 1 maio 2020.

TONG, Z. D. et al. Potential presymptomatic transmission of SARS-CoV-2, Zhejiang Province, China, 2020. **Emerging Infectious Diseases**, v. 26, n. 5, p. 1052–1054, 1 maio 2020.

TORRES, P. H. M. et al. Key Topics in Molecular Docking for Drug Design. **International Journal of Molecular Sciences**, v. 20, n. 18, p. 4574, 15 set. 2019.

TREANOR, J. J. et al. Efficacy and Safety of the Oral Neuraminidase Inhibitor Oseltamivir in Treating Acute InfluenzaA Randomized Controlled Trial. **JAMA**, v. 283, n. 8, p. 1016–1024, 23 fev. 2000.

TROGEN, B.; OSHINSKY, D.; CAPLAN, A. **Adverse Consequences of Rushing a SARS-CoV-2 Vaccine: Implications for Public Trust**. **JAMA - Journal of the American Medical Association** American Medical Association, , 23 jun. 2020.

TROPSHA, A. Best Practices for QSAR Model Development, Validation, and Exploitation. **Molecular Informatics**, v. 29, n. 6–7, p. 476–488, 2010.

TROPSHA, A.; GRAMATICA, P.; GOMBAR, V. The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. **QSAR & Combinatorial Science**, v. 22, n. 1, p. 69–77, abr. 2003.

TSENG, A.; SEET, J.; PHILLIPS, E. J. The evolution of three decades of antiretroviral therapy: challenges, triumphs and the promise of the future. **British Journal of Clinical Pharmacology**, v. 79, n. 2, p. 182–194, fev. 2015.

TSENG, C.-T. et al. Immunization with SARS Coronavirus Vaccines Leads to Pulmonary Immunopathology on Challenge with the SARS Virus. **PLoS ONE**, v. 7, n. 4, p. e35421, 2012.

UEHA, R. et al. Oral SARS-CoV-2 Inoculation Causes Nasal Viral Infection Leading to Olfactory Bulb Infection: An Experimental Study. **Frontiers in Cellular and Infection Microbiology**, v. 12, p. 924725, 2022.

ULLAH, I. et al. Myths and conspiracy theories on vaccines and COVID-19: Potential effect on global vaccine refusals. **Vacunas (English Edition)**, v. 22, n. 2, p. 93–97, 2021.

VALLONE, A. et al. Antimalarial agents against both sexual and asexual parasites stages: structure-activity relationships and biological studies of the Malaria Box compound 1-[5-(4-bromo-2-chlorophenyl)furan-2-yl]- N -[(piperidin-4-yl)methyl]methanamine (MMV019918) and analogues. **European Journal of Medicinal Chemistry**, v. 150, p. 698–718, 2018.

VANGEEL, L. et al. Remdesivir, Molnupiravir and Nirmatrelvir remain active against SARS-CoV-2 Omicron and other variants of concern. **Antiviral Research**, v. 198, p. 105252, 1 fev. 2022.

VÁZQUEZ, J. et al. Merging Ligand-Based and Structure-Based Methods in Drug Discovery: An Overview of Combined Virtual Screening Approaches. **Molecules**, v. 25, n. 20, p. 4723, 15 out. 2020.

VERDONK, M. L. et al. Improved protein-ligand docking using GOLD. **Proteins: Structure, Function and Genetics**, v. 52, n. 4, p. 609–623, 1 set. 2003.

VERÍSSIMO, G. C. et al. HQSAR and random forest-based QSAR models for anti-T. vaginalis activities of nitroimidazoles derivatives. **Journal of Molecular Graphics & Modelling**, v. 90, p. 180–191, jul. 2019.

VERÍSSIMO, G. C. et al. The Brazilian compound library (BraCoLi) database: a repository of chemical and biological information for drug design. **Molecular Diversity**, 28 jan. 2022.

VERMA, D. K. et al. Potential Inhibitors of SARS-CoV-2 Main Protease (Mpro) Identified from the Library of FDA-Approved Drugs Using Molecular Docking Studies. **Biomedicines**, v. 11, n. 1, p. 85, jan. 2023.

V'KOVSKI, P. et al. Coronavirus biology and replication: implications for SARS-CoV-2. **Nature Reviews Microbiology**, v. 19, n. 3, p. 155–170, mar. 2021.

VOEGLER SMITH, A.; HALL, C. K. alpha-helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model. **Proteins**, v. 44, n. 3, p. 344–360, 15 ago. 2001.

WACHARAPLUESADEE, S. et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. **Nature Communications**, v. 12, n. 1, p. 972, 9 fev. 2021.

WAN, Y. et al. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. **Journal of Virology**, v. 94, n. 7, p. e00127-20, 17 mar. 2020.

WANG, M. et al. **Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro.** **Cell Research** Springer Nature, , 1 mar. 2020a.

WANG, W. et al. **Detection of SARS-CoV-2 in Different Types of Clinical Specimens.** **JAMA - Journal of the American Medical Association** American Medical Association, , 12 maio 2020b.

WANG, Z. et al. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. **Physical Chemistry Chemical Physics**, v. 18, n. 18, p. 12964–12975, 2016.

WARREN, G. L. et al. A Critical Assessment of Docking Programs and Scoring Functions. **Journal of Medicinal Chemistry**, v. 49, n. 20, p. 5912–5931, 5 out. 2006.

WARREN, T. K. et al. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. **Nature**, v. 531, n. 7594, p. 381–385, 2016.

WAYBRIGHT, T. J.; BRITT, J. R.; MCCLOUD, T. G. Overcoming problems of compound storage in DMSO: solvent and process alternatives. **Journal of biomolecular screening**, v. 14, n. 6, p. 708–15, jul. 2009.

WEINGARTL, H. et al. Immunization with Modified Vaccinia Virus Ankara-Based Recombinant Vaccine against Severe Acute Respiratory Syndrome Is Associated with Enhanced Hepatitis in Ferrets. **Journal of Virology**, v. 78, n. 22, p. 12672–12676, 2004.

WHO. **WHO Coronavirus (COVID-19) Dashboard.** Disponível em: <<https://covid19.who.int>>. Acesso em: 3 abr. 2022.

WINKLER, D. A. The role of quantitative structure--activity relationships (QSAR) in biomolecular discovery. **Briefings in bioinformatics**, v. 3, n. 1, p. 73–86, 2002.

WLODAWER, A. Rational approach to AIDS drug design through structural biology. **Annual Review of Medicine**, v. 53, p. 595–614, 2002.

WOJDYLA, J. A. et al. Papain-Like Protease 1 from Transmissible Gastroenteritis Virus: Crystal Structure and Enzymatic Activity toward Viral and Cellular Substrates. **Journal of Virology**, v. 84, n. 19, p. 10063–10073, out. 2010.

WOUTERS, O. J. et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. **Lancet**, v. 397, n. 10278, p. 1023–1034, 13 mar. 2021.

WU, C. et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. **Acta Pharmaceutica Sinica. B**, v. 10, n. 5, p. 766–788, maio 2020.

WU, Z. et al. The inhibition of Mpro, the primary protease of COVID-19, by Poria cocos and its active compounds: a network pharmacology and molecular docking study. **RSC Advances**, v. 11, n. 20, p. 11821–11843, 2021.

XIA, S. et al. Safety and immunogenicity of an inactivated SARS-CoV-2 vaccine, BBIBP-CorV: a randomised, double-blind, placebo-controlled, phase 1/2 trial. **The Lancet Infectious Diseases**, v. 0, n. 0, 15 out. 2020.

XU, X. et al. **Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission.** **Science China Life Sciences** Science in China Press, , 1 mar. 2020.

XU, Z. et al. SARS-CoV-2 impairs interferon production via NSP2-induced repression of mRNA translation. **Proceedings of the National Academy of Sciences**, v. 119, n. 32, p. e2204539119, 9 ago. 2022.

YAN, R. et al. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. **Science**, v. 367, n. 6485, p. 1444–1448, 27 mar. 2020.

YANG, H. et al. admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. **Bioinformatics**, v. 35, n. 6, p. 1067–1069, 15 mar. 2019.

YAP, C. W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. **Journal of Computational Chemistry**, v. 32, n. 7, p. 1466–1474, 2011.

YEAGER, C. L. et al. Human aminopeptidase N is a receptor for human coronavirus 229E. **Nature**, v. 357, n. 6377, p. 420–422, 4 jun. 1992.

YI, Y. et al. Natural triterpenoids from licorice potently inhibit SARS-CoV-2 infection. **Journal of Advanced Research**, v. 36, p. 201–210, 1 fev. 2022.

YU, I. T. S.; SUNG, J. J. Y. **The epidemiology of the outbreak of severe acute respiratory syndrome (SARS) in Hong Kong - What we do know and what we don't.** **Epidemiology and Infection** Cambridge University Press, , 2004.

ZANG, Y. et al. High-throughput screening of SARS-CoV-2 main and papain-like protease inhibitors. **Protein & Cell**, v. 14, n. 1, p. 17–27, 1 jan. 2023.

ZHANG, C.-H. et al. Potent Noncovalent Inhibitors of the Main Protease of SARS-CoV-2 from Molecular Sculpting of the Drug Perampanel Guided by Free Energy Perturbation Calculations. **ACS central science**, v. 7, n. 3, p. 467–475, 24 mar. 2021.

ZHANG, L. et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors. **Science**, v. 368, n. 6489, p. 409–412, 2020a.

ZHANG, W. et al. Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. **Emerging Microbes and Infections**, v. 9, n. 1, p. 386–389, 2020b.

ZHANG, Z. et al. Structure of SARS-CoV-2 membrane protein essential for virus assembly. **Nature Communications**, v. 13, n. 1, p. 4399, 5 ago. 2022.

ZHAO, M.-M. et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. **Signal Transduction and Targeted Therapy**, v. 6, n. 1, p. 1–12, 27 mar. 2021.

ZHOU, Y. et al. Nirmatrelvir-resistant SARS-CoV-2 variants with high fitness in an infectious cell culture system. **Science Advances**, v. 8, n. 51, p. eadd7197, 21 dez. 2022.

**APÊNDICE A – Atividade inibitória (%) da série de moléculas análogos dos *hits* (WRR)
contra M^{pro} de SARS-CoV-2**

Molécula	% inibição (10 µM^a)	% inibição (5 µM^a)	Molécula	% inibição (10 µM^a)	% inibição (5 µM^a)
129	0,0 ± 0,0	0,21 ± 0,17	182	0,00 ± 0,00	6,17 ± 0,79
130	0,0 ± 0,0	5,45 ± 4,45	183	0,00 ± 0,00	5,76 ± 3,08
139	0,0 ± 0,0	3,53 ± 2,88	193	0,81 ± 0,81	4,12 ± 1,52
145	0,0 ± 0,0	3,85 ± 3,14	195	0,00 ± 0,00	4,99 ± 3,57
146	0,0 ± 0,0	0,64 ± 0,52	196	0,54 ± 0,54	4,48 ± 0,7
148	0,8 ± 0,8	7,26 ± 5,93	197	0,00 ± 0,00	7,53 ± 4,52
150	0,0 ± 0,0	2,56 ± 2,09	198	0,00 ± 0,00	5,24 ± 1,48
152	2,98 ± 0,29	4,51 ± 0,85	200	0,45 ± 0,45	4,18 ± 3,15
153	0,08 ± 0,08	3,42 ± 2,79	201	0,00 ± 0,00	7,38 ± 1,98
157	0,0 ± 0,0	0,00 ± 0,00	202	0,09 ± 0,09	5,61 ± 3,11
158	0,0 ± 0,0	1,92 ± 1,57	209	0,00 ± 0,00	3,1 ± 1,76
159	0,08 ± 0,08	2,24 ± 1,83	203	0,00 ± 0,00	6,19 ± 3,21
160	0,92 ± 0,92	15,47 ± 3,77	229	0,00 ± 0,00	4,39 ± 1,37
161	0,0 ± 0,0	1,6 ± 1,31	230	0,36 ± 0,36	3,97 ± 3,97
163	0,00 ± 0,00	0,00 ± 0,00	235	0,97 ± 0,97	1,01 ± 1,01
164	0,25 ± 0,25	0,00 ± 0,00	236	0,43 ± 0,43	0,43 ± 0,43
166	1,27 ± 0,8	5,24 ± 4,27	241	0,00 ± 0,00	4,71 ± 2,27
168	4,05 ± 1,27	5,13 ± 4,19	261	0,00 ± 0,00	3,22 ± 0,52
169	12,53 ± 0,26	20,08 ± 0,18	267	0,00 ± 0,00	1,36 ± 0,84
170	0,00 ± 0,00	0,00 ± 0,00	281	0,00 ± 0,00	1,66 ± 0,88
171	0,00 ± 0,00	0,00 ± 0,00	286	0,00 ± 0,00	1,29 ± 1,19
172	0,00 ± 0,00	1,28 ± 1,05	287	0,34 ± 0,34	1,22 ± 0,92
173	0,36 ± 0,36	3,33 ± 3,22	288	1,41 ± 0,73	0,58 ± 0,58
174	1,59 ± 1,59	5,98 ± 3,39	292	0,00 ± 0,00	2,28 ± 0,61
176	0,00 ± 0,00	8,74 ± 5,25	293	1,62 ± 1,15	1,45 ± 0,96
177	7,45 ± 0,6	34,24 ± 2,9	297	1,19 ± 0,83	0,93 ± 0,83
179	2,17 ± 2,17	5,81 ± 3,22	299	2,29 ± 1,25	2,53 ± 1,35
180	96,75 ± 1,04	98,70 ± 1,04	300	5,96 ± 2,9	1,26 ± 0,63
181	100,00 ± 0,00	100,00 ± 0,00	301	1,7 ± 0,96	1,56 ± 0,79

^a O percentual de inibição é representado como a média e o erro do desvio-padrão calculados entre um ensaio independente em triplicata. O erro é apresentado como a razão entre o desvio-padrão e a raiz quadrada do número de replicatas/leituras (n = 3).

A.1. Atividade inibitória (%) da série de moléculas análogos dos *hits* (WRR) contra M^{pro} de SARS-CoV-2 (continuação)

Molécula	% inibição (10 μM^a)	% inibição (5 μM^a)	Molécula	% inibição (10 μM^a)	% inibição (5 μM^a)
302	11,02 \pm 2,36	5,94 \pm 3,4	366	2,32 \pm 2,32	4,4 \pm 2,22
303	0,34 \pm 0,34	2,10 \pm 1,07	368	1,19 \pm 0,72	4,71 \pm 1,3
304	1,02 \pm 0,59	1,65 \pm 3,38	370	6,02 \pm 3,56	6,45 \pm 0,35
305	2,54 \pm 1,46	3,19 \pm 2,18	372	4,70 \pm 2,39	7,07 \pm 1,19
306	0,59 \pm 0,59	0,00 \pm 0,00	376	4,33 \pm 2,51	5,81 \pm 2,9
307	0,00 \pm 0,00	0,93 \pm 0,47	377	4,37 \pm 2,22	6,86 \pm 1,87
308	2,17 \pm 1,44	3,33 \pm 1,89	378	8,82 \pm 4,17	4,41 \pm 1,93
309	0,76 \pm 0,52	3,06 \pm 1,32	379	6,55 \pm 3,96	9,86 \pm 0,71
310	2,72 \pm 1,38	4,14 \pm 2,14	381	11,42 \pm 3,67	14,3 \pm 1,03
311	3,97 \pm 1,94	1,46 \pm 1,46	382	13,15 \pm 2,5	10,53 \pm 5,3
312	2,38 \pm 2,38	7,35 \pm 4,02	384	19,24 \pm 4,18	23,88 \pm 4,43
315	4,82 \pm 4,58	8,09 \pm 3,76	385	51,55 \pm 1,38	39,80 \pm 1,51
316	4,94 \pm 2,71	7,3 \pm 0,51	389	3,65 \pm 3,65	5,46 \pm 1,43
317	53,97 \pm 1,34	31,67 \pm 1,34	390	11,07 \pm 5,85	5,25 \pm 0,3
318	30,16 \pm 0,88	15,68 \pm 2,74	391	7,11 \pm 3,63	9,65 \pm 1,2
319	4,07 \pm 4,07	3,89 \pm 3,66	392	10,04 \pm 3,27	5,78 \pm 0,43
320	4,71 \pm 2,54	3,46 \pm 1,11	397	6,08 \pm 3,1	11,49 \pm 1,82
333	1,0 \pm 0,87	4,47 \pm 1,62	398	5,26 \pm 2,72	5,36 \pm 2,0
342	6,93 \pm 3,5	5,48 \pm 2,49	400	3,30 \pm 2,11	9,92 \pm 3,67
343	0,67 \pm 0,67	5,99 \pm 1,82	412	1,07 \pm 0,68	1,19 \pm 0,89
346	5,37 \pm 2,77	5,38 \pm 2,22	413	0,00 \pm 0,00	1,61 \pm 1,61
347	9,33 \pm 5,5	9,48 \pm 3,57	414	0,37 \pm 0,37	1,35 \pm 1,18
350	0,83 \pm 0,83	2,97 \pm 2,97	415	1,3 \pm 1,3	2,85 \pm 1,41
352	3,59 \pm 1,98	1,95 \pm 1,22	416	1,58 \pm 0,88	1,94 \pm 1,77
353	3,29 \pm 1,85	4,97 \pm 1,12	417	0,00 \pm 0,00	2,30 \pm 0,99
356	1,74 \pm 1,09	5,82 \pm 1,6	418	2,97 \pm 1,68	8,01 \pm 1,33
358	1,39 \pm 1,39	5,33 \pm 3,03	419	0,00 \pm 0,00	3,06 \pm 1,53
360	4,23 \pm 2,68	8,82 \pm 0,2	431	0,00 \pm 0,00	2,93 \pm 2,11
361	2,69 \pm 2,05	3,72 \pm 1,16	436	1,57 \pm 1,57	1,77 \pm 1,18

^a O percentual de inibição é representado como a média e o erro do desvio-padrão calculados entre um ensaio independente em triplicata. O erro é apresentado como a razão entre o desvio-padrão e a raiz quadrada do número de replicatas/leituras (n = 3).

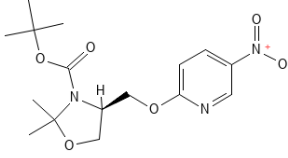
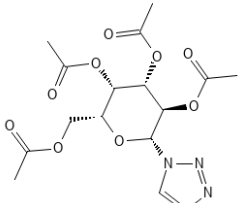
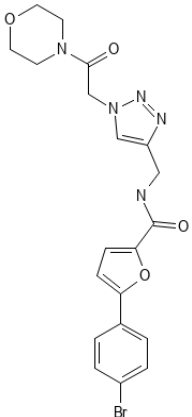
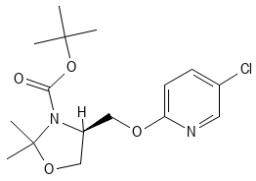
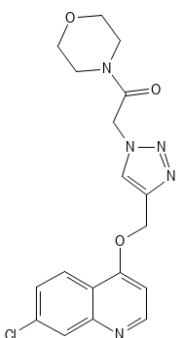
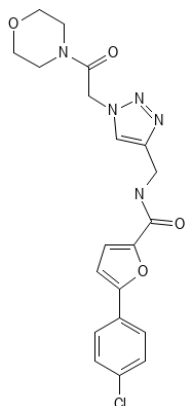
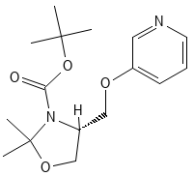
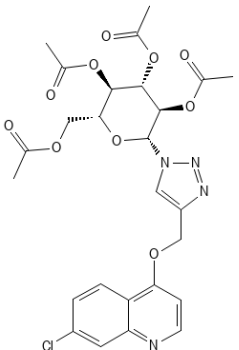
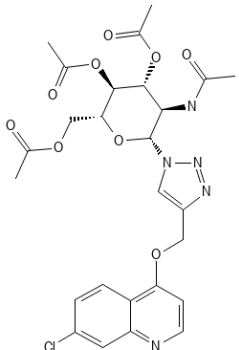
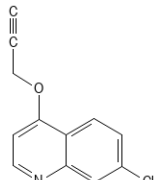
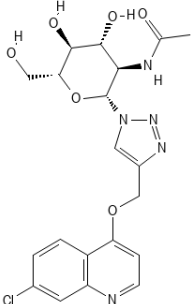
A.1. Atividade inibitória (%) da série de moléculas análogos dos *hits* (WRR) contra M^{pro} de SARS-CoV-2 (continuação)

Molécula	% inibição (10 μ M ^a)	% inibição (5 μ M ^a)	Molécula	% inibição (10 μ M ^a)	% inibição (5 μ M ^a)
438	52,55 \pm 3,61	45,09 \pm 0,53	463	2,61 \pm 2,61	1,68 \pm 0,85
439	3,99 \pm 2,61	8,45 \pm 1,28	464	1,23 \pm 1,23	2,84 \pm 2,74
441	5,75 \pm 2,72	7,52 \pm 0,76	465	0,84 \pm 0,84	2,91 \pm 2,31
442	7,55 \pm 4,94	9,4 \pm 1,08	466	2,53 \pm 2,53	3,18 \pm 3,08
443	2,87 \pm 1,4	4,03 \pm 2,13	467	3,87 \pm 2,44	3,17 \pm 3,17
444	1,94 \pm 1,94	8,69 \pm 2,48	468	2,95 \pm 2,25	3,36 \pm 2,07
445	5,07 \pm 2,54	7,38 \pm 1,8	477	0,23 \pm 0,23	1,36 \pm 0,97
446	0,00 \pm 0,00	1,85 \pm 1,18	478	0,00 \pm 0,00	4,05 \pm 2,97
447	5,26 \pm 2,64	4,03 \pm 2,17	479	0,23 \pm 0,23	1,53 \pm 1,53
453	0,00 \pm 0,00	5,11 \pm 1,66	482	2,01 \pm 1,66	3,97 \pm 3,46
454	2,77 \pm 1,38	3,95 \pm 1,29	500	1,15 \pm 1,15	4,02 \pm 4,02
459	0,00 \pm 0,00	4,23 \pm 2,33	605	5,65 \pm 1,18	8,72 \pm 4,84
460	0,00 \pm 0,00	2,38 \pm 2,05	-	-	-
Nirmatrelvir	95,79 \pm 2,12 (100 nM)	93,43 \pm 0,13 (100 nM)	Nirmatrelvir	95,78 \pm 2,12 (100 nM)	92,01 \pm 0,07 (100 nM)

^a O percentual de inibição é representado como a média e o erro do desvio-padrão calculados entre um ensaio independente em triplicata. O erro é apresentado como a razão entre o desvio-padrão e a raiz quadrada do número de replicatas/leituras (n = 3).

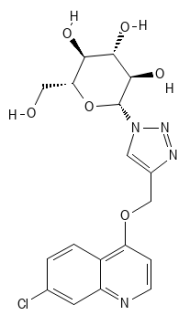
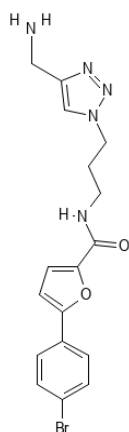
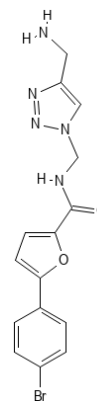
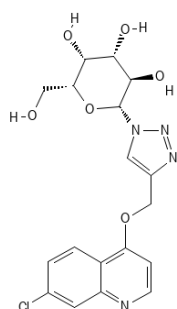
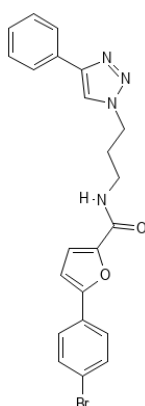
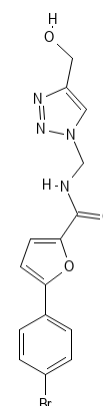
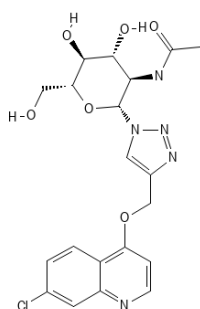
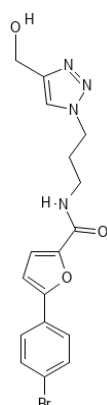
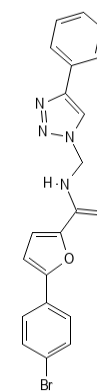
APÊNDICE B – Moléculas (BraCoLi) selecionadas no modelo de HQSAR

B.1. Moléculas oriundas da seleção por PubChem

PubChem (código BraCoLi)		
		
QF1 (BR010480)	QF6 (BR020100)	QF9 (BR020117)
		
QF2 (BR010479)	QF7 (BR020098)	QF10 (BR020119)
		
QF3 (BR010481)	QF5 (BR020099)	QF8 (BR020101)
		
QF4 (BR020097)	QF11 (BR020127)	

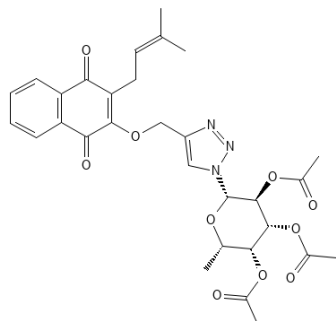
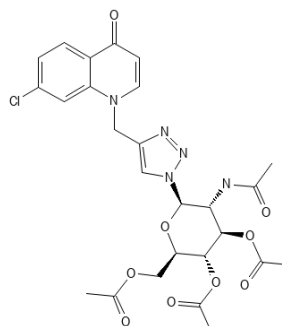
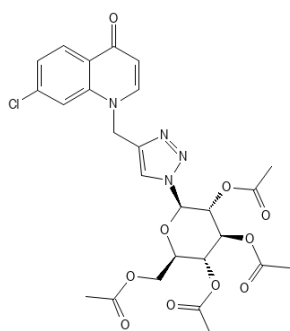
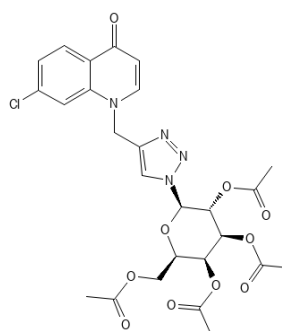
B.1. Moléculas oriundas da seleção por PubChem (continuação)

PubChem (código BraCoLi)

**QF12****QF15****QF18****QF13****QF16 (BR020113)****QF19****QF14****QF17****QF20**

B.2. Moléculas oriundas da seleção por AtomPairs2DCount

AtomPairs2DCount (código BraCoLi)

**QF21 (BR010185)****QF22****QF23****QF24**

APÊNDICE C – Alinhamento de seqüências das proteases de SARS-CoV-2

C.1. Alinhamento das seqüências de M^{pro} de SARS-CoV-2 e dos demais coronavírus

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```
TGEV          SGLRKKMAQPSGLVEPCIVRVSYGNNVLNGLWLGDVEICPRHVIASDTRRV-INYENEMSS
HCoV-229E     AGLRKKMAQPSGFVEKCVVRVCYGNLVLNGLWLDIVYVCPRHVIASNTTSA-IDYDHEYSI
HCoV-NL63     SGLRKKMAQPSGCVRCVVRVCYGNLVLNGLWLDIVYVCPRHVIAPSTTVL-IDYDHAYST
SARS-CoV-2    SGFRKMAFSPGKVEGCMVQVTCGTTLNGLWLDVVYVCPRHVICTSEDMLNPNYEDLLIR
SARS-CoV      SGFRKMAFSPGKVEGCMVQVTCGTTLNGLWLDVVYVCPRHVICTAEDMLNPNYEDLLIR
HCoV-HKU1     SGIVKMVSPSTKIEPCIVSVTYGSMNLNGLWLDKVVYVCPRHVICSSNMNEPDYSALLCR
MHV           SGIVKMVSPSTKIEPCIVSVTYGSMNLNGLWLDKVVYVCPRHVICSSADMTDPDYPNLLCR
MERS-CoV      SGLVKMSHPSGDVEACMVQVTCGSMNLNGLWLDNTVWVCPRHVMCPADQLSDPNYDALLIS
HKU4          SGLVKMSAPSGAVENCIVQVTCGSMNLNGLWLDNTVWVCPRHIMCPADQLSDPNYDALLIS
:*: ** *.: :* *: * * . **:*: * * * * : : *
```

```
TGEV          VRLHNFVSK---NNVFLGVVSARYKGVNLVVKVNPNTPEHKFKSIKAGESFNILAC
HCoV-229E     MRLHNFVSIIS---GTAFLGVVGATMHGVTLKIKVSQTNMHTPRHSFRTLKSGEGFNILAC
HCoV-NL63     MRLHNFVSSH---NGVFLGVVGVTMHGSLRIRKVSQSNVHTPKHVFKTLKPGDSFNILAC
SARS-CoV-2    KSNHNFVQA---GNVQLRVIIGHSMQNCLLRLKVDTSNPKTPKYKRVRIQPGQTFVLAAC
SARS-CoV      KSNHNFVQA---GNVQLRVIIGHSMQNCLLRLKVDTSNPKTPKYKRVRIQPGQTFVLAAC
HCoV-HKU1     VTLGDFTIMS---GRMSLTVVSYQMGGCQLVLTLSLQNPYTPKYTFGNVVKPGETFTVLAAC
MHV           VTSSDFCVMS---GRMSLTVMSYQMGGCQLVLTLSLQNPYTPKYTFGNVVKPGETFTVLAAC
MERS-CoV      MTNHSFVQKHIQGANLRVVGHAMQGTLLKLTVDVANPSTPAYFTTVPKGAASFVLAAC
HKU4          KTNHSFIVQKHIGAQANLRVVAHSMVGLLKLTVDVANPSTPAYFTTVPKGAASFVLAAC
.*: * : * * : . * * : * * : * * : * * : * * : * * : * * : * * : *
```

```
TGEV          YEGCPGSVYGVNMRSQGTIKGSFIAGTCGSVGYVL-ENGILYFVYMHLELNGSHVGSN
HCoV-229E     YDGAQGVFGVNMRTNWTIRGSFINGACGSPGYNL-KNGEVFVYMHQIELGSGSHVGSN
HCoV-NL63     YEGIASGVFVGNLRTNFTIKGSFINGACGSPGYNVRNDGTVEFCYLHQIELGSGAHVGSN
SARS-CoV-2    YNGSPSGVYQCAMRPNFTIKGSFLNGSCGSGVFNIDYDCVSFCYMHMELPTGVHAGTD
SARS-CoV      YNGSPSGVYQCAMRPNHTIKGSFLNGSCGSGVFNIDYDCVSFCYMHMELPTGVHAGTD
HCoV-HKU1     YNGRPPQAFHVTMRSSYTIKGSFLCGSCGSGVGYL-TGDSVKFVYMHQELSTGCHTGTD
MHV           YNGRPPQAFHVTLRSSHTIKGSFLCGSCGSGVGYL-TGDSVRFVYMHQELSTGCHTGTD
MERS-CoV      YNGRPTGTFTVVMRPNYTIKGSFLCGSCGSGVGYK-EGSVINFCYMHQELANGTHTGSA
HKU4          YNGKPTGVFTVNLRHNSTIKGSFLCGSCGSGVGYTE-NGGVINFCYMHQELANGTHTGSS
*.* . . . : * . ** : * : * * : * : * * : * : * * : * * : * * : *
```

```
TGEV          FEGEMYGGEYDQPSMQLGNTVMSSDNVAVFLYAALINGERWVFNNTSMLSLEYNTWAKT
HCoV-229E     FDGVMYGGFEDQPNLQVESANQMLTVNVVAVFLYAAILNGCTWLLKGEKLFVEHYNEWAQA
HCoV-NL63     FTGSVYGNFDDQPSLQVESANLMLSDNVAVFLYAAILNGCRWLLCSTRVNVDFNEWAMA
SARS-CoV-2    LEGNFYGFVDRQTAQAAGTDTTITVNLVLAWLAAVINGDRWFLNRFTTTLNDFNLVAMK
SARS-CoV      LEGKFYGFVDRQTAQAAGTDTTITLNVLAWLAAVINGDRWFLNRFTTTLNDFNLVAMK
HCoV-HKU1     FTGNFYGPRDAQVQVQLPVKDYVQTVNVIAWLAAILNCAWFVQNDVCSDEDFNVWAMA
MHV           FSGNFYGPYRDAQVQVQLPVQDYQTVNVVAVWLAAIFNRCNWFVQSDSCSLEEFNVWAMT
MERS-CoV      FDGTYMGAFMDKQVHQVQLTDKYCSVNVVAVWLAAILNGCAWFVKNRSTVSVSFNEWALA
HKU4          FDGVMYGAFFDKQTHQLQLTDKYCTINNVAVWLAAVINGCKWVFKPTRVGIPTYNEWALS
: * . ** : * * : : * * : * * : * * : * * : * * : * * : * * : *
```

```
TGEV          NSFTELSST--DAFSMLAAKTGQSVKLLDSIVRLNK-GFGGRTILSYGSLCDEFTPEV
HCoV-229E     NGFTAMNGE--DAFSILAAGTGVVERLLHAIQVLLN-GFGGKQILGYSSLNDEFINEV
HCoV-NL63     NGYTVSSV--ECYSILAAGTGVSEQLLASIQHLHE-GFGGKQILGYSSLCDEFTLAEV
SARS-CoV-2    YNYPELTQDHDVILGPLSAQTGIAVLDMCASLKELLQNGMNGRTILGSALLEDEFPPFDV
SARS-CoV      YNYPELTQDHDVILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTIIEDEFPPFDV
HCoV-HKU1     NGFSQVKAD--LVLDALASMTGVSITLLAAIKRLYM-GFQGRQILGSCTFEDELAPSDV
MHV           NGFSSIKAD--LVLDALASMTGVTVEQLLAAIKRLHS-GFQGKQILGSQVLEDETPSDV
MERS-CoV      NQFTEFVGT--QSVDMIAVKTGVAVIEQLLYAIQQLYT-GFQGKQILGSTMLEDEFPPEDV
HKU4          NQFTEFVGT--QSIDMLAHRGTGVSVEQLAAIQSLHA-GFQGKQILGSTLEDEFPPDDV
: . * : * * : : * * : * * : * * : * * : * * : * * : * * : *
```

```
TGEV          IRQMYGVNLQ
HCoV-229E     VKQMFVNLQ
HCoV-NL63     VKQMYGVNLQ
SARS-CoV-2    VRQCSGVTFQ
SARS-CoV      VRQCSGVTFQ
HCoV-HKU1     YQQLAGVKLQ
MHV           YQQLAGVKLQ
MERS-CoV      NMQIMGVVMQ
HKU4          NMQVMGVVMQ
* * * *
```

C.2. Alinhamento das sequências de PL^{pro} de SARS-CoV-2 e dos demais coronavírus

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```

pdcov      ---LNRTIVRITTTDYDQVTTKALTPQGVLEANL  FDGEDF--V--QEPKPGQIYLEV
idv        --TKQKTIYLTEDGVKYRSIVLKPGLSLG-QFGQVYAKNKIV--FTADDVEDKEILYVP
tgev
mers----- QLTIEVLVTVDGVNFRVTVLNNKNTYRSQLGCVFFNGADISDIPDEKQNGHSLYLA
sars      --EVKTIKVFVTTVDNLTNLHTQLVDMSTMTYGQQFGPTYLDGADVTKIKPHVNHGKTFEVL
cov2      LREVRTIKVFVTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPHNSHEGKTFYVL
    
```

```

pdcov      TEEVQNQAKEL-----DLNLQQYCVYLKCTHH--KWVVSRTNGLMHLKQKDNCFVSAGV
idv        TTD-----KSILEYYG-LDAQKYVIYLQTLAQ--KWNVQYRDNFLILEWRDGNCWISSAI
tgev      -----SLPP-----FKTTNLNGKIIILKQGDNNCWINACC
mers      DNLTADETKALKELYGPVDPFTLHRFYSLKAHVHGKMWVCDKVRSLKLSDNNCYLNAVI
sars      PSDDTLRSEAF-EYYHTLDESFLGRYMSALNHTKKWKFPQVGGLTSIKWADNNCYLSSVL
cov2      PNDDTLRVEAF-EYYHTTDPSEFLGRYMSALNHTKKWKYPQVNGLTSIKWADNNCYLATAL
          : . : : * . * : : :
    
```

```

pdcov      NLFQNT-AYQFR-PAIDALYREYLNNGNPNRFVAWIYASTNRRVGMGCPQQVISLLVSNS
idv        VLLQAA-KIRFK-GFLTEAWAKLLGGDPTDFVAVCYASCTAKVGFSDANWLLANLAEHF
tgev      YQLQAF-DF-FN  NEAWKFKKGDVMDVFNLCYAATTLARGHSGDAEYLLELMLNDY
mers      MTLDLLKDIKVFIPALQHAFMKHKGGSTDFIALIMAYGNCTFGAPDDASRLHTVLAKA
sars      LALQQL-EVKFNAPALQEAYYRARAGDAANFCALILAYSNKTVGELGDVRETMTHLLQHA
cov2      LTLQQI-ELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELGDVRETMSYLFQHA
          : : * : * : * . * . : :
    
```

```

pdcov      DAAFSATTA-----CCNTYFNHTGVISVAR  EYDPIQPKV-YCMKCDVWTPFTP
idv        DADYTN AFLKRVSCNCGIKSYELRGL EACIQPVRATNLLHFKTQYSNCPCTCGANNTDEV
tgev      ---STAKIVLAAKCG-CGEKEIVL  ERAVF-KLTPLKESFNYGVCDCMQVNTCRFLS
mers      ELCCSARMVWREWCNVCGIKDVVLQGLKACCY-VGVQTVEDLRARMTYVQCGERHRQL
sars      NLESAKRVL-NVVCKHCGQKTTTLTGVEAVMY-MGTLSDNLKTGVSIPCVCGRDATQYL
cov2      NLDSCKRVL-NVVCKTCGQQQTTLKGV EAVMY-MGTLSEYQFKKGVQIPCTCGKQATKYL
          * : :
    
```

```

pdcov      QSGKGAVAIGTSADEPTGPAIKFAA  AHCWYTNKKTVNGYDTKANVVAITYHRFD-
idv        IEASLPYLLL FATDGPATVDCDEDA----VGTVVFVG-STNSGHCYTQAAAGQA  FDN
tgev      VEGSGVFDHILSKQTP EAMFVKP---VMHAVYTG-TTQNGHYMVDIEHG--YCVDG
mers      VEHTTPWLLLSGT--PNEKLVTTSTAPDFVAFNVFQGIETAVGHYVHARLKGGLILKFDS
sars      VQQESSFVMSAP--PAEYKLQQGT  FLCANEYTG-NYQCGHYTHITAKETL-YRIDG
cov2      VQQESSPFVMSAP--PAQYELKHGT  FTCASEYTG-NYQCGHYKHITSKETL-YCIDG
          . . . . : . . *
    
```

```

pdcov      ----VPKPQLVEDVVALPTKNDFE-----
idv        LAKDRKFGKKSPIYITAMYTRFAFKNETSLPVAKQ
tgev      MGI-KPLKKRCYTSTLFINANVMT-----
mers      GTVSKTSDWKCKVTDVLFPGQKYSSDCN-----
sars      AHLTKMSEYKGPVTDVDFYKETSyttTIK-----
cov2      ALLTKSSEYKGPITDVFYKENSyttTIKP-----
          :
    
```

APÊNDICE D – Percurso, atividades e produções acadêmicas

Durante os quase dois anos e seis meses de doutorado (08/2020 a 01/2023), sendo dois anos e um mês efetivamente atuando no projeto (01/2021 a 01/2023), diversas atividades e produções acadêmicas foram desenvolvidas. Ao longo do percurso, destacam-se:

1. Cumprimento de 35 créditos obrigatórios em disciplinas, sendo dois relacionados ao treinamento didático obrigatório pela CAPES;
2. Participação como voluntário na equipe de diagnóstico Rede COVID-19, prestando auxílio no diagnóstico de SARS-CoV-2 por RT-qPCR na UFMG (04/2021 a 11/2021);
3. Co-supervisão dos alunos de iniciação científica Diego Lanza Dias (Ciências Biológicas, UFMG) e Christian Souza Moreira (Farmácia, UFMG);
4. Co-orientação de trabalho de conclusão de curso de Desirée Cristina dos Santos Jacó (Bacharel em Farmácia, UFMG);
5. Participação em cursos: Estrutura, evolução e análise de genomas virais (SBV); Métodos de Docking Receptor-Ligante e Virtual Screening (LNCC); Dinâmica Molecular Básica (LNCC); Protocolos de Manejo Clínico do Coronavírus (COVID-19) (UFSC); Medidas de Proteção no Manejo da COVID-19 (UFSC); Manejo clínico da COVID-19 na Atenção Especializada (UFMG); Divulgação Científica (FIOCRUZ); Prevenção, Controle e Vigilância Epidemiológica das IRAS (Oswaldo Cruz);
6. Participação em eventos científicos remotos e presenciais: XXXI e XXXII Congresso Brasileiro de Virologia; VII e VIII Simpósio de Microbiologia da UFMG; I Semana da Biologia Sintética; 1º Simpósio Acadêmico De Evidências Clínicas na Covid-19; 1º Congresso Brasileiro de Evidências Clínicas na Covid-19; 1º Simpósio Mitos e Verdades Sobre a Covid-19; 2º Congresso Brasileiro de Evidências Clínicas na Covid-19; 2º Simpósio Mitos e Verdades sobre a Covid-19; 4º Encontro de Química Medicinal e Desenvolvimento de Fármacos; *6th Brazilian Student Council Symposium: Omics and Data Science*; II divulGAÇÃO: Jornada de Divulgação Científica; I Simpósio de Microbiologia de Rondônia: Saúde, Ambiente e Inovação; Epidemiologia de viroses emergentes, com ênfase na disseminação de SARS-CoV-2 e suas variantes no mundo;

Guia para publicar em *journals* de alto impacto em Microbiologia; SIG 2021: Machine Learning in Bioinformatics; X Escola de Modelagem Molecular em Sistemas Biológicos; SIG: DRUG-DESIGN 2020;

7. Participação em bancas examinadoras: Beatriz Murta Rezende Moraes (Graduação em Farmácia, UFMG); Anna Letícia Teotônio (Graduação em Farmácia, UFMG); Isabela Conceição Duarte (Graduação em Farmácia, UFMG); Laís Dias Ribeiro (Graduação em Biomedicina, UFMG);
8. Aulas e seminários: Variantes de SARS-CoV-2 entre vacinas contra COVID-19: uma questão de tempo? (Graduação em Biomedicina e Farmácia, Izabela Hendrix); Dengue e outras arboviroses: aspectos gerais e clínico-laboratoriais (Especialização em Microbiologia, UFMG);
9. CAPES/PrInt: Programa de Doutorado Sanduíche no Exterior (PDSE) desenvolvendo parte do projeto de doutorado por seis meses na University of California, San Diego (UCSD), sob supervisão dos Dr. Anthony J. O'Donoghue e Dr. Conor R. Caffrey;
10. Premiações: Melhor apresentação oral no SIG: DRUG-DESIGN 2020;
11. Manuscritos: Durante o doutorado foram publicados nove artigos em periódicos científicos internacionais, sendo dois de primeira autoria e cinco como segundo autor (anexos a seguir). No presente momento, outros doze trabalhos estão em diferentes níveis de preparação, sendo três relacionados a este projeto e três a serem submetidos ainda em março/abril de 2023.