UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

JOAN PETRUS OLIVEIRA LIMA

IN SILICO INHIBITION OF SARS-COV-2 AND CHIKV BY PHENOL AND CHROMONE DERIVATIVES

FORTALEZA
2024

JOAN PETRUS OLIVEIRA LIMA

IN SILICO INHIBITION OF SARS-COV-2 AND CHIKV BY PHENOL AND CHROMONE DERIVATIVES

Doctoral thesis submitted to the Programa de Pós-graduação em Química of the Universidade Federal do Ceará as a requirement for the degree of doctor in chemistry.
Concentration area: Chemistry.

Supervisor: Pierre Basílio Almeida Fechine.

Co-supervisor: Rafael Melo Freire.

FORTALEZA
2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

L698i Lima, Joan Petrus Oliveira.
In silico inhibition of SARS-CoV-2 and CHIKV by phenol and chromone derivatives / Joan Petrus Oliveira Lima. – 2024.
119 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Química, Fortaleza, 2024.
Orientação: Prof. Dr. Pierre Basílio Almeida Fechine.
Coorientação: Prof. Dr. Rafael Melo Freire.

1. SARS-CoV-2. 2. CHIKV. 3. Daldinia sp. 4. Modelagem molecular. I. Título.

CDD 540

JOAN PETRUS OLIVEIRA LIMA

IN SILICO INHIBITION OF SARS-COV-2 AND CHIKV BY PHENOL AND CHROMONE DERIVATIVES

Tese de doutorado apresentada ao programa de Pós-graduação em Química na Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Química.
Área de concentração: Química.

Aprovada em: 25/01/2024.

BANCA EXAMINADORA

_____
Prof. Dr. Pierre Basílio Almeida Fechine
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Emmanuel Silva Marinho
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Pedro de Lima Neto
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Claudenilson da Silva Clemente
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Aluísio Marques da Fonseca
Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB)

I dedicate to my family, professors and friends, that contributed, direct or indirectly, to this work and my human and professional development. To everyone involved in the fight against and who fought COVID-19. My condolences to those who lost a loved one and for those that passed by, may humanity never have to suffer the same again.

ACKNOWLEDGEMENTS

To my family that gave me support, my mother Maria Claudete, my father Josenir Alcântara, my stepfather Paulo Mosânio (that may rest in peace), my brothers Marcus Rodrigo and Cassius Iordano and my sister Isa Crisna and his fiancé Miguel. Thanks for the patience and care.

To my beloved wife, Camila Neiva, which helped me to be my better self, trusts me, filled my worst days with joy and took care of me when I needed the most during this long arduous work that you contributed so much, for your patience, comprehension, support, I am very grateful and happy to share my days and conquests (including this) at your side. Also, to our cats, Rhaenyra and Polly, that are our little daily happiness with fur that we love and care.

To my supervisor and friend, professor Pierre Basílio Almeida Fechine, from whom I learned the most to be a PhD and professor, an example to me and many other orientates due to have a high knowledge and human treat, that I admire and replicate in my way, with my students and orientates. For your patience, guidance, teachings, friendship, support and never give up on me no matter the difficulties, I am grateful.

To my co-supervisor, professor Rafael, for the teachings, conversations, support, patience and development, I am grateful.

To professor Emmanuel Marinho, the person responsible for the trail I am going in this work, that I always admired as a great researcher and professor, that demands hard work but its tender, professional virtues that are rare and difficult to conciliate, and I Admire. For the patience, the hard work and support besides any problem that we had, I am grateful.

To professor Samuel, that I always saw as a close example to me, for his many contributions in my qualification and for the friendship and lessons from a great researcher, I am grateful.

To professor Claudenilson, for the friendship, support, guidance, contributions to this work in the qualification and patience, I am grateful.

To professor Men de Sá, from whom I learned and developed as a professional, contributed to my teacher identity, for the friendship, care and support, I am grateful

To professors Sávio Macambira, Hugo Leonardo, Rita Mikaela, Sérgio Matos, Jorge, Rinaldo and Francisco Assis, from the Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), for the many contributions to my formation, in special professor Sávio, that taught my first steps in programing, that helped this work. I am grateful.

To my professors in the Universidade Federal do Ceará – UFC, my great admiration for the professionals that had encouraged me and my colleagues to be the best in our fields, my academic and professional evolution through hard work and the support of yours, I owe to professors Izaura Cirino, Luiz França, Paulo Naftali, Alcinéia Oliveira, Goretti Vasconcelos, Nágila Ricardo, Maria Conceição, Marcos Mattos, Mary Anne, Belmino Romero, Giselle Lopes, Selma Mazzetto, Cristiane Oliveira. Thanks for the patience, teachings and many contributions.

To Matheus Rocha, Victor Moreira, Hélcio, Aluísio, Caio, Gabrielle, Emanuelle and professor Pedro de Lima Neto, my colleagues and friends that published together and helped this work, I have many thanks due to solve my questions, helped me trail this area and for the friendship.

To my friends Filipe, Helem, Allex, Nicholas, Isaac, Sarah, Ana Lima, Nara, Fernanda, Lucas, Viviane, Levi Nogueira, Vanessa, Filipe, Waleska, Matheus, Emanuel, Paula Bianca, Lucas Fernandes, Diego, for our friendship that I am grateful.

To Elayne, Victor Moreira, Alvernes, Lorena, Alinne, Vanessa Abreu and Vanessa Abreu, Andreza, Fabrizia, for all the support, friendship and guidance. I am grateful.

To my friends, colleagues and ex-students from Limoeiro do Norte — CE, from the Universidade Estadual do Ceará (UECE), campus Faculdade de Filosofia Dom Aureliano Matos (FAFIDAM) Damião, Dalila, Cosmo, Matheus, Laura, Bruna, Samara, Tallia, Amanda, Raiane, João Paulo, Celena, Raimundo Nonato, Anthony, Osmar Júnior, Gustavo, Marcelo, Miquéias, Robson, Margarida, Zilvanir, Lailton, Júlio, Humberto, Sérgio, Érika, Rondinelle, Ranulfo, Brenna, Diego, Celene, Adriana, Alex, Hélio, Robson Maia, Miguel, Robson Sanábio, Salomão, Wager, Jean and among other students and professors, for the friendship, help, care and professional development, that I am very proud of my ex-students. I am grateful to all of you.

To everyone that helped and contributed for the realization of this work.

To the Centro Nacional de Processamento de Alto Desempenho da UFC (CENAPAD-UFC), that was essential for this work.

To the Programa de Pós-graduação em Química, Universidade Federal do Ceará and Universidade Estadual do Ceará, for all the infraestructure and public high quality education and to the scholarship of CAPES in the emergentil actions for COVID-19.

This work was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance code 001.

RESUMO

Em um novo contextos, como o das mudanças climáticas e a expansão de doenças endêmicas tropicais, a exemplo da recente pandemia de COVID19, a desinformação e a não aderência ao método científico pode potencializar os impactos socio-econômicos da situação, o que justifica a relevância de novas formas de agilizar pesquisas na área. O método tradicional para desenvolvimento de fármacos é por tentativa e erro, o que é demorado e custoso, portanto não é o ideal dentro do novo cenário da saúde mundial. Uma forma de afunilar e filtrar possíveis fármacos, de modo a agilizar o processo e diminuir os custos é por meio da química computacional. Diante disto, nesta tese buscou-se por um conjunto de moléculas que possam inibir os vírus SARS-CoV-2 e o vírus da Chikungunya. Assim, o primeiro trabalho consistiu em utilizar um conjunto de moléculas relatadas na literatura, derivados de fenóis e cromonas de *Daldinia sp.*, para a inibição da glicoproteína espícula do SARS-CoV-2, importante para a entrada viral nas células. Baseando-se na afinidade de ligação, realizou-se *docking molecular* obtendo então o Derivado 4 (Der4) com menor energia, removeu-se PAINS e interferentes para a modelagem molecular. Foram obtidos 557 Hits com afinidade entre -7 kcal/mol e -13 kcal/mol, dos quais se selecionou o de maior acessibilidade de síntese (Hit 48, 80%) e o de menor energia de ligação (Hit 250, -13 kcal/mol) e realizou-se novamente o método de *docking molecular* junto com a dinâmica molecular e ADMET. Ambos apresentaram propriedades que indicam boa ação inibitória dos resíduos importantes da glicoproteína espícula assim como viabilidade de uso como droga oral, apesar da dificuldade em sintetizá-los. O segundo estudo avaliou este mesmo grupo de moléculas de *Daldinia sp.* para a inibição do vírus da Chikungunya (CHIKV). Foram escolhidos como alvos as proteínas não estruturais nsP2 e nsP3. Os resultados de *docking* mostraram que todos os derivados têm alta energia de ligação com o nsP2 e baixa com o nsP3, com resultados abaixo de -6,2 kcal/mol, com todos os resíduos importantes com interações moderadas a fortes, com destaque para os Der9 a Der12. Apesar da baixa energia, os resultados de ADMET indicam desvantagens no seu uso como fármaco, tendo alta probabilidade de gerar metabólitos tóxicos, porém destaca-se o Der8 por apresentar boa energia de ligação, bons descritores medicinais e já ser relatado como um inibidor *in vitro* do vírus da zika. Assim, as moléculas selecionadas e modeladas são viáveis para as etapas seguintes do desenvolvimento como fármacos.

Palavras-chave: SARS-CoV-2; CHIKV; *Daldinia sp.*; modelagem molecular; Cromonas.

ABSTRACT

In a new context, such as climate change and the spread of endemic tropical diseases, such as the recent COVID-19 pandemic, misinformation and non-adherence to the scientific method can exacerbate the socio-economic impact of the situation, which justifies the relevance of new ways of streamlining research in this area. The traditional method of drug development is trial and error, which is time-consuming and expensive, and therefore not ideal in this new scenario of global health. One way to narrow down and screen potential drugs, streamline the process and reduce costs is through the computational chemistry. With this in mind, the aim of this thesis was to find a group of molecules capable of inhibiting the SARS-CoV-2 and Chikungunya viruses. Thus, the first work consisted in the use of a group of molecules reported in the literature, derivatives of phenol and chromones from *Daldinia sp*., for the inhibition of the spike glycoprotein of SARS-CoV-2, important for the viral entry into the cells. Based on the binding affinity, a molecular docking simulation was perfomed to obtain the lowest energy derivative (Der4), the PAINS and interferents were filtered out for molecular modelling, yielding 557 hits with binding affinities between -7 kcal/mol and -13 kcal/mol, from which the one with the highest synthetic accessibility (hit 48, 80%) and the one with the lowest energy (hit 250, -13 kcal/mol) were selected, a new docking simulation was perfomed, then a molecular dynamics and ADMET. Both presented properties indicating a good inhibitory action in important residues of the spike glycoprotein as well as the viability to be an oral drug, despite the difficulty to synthesise. After the first work, the second is the evaluation of the same group of molecules from *Daldinia sp.* for the inhibition of the Chikungunya virus (CHIKV), the targets chosen being the non-structural proteins nsP2 and nsP3. The docking results showed that all the derivatives have high binding energy with nsP2 and low with nsP3, with results below -6,2 kcal/mol, with all major residues having strong to moderate interactions, highlighted for Der9 to Der12. Despite the low energy, the ADMET results show some drawbacks of these derivatives as a drugs, with high probability of producing toxic metabolites, nevertheless Der8 is highlighted due to good binding energy, good medicinal descriptors and has already been reported as an inhibitor of Zika virus *in vitro*. Therefore, the selected and modelled molecules are viable for the next steps for drug development.

Keywords: SARS-CoV-2; CHIKV; *Daldinia sp.*; molecular modelling; Chromones.

LIST OF FIGURES

LIST OF TABLES

LIST OF CHARTS

LIST OF ACRONYMS AND ABBREVIATIONS

ADMET	Absorption, Distribution, Metabolism, Excretion and Toxicity
AI	Artificial Intellifence
ANVISA	Agência Nacional de Vigilância Sanitária
BBB	Blood-brain barrier
BCE	Before Common Era
BLA	Biliverdin alpha IX
CDC	Center for disease control and prevention
CE	Common Era
CHIKV	Chikungunya Virus
$CL_{int,u}$	Intrisic clearance rate
CNS	Central Nervous System
COVID-19	Coronavirus Disease 2019
FDA	Food and Drug Administration
GP spike	Glycoprotein Spike
GSK	GlaxoSmithKline™
GT	Golden Triangle
GUI	Graphical User Interface
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HDL	High Density Lipoprotein
H-HT	Human hepatotoxicity
HIA	Human Intestinal Absorption
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
IAV	Influenza A virus
LC_{50}	Median lethal concentration
LD_{50}	Median lethal dose
LogP	Partition coeficient
MCE-18	Medicinal Chemistry Evolution 2018
MERS	Middle East Respiratory Syndrome
MM/GBSA	Molecular mechanics/generalized Born surface area

MO	Molecular orbital
mRNA	Mensenger Ribonucleic acid
MW	Molecular Weight
nAR	Number of aromatic rings
NMR	Nuclear Magnetic Ressonance
nRot	Number of rotatable bonds
PAINS	Pan-assay interference compounds
P_{app}MDCK	Madin-Darby Canine Kidney cells model
Pgp	P-glycoprotein substrate
PPB	Plasma protein binding
QED	Quantitative Estimation of Druglikeness
QSAR	Quantitative structure-activity relationship
R&D	Research and Development
RBD	Receptor-binding domain
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic acid
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SASA	Solvent Accessible Surface Area
SULT	Sulfotransferase
$T_{1/2}$	Half-life
TMPRSS2	Transmembrane protease serine 2
TPSA	Topological Polar Surface Area
UGT	Uridine diphosphate glucuronosyltransferase
VD	Volume of Distribution
WHO	World Health Organization
XRD	X-ray diffraction
ZIKV	Zika virus

SUMMARY

CHAPTER I

1 INTRODUCTION AND OBJECTIVES

1.1 Introduction

Humanity has perennially grappled with pandemic and endemic viral diseases. In the 20th century, notable instances include the Spanish flu (influenza virus, coinciding with the conclusion of the First World War) (Martini *et al.*, 2019), Ebola (caused by the eponymous virus, discovered in the 1970s, with a pandemic occurrence in the last decade) (Li & Chen, 2014), and AIDS (HIV, emerging in the 1980s) (Sharp & Hahn, 2011). In the 21st century, challenges persist with outbreaks such as SARS (coronavirus, 2002) (Cherry & Krogstad, 2004) and COVID-19 (coronavirus, 2019) (Rabi *et al.*, 2020), which have occurred more frequently in recent years, inflicting significant health and social impacts. Among the myriad factors contributing to this increased frequency, it is noteworthy to highlight globalization, particularly facilitated by transportation, urbanization leading to denser populations—sometimes residing in unsanitary conditions with limited access to healthcare—and climate change, which facilitates the spread of vectors to new regions and fosters human-animal interactions. Examples of the species barrier being breached by certain viruses include HIV and coronaviruses. Additionally, immigration of healthcare professionals exacerbates shortages in countries with low to moderate income levels, as these professionals often migrate to countries with higher income levels (Haileamlak, 2022), thereby amplifying the probability of disease transmission.

Viral infections are difficult to treat due to the nature and diversity of viruses, which are not living and use the cell as a means of replication, and the specificity of antiviral drugs. To develop these antivirals, it is necessary to know the replication cycle of the viruses, or at least some of its steps, and the role of the proteins and enzymes involved in each step, then a group of potential drugs are tested to inhibit one of the steps of the cycle, effectively ending the replication cycle. Because each virus has a different preferred cell, different ways to infect cells, different protein targets, and a different DNA or RNA delivery route, the development of antivirals is a difficult task, which is why it is more common to develop antibiotics and viral

vaccines than antivirals (CHENG et al., 2016), with the exception of increasing antibiotic resistance in bacteria and collateral effects, respectively.

Drug research and development (R&D) is divided into three phases: drug discovery, preclinical development and clinical trials. The traditional method, also the oldest, of drug discovery is trial and error, it is, as the name suggests, to discover new drugs by chance, with historical examples such as penicillin as the first antibiotic, nitroglycerin (an explosive molecule) used as a vasodilator and to treat angina, and chlordiazepoxide as a result of work on a class of dye molecules, obtaining the first benzodiazepine (DOYTCHINOVA, 2022). It is clear that this method was overcome, and had to be, since the first molecules were the result of accidents. In the case of antivirals, it would be a more difficult task because of the characteristics mentioned above.

New molecular modelling methods, particularly computational methods, are now widely used to identify molecular targets involved in one or more pathogenic processes and to test and evaluate the effect of a candidate (potential drug) on that target. The target can vary depending on the disease, but the most common is a viral protein or a cellular receptor. The initial ligand to be used can be very diverse, but certain criteria are usually used to define it, usually the similarity of diseases (to be able to treat both), pathogens (to be able to inhibit both), targets and so on. In the event of failure, an assessment of the binding affinity, structural interactions, medical prediction and stability of the target-ligand complex forms the basis of an optimisation process that can be carried out to generate new molecules from the initial one or to test new classes of molecules.

Among drugs, the most common and successful source is natural products, due to the history of drug development, which only in recent years (last century) has reached the achievement of new synthetic drugs, thanks to new methodologies and drug modifications. With a diverse biodiversity and undiscovered biomolecules, this is a promising area for the search, research and discovery of new drugs or a molecular framework (DIAS; URBAN; ROESSNER, 2012), some of which are widely used today in the form of extracts, such as Papaver somniferum (marketed in Brazil as "Elixir Paregórico"), Valeriana officinalis, Aloe vera, Cannabis sativa, Ginkgo biloba and Salvia officinalis, to name a few. The advantages over synthetic drugs are that they have fewer side effects (less frequent and less severe) and are generally cheaper to obtain than synthetic drugs (SINGH; BHARDWAJ; PUROHIT, 2022). With these characteristics, there is a high probability of finding an inhibitor molecule or its derivatives that are safe for use in the treatment of viral diseases.

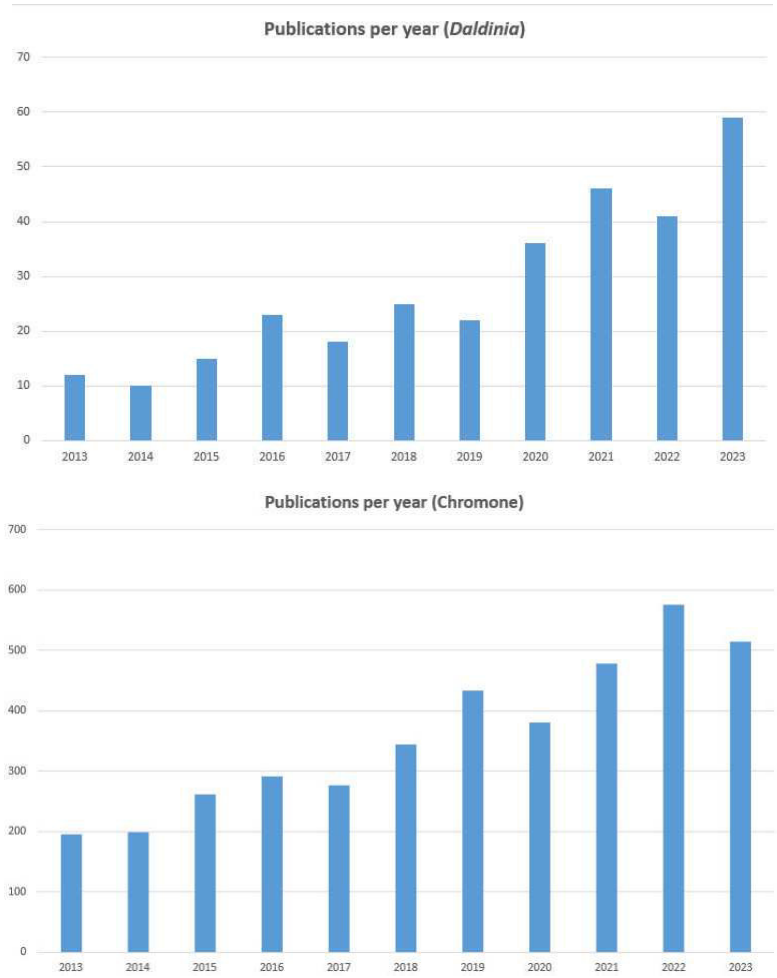
To develop new medicines from natural products, it is necessary to explore molecules from sources that have been little studied. Among several sources, the fungus of the genus Daldinia emerges as a potential source as it is little studied, widely distributed worldwide (in South America, Central America, North America, East and Southeast Asia, Oceania, Sub-Saharan Africa and Europe) and recent studies have identified the presence of immunosuppressive polyketones in D. eschscholzii and the antioxidant isoindolinone in D. concentrica. Some of these biomolecules have shown antibacterial and anti-H1N1 activity (LI *et al*., 2021). Therefore, natural products derived from Daldinia species represent a promising line of research for the development of antiviral agents against a variety of viruses.

Among the scarce literature (Figure 1) on the subject, the work of Zhang and collaborators stands out for the novel molecules, chromone and phenols, which are well characterised and which show in vitro activities against influenza A virus (IAV) and Zika virus (ZIKV, which causes the eponymous disease), which is important to note that it is spread concomitantly with dengue and Chikungunya (ZHANG et al., 2021a).

This study explores the potential of chromones and phenolic derivatives as drug candidates to inhibit both SARS-CoV-2 and CHIKV. Although the SARS-CoV-2 pandemic may have subsided, the occasional emergence of variants and sporadic cases highlights the continuing threat. Moreover, the inevitability of future SARS pandemics, given historical precedent, underscores the importance of existing drugs for treatment or as lead molecules for drug development. In addition, diseases transmitted by the same vector, such as dengue, Zika and chikungunya, remain endemic in tropical regions and are neglected diseases that primarily affect economically disadvantaged regions. However, with the increasing impact of climate change, these diseases are attracting increased global attention as they spread to new geographical areas. Figure 1 illustrates the trend in publications related to daldinia and chromone, indicating a growing interest in these topics.

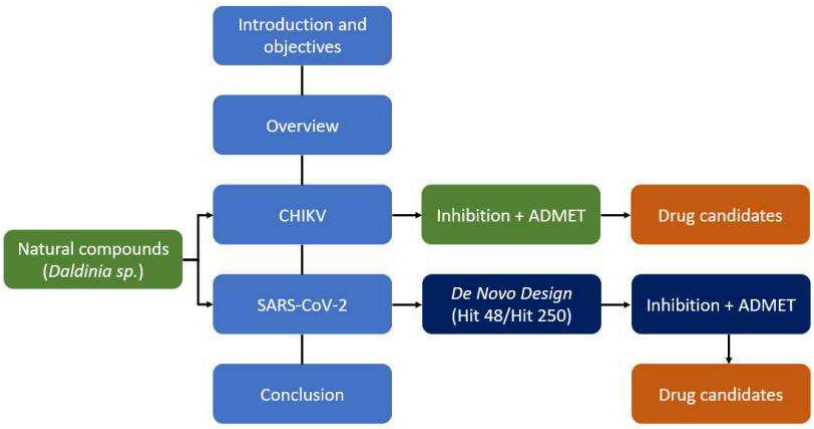
Chapter II provides an overview of the diseases, computational tools, Daldinia derivatives and the state of the art. Chapter III is the paper "Phenol and chromone compounds for in silico inhibition of nsP2 and nsP3 of Chikungunya virus" submitted to the DARU Journal of Pharmaceutical Sciences. Chapter IV presents the results of the paper "De novo Design of bioactive phenol and chromone derivatives for Inhibitors of Spike glycoprotein of SARS-CoV-2 in silico", published in 3 Biotech. Chapter V presents a general conclusion of the work and the perspective of further work to develop the drug at the end of the research and development line. The chapters and their contents are shown in Figure 2.

Figure 1 – Scientific production involving "*Daldinia*" (top) and "Chromone" (bottom), on ScienceDirect website on December 15, 2023.



Source: Author.

Figure 2 – Thesis structure.



Source: Author.

1.2 Objectives

1.2 General

To evaluate the potential antiviral activity of chromone and phenolic derivatives (ligands) and to discover new compounds for the treatment of Chikungunya fever and COVID-19 *in silico*.

1.2.2 Specifics

a) To perform the molecular docking of the ligands with non-structural protein targets of CHIKV;
b) To evaluate the ADMET of the ligands;
c) To perform molecular docking of the ligands with GP spike;
d) To propose novel synthetic molecules from the best ligand obtained from molecular docking;
e) To evaluate the synthetic accessibility and oral safety of the proposed ligands.

CHAPTER II

2 OVERVIEW

2.1 Epidemic and Pandemic diseases

Reports of infectious diseases in human history date back to ancient times, with the detection of smallpox or variola (caused by the variola virus) through genome sequencing in mummies (from the 17th century onwards) and lesions resembling smallpox rashes (12th century BCE) (MCCOLLUM et al., 2014; WERTHEIM, 2017). Additionally, there are numerous accounts of epidemics, such as the plague in Athens (5th century BCE), the Antonine Plague (Roman Empire, 165 CE), the Plague of Justinian (Byzantine Empire and parts of Europe, 541 CE), and the Black Death (also known as the bubonic plague, which affected Europe, parts of West Asia, and North Africa in 1346 CE) (SAKAI; MORIMOTO, 2022), some of which resulted in death tolls estimated at up to 60% in Western Eurasia over eight years, in the case of the Black Death (SPYROU et al., 2022).

In the context of the Black Death, as depicted in Figure 3, efforts to safeguard public health from a highly contagious disease led to the implementation of mandatory measures aimed at physically separating individuals, animals, and objects that had been in contact with or exposed to infection. This practice, now known as quarantine (from the Italian "quaranta," referring to the forty days of isolation), was first enforced in 1377 in Croatia and later in Venice in 1423, when the establishment of dedicated plague hospitals, such as the Lazaretto (named after the biblical figure Lazarus), further underscored the stigma associated with infection during public health crises. Throughout history and up to the present day, it is evident that addressing public health challenges requires collective action, whether through the enforcement of quarantine, the development and administration of vaccines (including mandatory vaccination0 (GRAVAGNA et al., 2020), the advancement of pharmaceutical treatments, or vector suppression strategies (CARVALHO et al., 2015), among various other collaborative approaches. It is noteworthy that the stigmatization of the infected, the disruptions to daily life, the diminished social interactions (despite the advent of social media and remote technologies), and the scapegoating of particular ethnic or national groups (e.g., during the Black Death, the Jewish population (WINKLER, 2007), and amid the COVID-19 pandemic, individuals of

Chinese descent (HUANG et al., 2023)) persist in infectious disease outbreaks, as evidenced by the recent COVID-19 pandemic (LIMA, 2020).

Figure 3 – Burying the black death victims, Tournai, Belgium, circa 1353.

Source: *Tractatus quartus*. Image in public domain.
Note: Despite the sorrow, the illustrated scene and context may recur in the XXI century.

To further discuss diseases and their occurrence in modern times, it is necessary to clearly define the terms 'endemic', 'epidemic', and 'pandemic'. These terms are defined by the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC, 2012) and are described in Chart 1.

Chart 1 – Terms for the occurrence and distribution of a disease, accordingly to CDC.

Occurrence	**concept**	**details**	
Endemic	Common presence of the disease in a community	Sporadic	Irregular occurrence
		Hyperendemic	Persistent
Epidemic	Sudden growth of cases in a population of a determined area	outbreak	Delimited geographic region
		Cluster	Cluster of cases that may be higher than reported in a space and time
Pandemic	Epidemic that has spread globally	Affects a high amount of people in the world	

Source: CDC (2012).

Advances in sanitation techniques, medicine, and diagnostics, as well as the discovery, research, and development of antibiotics, from penicillin to azithromycin, have significantly contributed to mitigating bacterial diseases. Although the threat of antibiotic-resistant bacteria is growing, viruses have emerged as the biggest threat among the myriad of

pathogens that cause infectious diseases. Recent epidemics and pandemics have been predominantly viral in nature, including Spanish flu (Influenza A) (MARTINI et al., 2019), polio, smallpox, rubella, AIDS, Ebola, SARS, MERS and SARS-CoV-2. Some, such as smallpox, have been eradicated through global vaccination campaigns. When reflecting on virus outbreaks and their contagious nature, it is important to note that in recent decades the term 'viral' has acquired a broader connotation. In addition to its biological meaning, the term 'viral' can now refer to anything that is highly contagious and spreads quickly. This semantic evolution is exemplified in the common use of 'viral' as an adjective, such as in the phrase 'the article went viral', indicating a sudden increase in citations over a specific period of time.

Chart 2 presents some infectious viruses. It should be noted that while some have drugs, others such as herpes and AIDS do not have a cure, but their symptoms can be treated. However, an increase in resistant strains of herpes has made treatment ineffective. Nocchi et al. (CHENG et al., 2016; NOCCHI et al.) have discussed this scenario. In 2022, a topic formulation was developed using an extract from the bark of Schinus terebinthifolia, commonly known as 'Aroeira'. This extract is popular in folk medicine for its anti-inflammatory, antimicrobial, analgesic, and antiulcerogenic properties, and is used to treat urinary tract and mucosal injuries. The formulation was tested in vivo on mice and was found to be an effective antiviral, similar to acyclovir but less irritating.

Chart 2 – Some of the common human viral infections.

Family	Viruses	Transmission	Symptoms	Diagnostic	Treatment	Vaccine
Herpesviridae	HSV-1, HSV-2	Contact	Oral/genital herpes	*	Acyclovir, valacyclovir	–
Orthomyxoviridae	Influenza A	Contact Droplets	Pneumonia	Immunofluoresce	Neuraminidase inhibitor, Favipiravir	Trivalent or Quadrivalent influenza vaccine
	Influenza B		Encephalitis	*		
	Influenza C		Respiratory infection			
Filoviridae	Ebola virus	Contact	Hemorrhagic fever	*	–	Ebola vaccine (trial)
	Marburgvirus					–
Picornaviridae	Hepatovirus	Droplets	Hepatitis A	*	–	Hepatitis A vaccine
	Poliovirus		Poliomelitis			Polio oral vaccine
	Enterovirus		Diverse**			–
Rhabdoviridae	Rabies virus	Contact	Rabies	*	Immunoglobulin	Rabies vaccine
Flaviviridae	Dengue virus	Arthropod	Hemorrhagic fever	*	–	Dengue vaccine (trial)
	Hepatitis C	Blood contact	Acute and chronic hepatitis	*	interferon	–
	Zika virus	Arthropod	Zika fever	*	–	–
Paramyxoviridae	Measles virus	Droplets	Measles	*	–	MMR vaccine
Togaviridae	Rubella virus	Droplets	Rubella	*	–	MMR vaccine
	Chikunguny a virus	Arthropod	Chikunguny a			–
Retroviridae	HIV-1, HIV-2	Body fluid contact	AIDS	*	HAART	–
Hepadnaviridae	Hepatitis B virus	Body fluid Contact	Acute and chronic hepatitis B	*	Nucleoside, Immunoglobulin	Hepatitis B vaccine
Coronaviridae	SARS-CoV SARS-CoV-2	Droplets	Acute Pneumonia	***	–	Vaccines**
Papillomaviridae	Human papillomavirus	Contact	Warts	*	Imiquimod	HPV vaccine

Source: adapted from CHENG *et al.* (2016).
Note: The diagnostics noted with * indicates that the diagnostic methods are PCR, viral culture and/or serology; The ** notation includes: respiratory illness, meningitis, myocarditis, among others. The *** notation indicates a large group of diagnostics (for example RT-PCR, RT-LAMP, GICA and ELISA) and vaccines.

2.2.1 Coronavirus Disease-19

To understand the COVID-19 pandemic, it is important to note that it was a sudden global outbreak of cases. This crisis has posed significant challenges across all sectors and is considered one of the biggest health crises of the century (BRITO et al., 2020). Despite the WHO declaring an end to the COVID-19 public health emergency (WISE, 2023), its impacts are still being felt today and have brought about significant changes in our societies.

The COVID-19 pandemic has provided on the scientific community and healthcare workers with new perspectives on the search for vaccines and drugs. The first, vaccine was developed in record time of less than a year after the WHO declared the pandemic. mRNA vaccines were rapidly approved and manufactured, saving millions of lives (GOTE *et al.*, 2023). The latter treatment repurposed antiviral drugs, such as Nirmatrelvir with Ritonavir, Remdesivir, Molnupiravir, which were approved by the Food and Drug Administration (FDA) (CDC, 2023). However, SARS-CoV-2 is a RNA virus. As a less stable ribonucleic acid than DNA with only a single short strand, it is more susceptible to mutations, being able to evolve in a matter of months or years (MARKOV *et al.*, 2023). Additionally, its highl contagiousness, relatively low lethality, latent period, public health policies and social factors (BRITO *et al.*, 2020; XIN *et al.*, 2022) have resulted in a proliferation of variant strains, leading to a asynchronous epidemic waves.

Although vaccines and drugs can reduce the lethality, recovery time, and symptoms of COVID-19, their effectiveness is diminished by emerging variants (WU *et al.*, 2023). Therefore, it is expected that humanity will not eradicate SARS-CoV-2, but rather coexist with it in cases of epidemics and/or hyperendemics (CONTRERAS; IFTEKHAR; PRIESEMANN, 2023). In this scenario, it is likely that another outbreak of SARS-CoV will occur in the future, as history has demonstrated with previous outbreaks such as SARS (2002) (CHERRY, 2004; GOLDSMITH *et al.*, 2004) and MERS (RABI *et al.*, 2020). This outbreak could be a derivative of SARS-CoV-2 or have a sylvatic origin.

Numerous lessons have been gleaned from the COVID-19 pandemic. Among these lessons, pertaining to the rapid development of drugs and vaccines, is the imperative of research focus and diversity (VON DELFT et al., 2023). Additionally, there has been a reevaluation of the role played by professionals, research institutions, and scientific methodology. The urgency driven by governmental pressures and societal demands led to the publication of studies with rapid but inadequate peer review and diminished rigor. Regrettably, this hastiness resulted in

instances of misinformation disseminated via social media platforms, with repercussions for public health. For instance, the controversial work of Raoult and colleagues (GAUTRET et al., 2020) regarding the efficacy of (hydroxy)chloroquine contributed, among other factors, to elevated death tolls in countries such as Brazil (NOGUEIRA; BRITO, 2020) and the USA, with 702,116 and 1,144,877 total confirmed deaths respectively, by 2023 (WHO, 2023a). It is noteworthy that Raoult's work garnered considerable attention, with 1565 citations in 2020, 838 in 2021, 257 in 2022, and 100 in 2023 (according to PubMed, accessed on December 16, 2023), serving as both a subject of critique (KERRIDGE; SILVA; UPSHUR, 2023) and a reference for ongoing investigations (SOBNGWI et al., 2023). Furthermore, an intriguing aspect pertains to the business model adopted by journals and publishers. Critiques of certain studies often require paid or institutional access, while research that aligns with positive clinical outcomes, such as those related to (hydroxy)chloroquine, tends to be freely accessible. This discrepancy in access to knowledge has the potential to introduce bias into research, particularly in matters pertaining to public health. The multifaceted impacts of the COVID-19 pandemic on research and society continue to be extensively studied and documented, underscoring its significance in shaping humanity's response to health crises.

Before the pandemic 67% of approved antivirals targeted to HIV and HCV. Even with rapid phase I drug through computational tools and methods, phases II (preclinical) and III (clinical) remain lengthy and costly. As COVID-19 cases began to surge globally, researchers from diverse fields worked to solve the emerging problems. Some solutions, such as oral repurposed drugs, like Merck's molnupiravir and the vaccines, were developed in record time. In the field of COVID-19 design, there are two possible classes of targets: the virus itself (such as Mpro and GP spike) or the infected host (ACE2 and TMPRSS2). Significant effort has been made to discover host target ligands to prevent SARS-CoV-2 cell binding, with the reasoning that a host directed drug would have a broader spectrum against other pathogens. Cleary, binding to a cell target may prevent its proper function, leading to toxicity and worsening of the clinical condition. Therefore, the safest method is to target viral proteins. To achieve this, genome sequencing, viral cycle elucidation and protein analysis (XRD) should be conducted to propose a potential ligand (VON DELFT *et al.*, 2023)

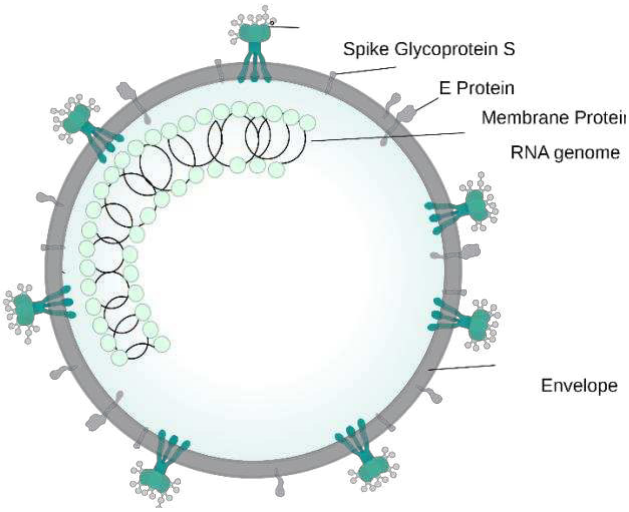
The SARS-CoV-2 is an Coronaviridae, ranging from 60 to 140 nm (BAR-ON *et al.*, 2020), with a single stranded positive-sense RNA (+ssRNA). The RNA is involved in the nucleocapsid protein (N), that in sequence is involved by the envelope, which has 3 components: membrane protein (Mpro), spike protein (Spro, also known as glycoprotein spike,

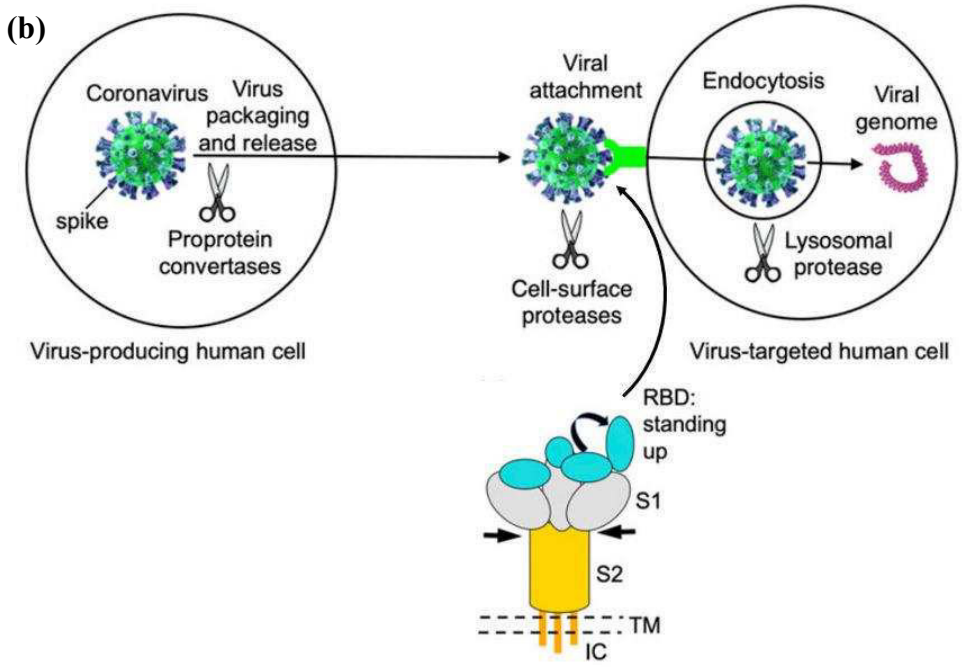
GP spike, what give coronavirus its name, as the spikes resembles a crown), and envelope protein (E) (WANG *et al.*, 2020), its structure is represented in Figure 4a.

The spike protein is composed of fused trimeric membrane stalk (S2) that have trimeric protein heads that act as a receptor-binding (S1) on top of S2. The receptor-binding domain (RBD), in S1, actuate in the angiotensin-converting enzyme 2 (ACE2, the receptor) specifically, the S1 can switch between an up or down position, for binding with the receptor or avoiding immune response, respectively. After a successful binding, to entry in the cell a protease, transmembrane protease serine 2 (TMPRSS2, a host cell surface protease), act on the boundary between S1 and S2, the later is induced a structural change to entry in the cell. The mechanism is illustrated in Figure 4b (SHANG *et al.*, 2020).

Figure 4 – SARS-CoV-2 Structure and schematic of the cellular infection mechanism

(a)





Source: user *SPRQ10* (a) and adapted from Shang *et al.* (SHANG *et al.*, 2020) (b). (a) obtained from <https://commons.wikimedia.org/wiki/File:Coronavirus_virion_structure.svg>. Accessed in December 16, 2023. Image licensed under *Creative Commons Attribution* 4.0 *International* (CC-BY-4.0), < https://creativecommons.org/licenses/by/4.0/deed.pt>, All the rights reserved to the original authors.

In front of this, surface proteins of SARS-CoV-2 are preferable as targets for drug discovery, with emphasis on the GP spike as it is the key to the entry in the host cells and common to coronaviruses. Concerning the infection, one of the defense mechanisms characteristics to the COVID-19 infection is a high release of pro-inflammatory cytokines, as a result of which there is a cascade of immunological defense and inflammation of the lung epithelial tissue (infected or not).

In this scenario, an individual infected with a weakened and occupied immune system is more susceptible to concurrent diseases. During quarantine and lockdowns, the population stayed at home, close to mosquitoes breeding sites, especially during rainy seasons. Some of these diseases are typical in tropical regions, like Dengue, Zika and Chikungunya, which, due to the effort in the COVID-19 pandemic, reports indicate that the cases of these diseases are underreported during this period (BORRE *et al.*, 2022; SILVA; MAGALHÃES; PENA, 2021; VICENTE *et al.*, 2021).

2.2.2 Chikungunya

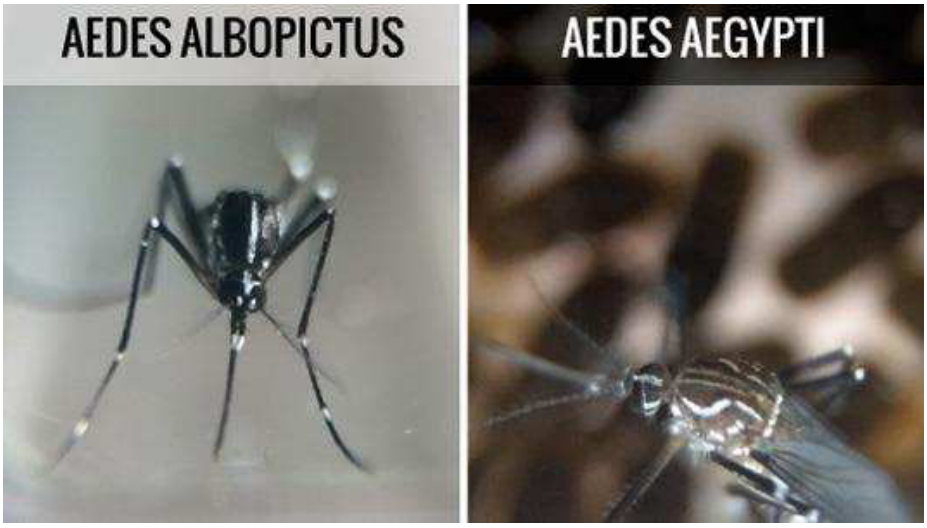
Arboviruses are infectious viruses, such as Dengue virus (DENV), Zika virus (ZIKV), and Chikungunya virus(CHIKV), which are transmitted by the anthropophilic

mosquitoes of the *Aedes* genus. The most common vectors are *Aedes aegypti* and *Aedes albopictus* (Figure 5a), and they are found in tropical regions (Figure 6). These diseases may have similar symptoms in the early stages, such as fever. Therefore, it can be difficult to diagnose them, especially during the COVID-19 pandemic. The lack of correct treatment may lead to a worsening of the clinical condition and sequelae (VICENTE *et al.*, 2021).

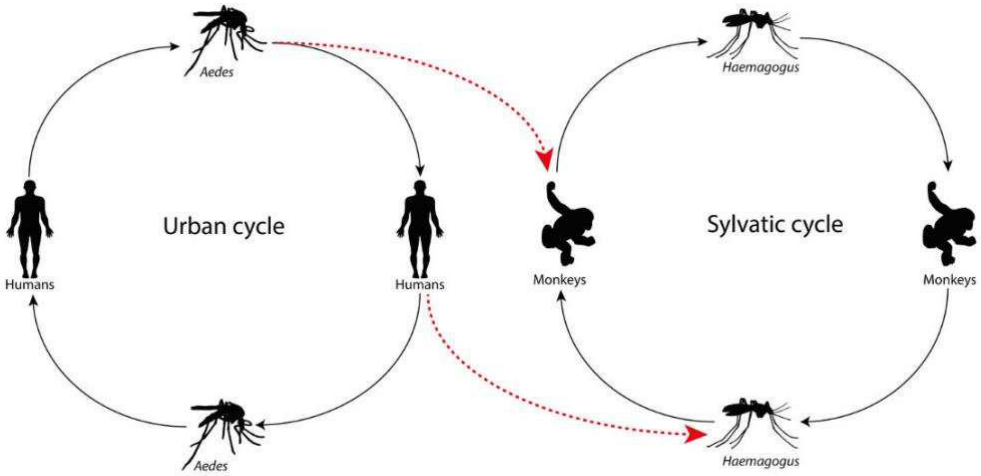
To control the disease, the vector must be removed from the cycle, so, the vector and its cycle must be studied in details to assess the transmission. The urban expansion and the sanitary conditions have allowed the proliferation of *Aedes*, of which humans are the main reservoir of the disease, and when the conditions for proliferation in urban areas cease , during the inter-epidemic period, animals are the main reservoir, especially monkeys. During this period, another genus of mosquitoes, *Haemagogus* (in particular *H. janthinomys* and *H. leucocelaenus*) (DE ABREU *et al.*, 2019), do the opposite of *Aedes*: the main reservoirs are animals and occasionally infect a human (Figure 5b). For this reason, the epidemics are persistent and to solve this public health challenge, the vectors should be removed from the cycle.

Figure 5 – *Aedes aegypti* and *Aedes Albopictus* mosquitoes (a), the main vectors of dengue, Zika and Chikungunya. For an easier species differentiation, note the different white strips pattern. The urban and sylvatic cycle or arboviruses (b).

(a)



(b)

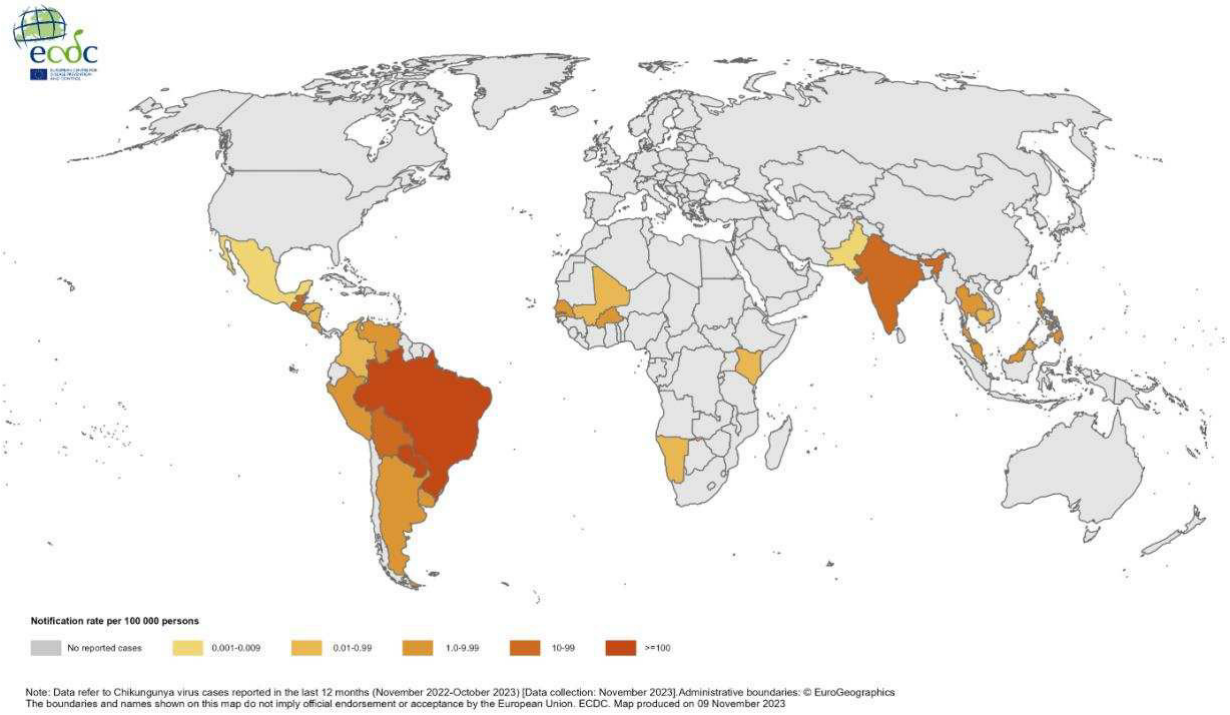


Source: (a) Institute Oswaldo Cruz/Fiocruz. Photo authorship: Josué Damacena, 2016. (b) Figueiredo (FIGUEIREDO, 2019). Image licensed under *Creative Commons Attribution* 4.0 *International* (CC-BY-4.0), < https://creativecommons.org/licenses/by/4.0/deed.pt>, All the rights reserved to the original authors.

Strategies of vector control, in the case of Brazil, are traditionally health awareness campaigns, which consist of mechanical control (prevention of possible spawning sites and prevention of contact with the vector), biological control (natural predators, such as fish, most common in Brazil the *Betta splendens* (CAVALCANTI *et al.*, 2007)), chemical control (use of chemical products to kill larvae and adult mosquitoes). Health agents for the mechanical and chemical control of the vector, to detect and prevent any possible water reservoir for the *Aedes* mosquito spawning. These methods are effective in reduzing the incidence, but are far from being the ideal solution to prevent the return of the outbreaks in the period of urban proliferation period, and the research is scarce for being a tropical disease, that affects low-income countries, a solution to their problems and for being of little interest to the majority of research powerhouses countries, thus being nominated as Neglected Tropical Diseases (NTD) by the WHO, and among the 20 different diseases, Chikungunya has no approved drug or vaccine (Chart 2) (MOLYNEUX *et al.*, 2021; WHO, 2023b).

The distribution of Chikungunya is shown in Figure 3, which shows that it is mainly found in tropical regions. Due of globalization, deforestation, climate change, population disorder, increasing rates of population density and sanitization, the vector can and does migrate to other regions, making it a global concern (BARTHOLOMEEUSEN *et al.*, 2023; BATTISTI; URBAN; LANGER, 2021; MOLYNEUX *et al.*, 2021).

Figure 6 – Notification rate of cases of Chikungunya virus per 100.000 habitants, from November 2022 to October 2023, produced by the European Center for Disease prevention and control (ECDC).



Source: ECDC (2023).

Most antivirals for CHIKV target its surface proteins, such as envelope (E1, E2), capsid protease (C), non-structural proteins 1, 2, 3 and 4. Of these, highlights the nsP2 stands out as a well-established target and the nsP3 as an understudied target, both of which are important for its replication cycle. The first is a common target because of the early XRD obtained, mostly molecules *in silico* that were capable of effective inhibition share some characteristics, as long chains, high number of aromatic rings and π staking sites (BATTISTI; URBAN; LANGER, 2021), the latter, a relatively larger protein with 4 different chains (A, B, C and D), has few published works, according to Science Direct ("nsP3 chikungunya"), from 2013 to 2023, no more than 50 papers per year were published with the exception of 2021 (51) and 2023 (52), in the work of Zhang *et al.* (ZHANG *et al.*, 2021c), small molecules (similar to chromones, but nitrogenated) had better fit in the receptor binding pocket than the larger molecules of nsP2, showing that it is possible to use small molecules, which are safer to use, may be used to treat chikungunya fever.

Despite the difficulty, history proves that a global effort to eradicate diseases is possible, as evidenced when in 1980 WHO declared that smallpox was eradicated. To do this, detection, monitoring of cases and vaccination has to be made. Smallpox had characteristics that made it easier to eradicate, for instance, only affect humans (in contrast to Chikungunya), easily diagnosticated, vaccination of the contacts, fast production of vaccine a single dose vaccine induces long lasting resistance. COVID-19 vaccines differ from the 2 later characteristics (MEYER; EHMANN; SMITH, 2020).

Each infectious disease presents unique challenges that researchers must address in a specific way. In drug discovery for these diseases, computational chemistry is emerging as a fast, economical and powerful method in the search for new drugs, given the diverse possibilities of finding molecules, but the low probability of finding a viable molecule that is capable of inhibiting the pathogen and is safe to use.

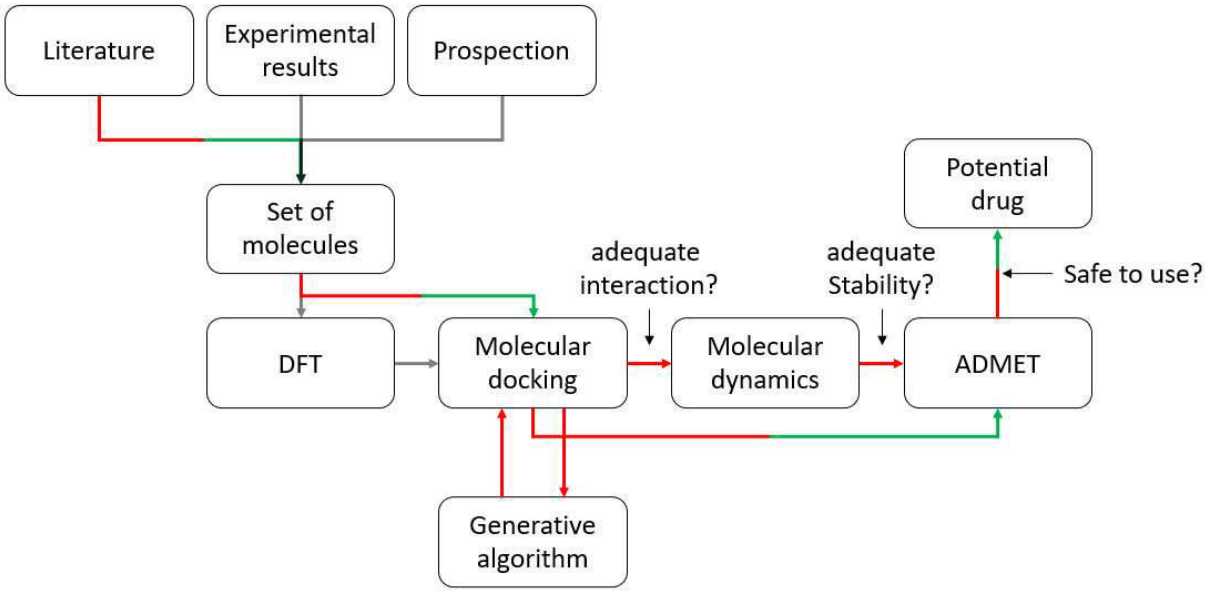
2.2 Computational chemistry

Computational chemistry can be defined as the use of computers and its tools to solve chemical problems. It envolves the use of theoretical equations and experimental data to model and optimeze chemical processes. This work focueses on the use of computational chemistry tools for drug development. Specifically, it involves testing a ligand (a molecule intended to be a drug) against a target, analyzing the binding and residues involved, and assessing the stability of the ligand-target complex. The ultimate goal é to ensure the safety of the drug before proceeding to *in vitro* and *in vivo* assays, regulatory approval and commercialization.

Generally, the pharmaceutical industry wastes a significant portion of its clinical trial budget due to the high failure rate of around 90%. The primary cause of this failure is attributed to the drug discovery stage, which includes poor properties and unsuitability for the recipient (Sadybekov; Katritch, 2023). To summarize the workflow of drug discovery, Figure 7 is presented. However, some studies may extend beyond this, incorporating in vitro assays (INDU et al., 2022) to validate results and confirm molecular docking and molecular dynamics simulations, as well as retrosynthesis to propose a synthetic route or synthesize and characterize compounds (BORN et al., 2021a). From now on in this work, retrosynthesis will refer to the reverse route of organic synthesis of a determined compound from simpler and available reagents, that may be ranked by number of steps to produce the compound of interest, price and delivery time of precursors. Regardless of some recent work on nanoparticles as ligands is

promising,on computational nanoparticles as ligands (MEHRANFAR; IZADYAR, 2020), only molecules are considered for drug discovery due to computational limitations.

Figure 7 – General workflow of computational drug discovery. In green the steps made in Chapter III, in red, the steps made in Chapter IV.



Source: Author

From Figure 7, the candidates may come from the literature, as the result of experiments (e.g. organic synthesis of novel compounds) or the prospection of products (usually extraction from plants, fungus and marine life) (LI *et al.*, 2021) that present some potential against a virus target. This set of molecules may be optimized to obtain accurate conformers (ALENCAR *et al.*, 2022) and screen reactivity (e.g. analysis of HOMO-LUMO) (HUSSAIN; AMIR; RASOOL, 2020). In sequence, submitted to a molecular docking simulation (docking), were the molecules interacts with the target and the binding affinity is measured and the target residues may be analyzed. After that, three paths are possible: 1) use of a generative approach to produce a new set of molecules or extend the existing one (e.g. virtual screening, *De Novo Design*) (BAI *et al.*, 2022) to be sent to docking again; 2) a molecular dynamics simulation to assess the stability and behavior, in physiological medium, of the ligand-receptor complex in time; 3) after a docking simulation to be sent to an Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) evaluation, to estimate the physicochemical properties, druglikeness, clearance rate, half-life, oral bioavailabilty, the probabilities of metabolism (and

its sites), hepatotoxicity, mutagenicity and more indicators and descriptors depending on the objective of the drug and the development of ADMET tools.

Some of the widely used webservers and softwares are presented in Chart 3, with the official websites presented in Appendix A. In this work context, AutoDock Vina, MolAICal, ADMET server, MarvinSketch and UCSF Chimera are highlighted. AutoDock Vina is to date one the most utilized docking softwares (BAI *et al.*, 2021; EBERHARDT *et al.*, 2021; SHITYAKOV; FÖRSTER, 2014; TORNG; ALTMAN, 2019; TROTT; OLSON, 2010), thus it is a standard in literature. UCSF Chimera is used to prepare and visualize proteins, removing interferents, targets non-related residues, water, ligands (from the PDB file) (PETTERSEN *et al.*, 2004), MarvinSketch is a chemical drawing software with many functionalities, like prediction of NMR spectra, ADME descriptors, generating SMILES string, .mol files, conformers, isomer analysis to name a few (CHEMAXON, 2022).

MolAICal is a powerful and versatile tool for drug discovery, it uses some other programs and algorithms as base (e.g. AutoDock Vina) with slight modifications to improve its accuracy, it has a friendly Graphical User Interface (GUI) in comparison with other software of the genre, it is easy to install, setup and use, with well described manual and forum support, and in a single package brings molecular docking, *De Novo Design*, virtual screening, molecular dynamics , quantitative structure-activity relationship (QSAR), synthetic accessibility (SA) and filters (Lipinski (LIPINSKI *et al.*, 2001), pan-assay interference compounds (PAINS) (BAELL; WALTERS, 2014)). Despite being a new software, MolAICal community and publications are increasing (BAI *et al.*, 2021).

Chart 3 – Softwares and webservers for *in silico* drug development. An * indicates some limitation in the function.

Type	Name	License	Functions and observations	References
Server	PaccMann[rl]	Free	Retrosynthesis, *De Novo Design* (need python ≥3.7)	(BORN *et al.*, 2020, 2021b)
	AiZynthFinder	Free	Retrosynthesis (need python 3.9)	(GENHEDEN *et al.*, 2020)
	Manifold	Free trial (15 days)	Retrosynthesis, medicinal chemistry*	(FLORESTA *et al.*, 2022)
	ADMETlab 2.0	Free	ADMET	(XIONG *et al.*, 2021)
	SwissADME	Free	ADMET	(DAINA; MICHIELIN; ZOETE, 2017)
	ProTox-II	Free	ADMET (toxicity)	(BANERJEE *et al.*, 2018)
	XenoSite	Free	ADMET (metabolism and toxicity)	(MATLOCK; HUGHES; SWAMIDASS, 2015)
	pkCSM	Free	ADMET	(PIRES; KAMINSKAS; ASCHER, 2018)
	ADMETboost	Free	ADMET	(TIAN; KETKAR; TAO, 2022)
	PreADMET	Free	ADMET	(KOVAČEVIĆ *et al.*, 2014)
	StopTox	Free	ADMET (metabolism and toxicity)	(BORBA *et al.*, 2022)
	C6H6 NMR	Free	IR*, NMR spectra prediction	(PATINY *et al.*, 2018)
	DockThor	Free	Docking	(SANTOS *et al.*, 2020)
Software	MolAICal	Free to noncommercial use	*De Novo Design*, virtual screening, molecular docking, MM/GBSA	(BAI *et al.*, 2021)
	UCSF Chimera	Free to noncommercial use	Visualization, structure analysis, sequence alignments and density maps	(PETTERSEN *et al.*, 2004)
	MarvinSketch	Free to noncommercial use	Structure drawing, NMR prediction, generate SMILES string and .mol files	(CHEMAXON, 2022)
	ORCA	Free for academic uses	DFT	(NEESE *et al.*, 2020)
	PyMOL	Open-source	Molecular visualization	(DELANO, 2002)
	AutoDock Vina	Open-source	Molecular docking	(TROTT; OLSON, 2010)
	GROMACS	Open-source	Molecular dynamics	(ABRAHAM *et al.*, 2015)
	CP2K	Free	DFT and molecular dynamics	(KÜHNE *et al.*, 2020)
	Gabedit	Package (ORCA)	Visualizaiton, MO, UV-Vis, IR, Raman spectra	(ALLOUCHE, 2012)

Source: Author.

An important observation must be made for ADMET, which is based on datasets of already known molecules or fragments, commercial or not, and every server uses different groups of datasets and algorithms, that may outcome some disparity in the results. In the case of Chikungunya diagnostics, to treat the symptoms in Brazil, the most common drugs are presented in Table 1.

Table 1 – Physicochemical descriptors of common medicine for the treatment of chikungunya, obtained from ADMETlab 2.0. The * notation indicates low to moderate probability.

Property	Dipyrone	Prednisone	Acetaminophen (paracetamol)
Physicochemical Properties			
logP	-0.541	1.694	0.608
logD	-0.933	1.043	0.789
MW	333.08 g/mol	358.18 g/mol	151.06 g/mol
HBA	7	5	3
HBD	0	2	2
TPSA	83.370 Å²	91.67 Å²	49.33 Å²
nRot	4	2	2
nRing	2	4	1
MaxRing	6	17	6
nHet	9	5	3
fChar	0	0	0
nRig	14	23	7
nStereo	0	6	0
Medicinal Chemistry			
Pfizer rule*	(–)	(–)	(–)
GSK filter	(–)	(–)	(–)
Golden Triangle	(–)	(–)	1 alert; MW<200 g/mol
QED	0.453	0.785	0.595
Fsp^3	0.308	0.667	0.125
MCE-18	14.0	73.6	6.0
Highlighted descriptors			
AMES	(+)	(–)	(–)*
H-HT	(+)	(–)*	(–)
BBB	(+++)	(+++)	(++)
VD	1.573 L/Kg	0.563 L/Kg	0.923 L/Kg

Source: Author.

Although dipyrone and paracetamol are already approved by Agência Nacional de Vigilância Sanitária (ANVISA), and are well-known and studied commercial drugs, they present low values of quantitative estimated druglikeness (QED), that is the similarity of some physico-chemical properties of the compound with already commercialized and well-studied drugs, with criteria of QED > 0,67 as excellent and ≤ 0,67 a poor result, that would mean that dipyrone and paracetamol are similar to commercially available drugs, even that they are already in this set. Also, the evaluation presented that they have moderate probability of being mutagenic (AMES test, allows for positive or negative results or probability to cause DNA damage), which is not the case of dipyrone *in vitro* studies (ARAGÃO; TOBIAS, 2019), and all three drugs presented a high probability of blood-brain barrier permeability, that is they are able to reach the central nervous system, usually for viral diseases it is an undesirable property as it may accumulate and/or cause nervous system damage, but in an analgesic it would be expected and desirable, also with excellent volume distribution (VD, concentration of the drug on the circulation). Therefore, even with good accuracy overall, according to the druglikeness

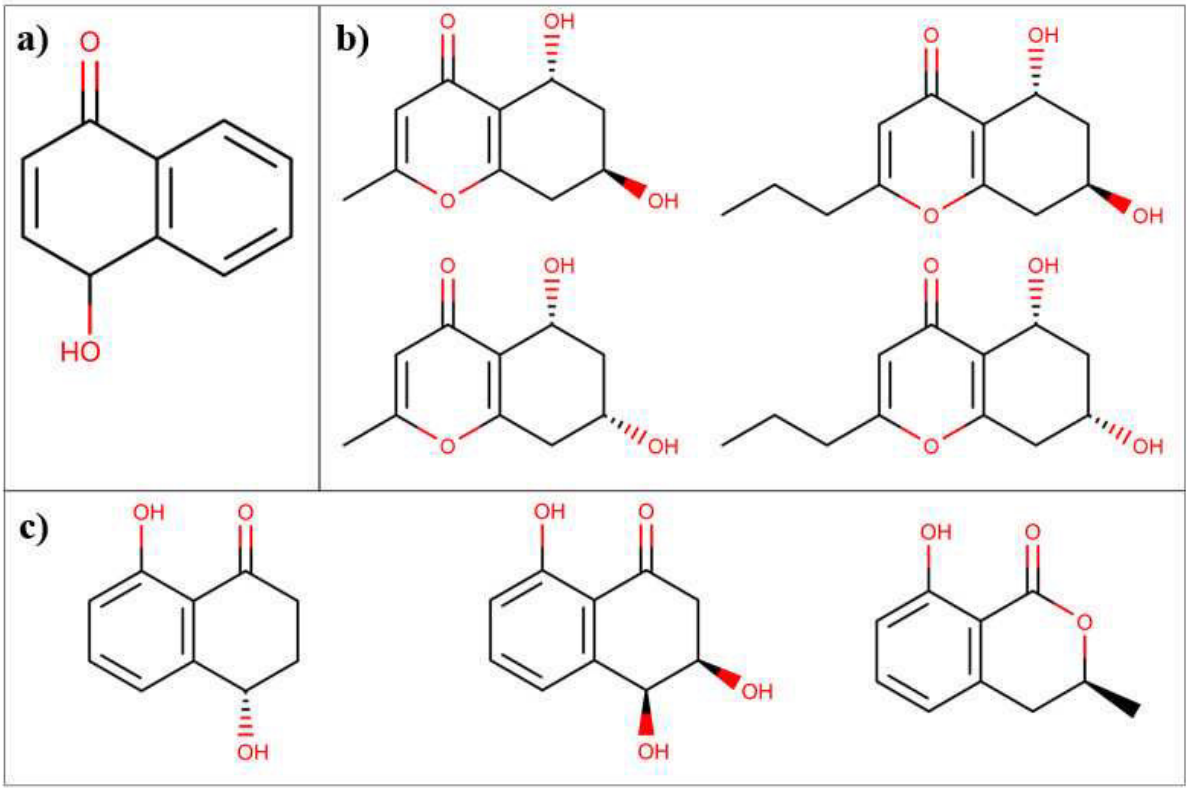
criteria rule (Pfizer, GSK and Golden triangle, with the exception of paracetamol that have 1 alert) that correlates the molecule to commercially available drugs or drug fragments from databases, the *in vitro* and *in* vivo assays are very important to assess the real effects and to feedback the ADMET algorithms and database, but also a method that considers different reliable algorithms to produce a consensus.

2.3 Natural products from *Daldinia*

Daldinia is an *Ascomycota* fungus, an important lignocellulolytic fungi with global distribution that has more than 40 species, with interesting metabolites reported as antioxidants, anti-HIV, anti-IAV, anti-ZIKV, and antibacterial, to name from the few studies of this genus. Some of the common studied species are *D. eschscholzii*, *D. concentrica*, *D. sp.*, *D. loculata*, *D. hawksworthii* and *D. childiae*, and despite this, it is understudied (LI *et al.*, 2021; YUYAMA *et al.*, 2013; ZHANG *et al.*, 2021a).

Some *Daldinia* extracted compounds are classified as ketenes (LU *et al.*, 2023), polyketide (DU; KING; CICHEWICZ, 2014; LIN *et al.*, 2022; WUTTHIWONG *et al.*, 2021), polyketones, lactone, naphtoquinones, naphtofuran, chromone, phenols and indole derivatives, the majority with interesting bioactive properties (LI *et al.*, 2021). Some of the chromones derivatives are shown in Figure 8.

Figure 8 – Chromone molecule (a), chromone derivatives from the works of Zhang *et al.* (b) and Wutthiwong *et al.* (c).



Source: Author (a), Zhang *et al.* (2021) (b), Wutthiwong *et al.* (2021) (c).

Thereby, as a wood-decaying fungus distributed globally, with bioactive molecules *in natura* and understudied, *Daldinia* natural products are potential candidates for drug discovery in the treatmen of antiviral diseases, as previously reported by Lin *et al.* and Zhang *et al.* (LIN *et al.*, 2022; ZHANG *et al.*, 2021a) for H1N1, Influenza A, and Zika.

Then, in this context, natural products from *Daldinia* can be used as a set for *in silico* development of antivirals for COVID-19 and CHIKV as innovative proposal for a base drug that could be used in future SARS epidemics or pandemics, as wll as for the treatment of Chikungunya.

CHAPTER III[1]

3 PHENOL AND CHROMONE COMPOUNDS FOR *IN SILICO* INHIBITION OF NSP2 AND NSP3 OF CHIKUNGUNYA VIRUS

Abstract

The rising concern about neglected tropical diseases imposes a global challenge, in this sense, this work brings 12 potential candidates, as the natural chromone and phenol compounds extracted from *Daldinia sp*. to inhibit the nsP2 and nsP3 of the Chikungunya virus. A molecular docking was carried to determine the energy affinity between ligand and receptor and a ADMET prediction was made. The docking simulations for the nsP2 showed mild binding, in the nsP3 all the derivatives presented -6 kcal.mol^{-1} binding affinity and interacts with crucial residues in the replication cycle of CHIKV, the 5 best were chosen as the main derivatives for ADMET. The ADMET results shows high QED values, with good oral and intestinal absorption, excretion, distribution and toxicity, with moderate (Der9 to Der12) and poor (Der8) metabolism. Therefore, the 5 derivatives are potential candidates do treat chikungunya, with a focus on Der8 as they also have *in vitro* inhibitory action against zika virus. Those candidates may progress to *in vitro* assays to treat an actual endemic tropical disease.

Keywords: Molecular Docking; ADMET; Neglected Tropical Disease; Natural products; *Daldinia*.

3.1. Introduction

Neglected tropical diseases (NTDs) are an ever-increasing world concern due to the impact in tropical and subtropical countries that struggles to eradicate them, but mostly due to climate change that may spread NTDs to novel regions and the possibility of new epidemics and pandemics in the world after the COVID-19 pandemic (ROUGERON *et al.*, 2015). Among the NTDs, Chikungunya is a disease caused by a namesake virus (CHIKV), an arbovirus like dengue and zika (SARKAR; GARDNER, 2016), which is transmitted mainly by *Aedes Aegypti* and *Aedes Albopticus*, common in tropical weather (FERREIRA; DE MORAES; ANDRICOPULO, 2022; ROUGERON *et al.*, 2015).

1. This chapter is based on the results of the submitted paper entitled “Phenol and chromone compounds for *in silico* inhibition of nsP2 and nsP3 of Chikungunya virus”, to DARU Journal of Pharmaceutical Sciences.

Dengue, zika and chikungunya have fever and rash as common symptoms, but dengue and chikungunya usually causes arthritis, arthralgia and joint pain (DA SILVA NETO *et al.*, 2022; HUSSAIN; AMIR; RASOOL, 2020). Due to the lack of antiviral for these diseases, the traditional public health actions are to prevent one or more steps of the cycle of the *Aedes* mosquitoes (entomological surveillance and vector control) and treat the symptoms of the infected. In Brazil, despite its efficiency (ZARA *et al.*, 2016), the epidemic persists, attributed to the role of the population in the vector control and the intersectoral of the health system to eradicate and treat the disease (RAMOS *et al.*, 2021).

Proposals for treatment and vector control have arisen, the majority are the genetic modification of *Aedes* (CARVALHO *et al.*, 2015; EVANS *et al.*, 2019), the development of vaccines and antivirals. In the sense of antivirals, the *in vitro* testing with already known antiviral compounds, as interferon, ribavirin, favipiravir, arbidol and amantadine, among others repurposed drugs. Nonetheless, some are not approved yet to use in humans or their concentrations are too high (POWERS, 2018). Another approach is the development of novel drugs through computational drug design, targeting a protein crucial in the replication cycle of the CHIKV, that may find a potential candidate to inhibit quickly, at the expense of starting the *in vitro* and *in vivo*, in contrast to repurposed drugs (BATTISTI; URBAN; LANGER, 2021).

Natural compounds usually shows less side effects, are cheaper and easy to obtain when compared to synthetic ones (SINGH; BHARDWAJ; PUROHIT, 2022), in the CHIKV context, the works of Puranik *et al.* (PURANIK *et al.*, 2019), Indu *et al.* (INDU *et al.*, 2022) brings the computational docking, the first targets the nsP3 (non-structural protein 3) protease with flavonoid derivatives, the second the envelope glycoprotein and nsP2 protease with phytocompounds from online databases, both works showed high potential for inhibition or to be a molecular scaffold to novel drugs.

Therefore, the proposal is to evaluate *in silico* the antiviral activity in the nsP2 and nsP3 macrodomains and the absorption, distribution, metabolism, excretion and toxicity (ADMET) of the chromone and phenol compounds extracted from *Daldinia sp.*, based on the work of Zhang and coworkers(ZHANG *et al.*, 2021a), as these natural compounds presented anti-ZIKV and anti-influenza A virus (IAV) activity, some are relatively small molecules that may interact with the targets, as described in the work of Zhang *et al.*(ZHANG *et al.*, 2021c), that was able to inhibit the nsP3. Also, the possibility to achieve a compound that is able to inhibit both CHIKV and ZIKV, which are similar diseases from the same vector, is desirable.

3.2. Methodology

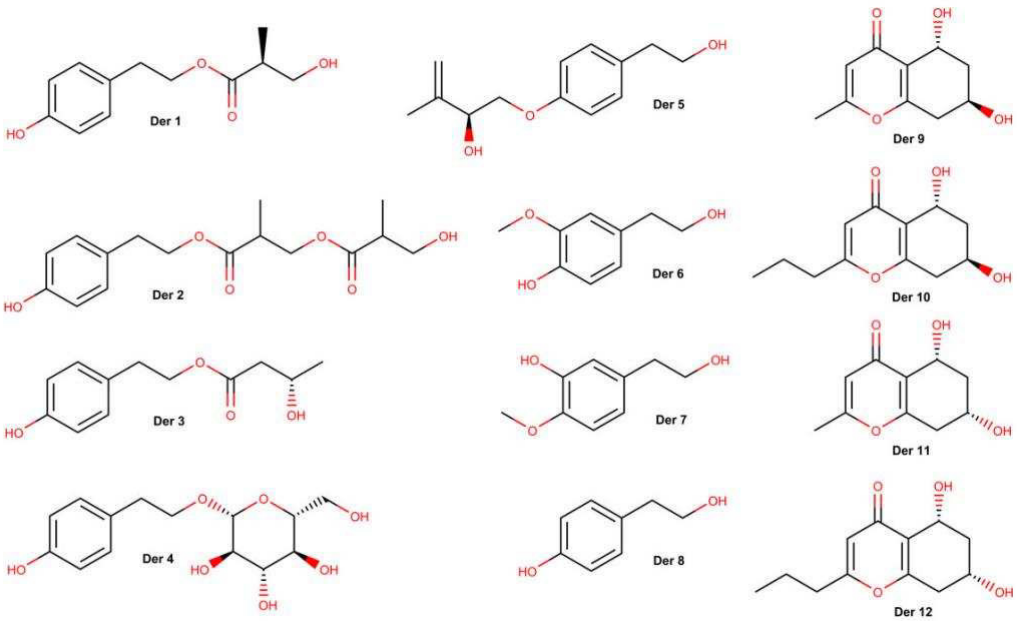
3.2.1. Ligands

The chromone and phenols molecules for Docking and ADMET, Figure 9, is based on the work of Zhang *et al.* (ZHANG *et al.*, 2021a), which extracted (HPLC, ethyl acetate) and characterized novel compounds from *Daldinia sp*. and tested their antiviral properties against zika virus (ZIKV) and influenza A virus (IAV). In this sense a computational simulation may indicate a potential anti-CHIKV for further *in vitro* testing.

Initially, the molecules can be classified in phenol and chromone-like: from Der1 to Der8 are the phenolic compounds and Der9 to Der12 are the chromone derivatives. Among the derivatives, some molecules stand out. Using the PubChem database, using the SMILES code obtained from drawn in MarvinSketch 22.22 (CHEMAXON, 2022), were not found molecules for the phenol derivatives Der1, Der2, Der3, and the chromone derivatives Der9 and Der10, wherefore those are novel molecules. Also, Der8 and Der4 are already well-known molecules, Der8 is known as Tyrosol (IUPAC name: 4-(2-hydroxyethyl)phenol), a plant (e.g. olive oil) and fungus metabolite that has antiarrihtimia, neuroprotective and antioxidant activities, for that cause it is a promising molecule for ischemic stroke (PLOTNIKOV; PLOTNIKOVA, 2020), were as Der4 is the salidroside (IUPAC name: (2R,3S,4S,5R,6R)-2-(hydroxymethyl)-6-[2-(4-hydroxyphenyl)oxane-3,4,5-triol), a glycoside commonly found on *Rhodiola rosea* (also known as golden root), that has been vastly reported its neuroprotective, in the work of Liu *et al.* (LIU *et al.*, 2017) salidroside has been tested *in vitro* (Schwann cells) and *in vivo* (rats, injured in the sciatic nerve), demonstrating good results in nerve regeneration. Also, there is reports of *in vitro* testing of salidroside as an anti-dengue virus (anti-DENV), with good indications of potential blocking of the viral cycle (LOAIZA-CANO *et al.*, 2021).

The compounds obtained and characterized by Zhang *et al.*, which is worth mentioning the characterization analysis (antiviral activities, ^{1}H NMR, ^{13}C NMR, IR and UV in particular) in the supplementar material of the original paper. The SMILES of the derivatives, link to the PubChem pages and the supplementar material of the original paper are present in Appendix B.

Figure 9 – Chromone and phenol derivatives from *Daldinia sp.*



Source: LIMA *et al.*(2023).

3.2.2. Docking

Molecular docking simulations were performed to evaluate the potential action of the novel derivatives of phenol and chromones molecules in the replication enzymes nsP2 and nsP3 of CHIKV. The structures were obtained from the repository Protein Data Bank (PDB) and identified as “Structure of Chikungunya virus nsP2 protease” (PDB 3TRK), submitted without mutations with 2.4 Å of resolution determined by x-ray diffraction (XRD), classified as a hydrolase expressed in *Escherichia coli* BL21 (R. P. D. BANK, [s.d.]) and “Crystal structure of macro domain of Chikungunya virus” (PDB 3GPG) with 1.65 Å resolution determined by XRD, expressed in *Eschericia coli* without mutations (MALET *et al.*, 2009). Both targets are presented in Appendix C. It is important to note that the original PDB file of nsP3 already nominates the chains with A (red), B (orange), C (light cyan), and D (blue).

The AutoDock Vina software was used to perform the molecular docking simulations, set to execute a Lamarckian Genetic Algorithm (LGA) (FUHRMANN *et al.*, 2010). The grid box for nsP2 was centered in 16.472, 24.972. and 24.389 Å for the x, y and z axis respectively with size parameters of 116 Å (x), 126 Å (y) and 120 Å (z) and the nsP3 in the coordinates 16.667 (x), -25,139 (y) and -1,722 (z) Å with size parameters of 126 Å (x), 126 Å (y) and 120 Å (z). To further enhance the pocket, both grid were confirmed with

DoGSiteScorer (DA FONSECA *et al.*, 2022; VOLKAMER *et al.*, 2010, 2012a), to determine the potential active sites based on the protein structure. The DoGSiteScorer was used to evaluate only the pockets, without ligands, with properties and druggability and on all the chains (A, B, C and D, given by the PDB). To select the pockets, a simple score of 0.4 or above was used.

As criteria for the protein structure, the methodology proposed by Yan and coworkers (YAN *et al.*, 2014), which were added Gasteiger charges, essential hydrogen atoms and removed water molecules. The code for the preparation was made through AutoDockTools (ADT) (MORRIS *et al.*, 2009b). A total of 30 independent simulations were realized, being able to obtain 20 poses per simulation.

To improve the partial refinement of each calculus, the exhaustiveness criteria was set to 64, all bonds and torsion of the ligands were adjusted to twist while the protein structure kept rigid (NGUYEN *et al.*, 2017). As the criteria to select the best pose, the root mean square deviation (RMSD) was carried, with ideal valors bellow 2 Å (YUSUF *et al.*, 2008). To evaluate the stability of the complex binder-ligand through the simulations, the energy affinity (ΔG) was used, with desirable values close to -6.0 kcal/mol (SHITYAKOV; FÖRSTER, 2014).

3.2.3 ADMET

For the ADMET evaluation, an approach similar to the work of Rocha (NUNES DA ROCHA *et al.*, 2022), Lima(LIMA *et al.*, 2022) and Lima(LIMA *et al.*, 2023b) was used, as the online websites utilizes different drug databases and algorithms, a consensus based on the numerical (like physico-chemical descriptors) and empirical (*in vitro* assays) can be made to predict the properties of the drugs.

The compounds were drawn in MarvinSketch v22.22 as described in the work of Zhang and coworkers(ZHANG *et al.*, 2021a), then converted to a SMILES format and .mol to submit to ADMETlab 2.0, ADMETboost and SwissADME (http://www.swissadme.ch/). In order to produce a .mol file in MarvinSketch, the force field used was the MMFF94, with maximum number of conformers of 10.

In the sense of the state of the art, AI Drug Lab was used to further improve the results of the consensus, due to the recent implementation of the ADMETboost algorithm trained in the Therapeutics Data Commons ADMET Benchmark, achieving high rank in several tasks and also has a higher degree of specificity in the results(TIAN; KETKAR; TAO, 2022), and also, uses the same limits for optimal data as ADMETlab 2.0, then being easily comparable.

For druglikeness, the Quantitative Estimation of Druglikeness (QED) and Medicinal Chemistry Evolution 2018 (MCE-18) were performed with the rules of Pfizer (logP < 3 and 40 < TPSA < 90 Å²) (HUGHES *et al.*, 2008), , GSK (MW ≤ 400 g/mol, logP ≤ 4) (GLEESON, 2008) and the Golden Triangle (GT) rule (200 < MW ≤500 g/mol, -2 < logD ≤ 5) (JOHNSON; DRESS; EDWARDS, 2009), from the ADMETlab 2.0.

The metabolism sites, reactive fragments and excretion descriptors were determined from XenoSite and ADMETlab 2.0 webservers, while the toxicity was evaluated at pkCSM and ProTox II.

3.3 Results and discussion

3.3.1. Molecular Docking

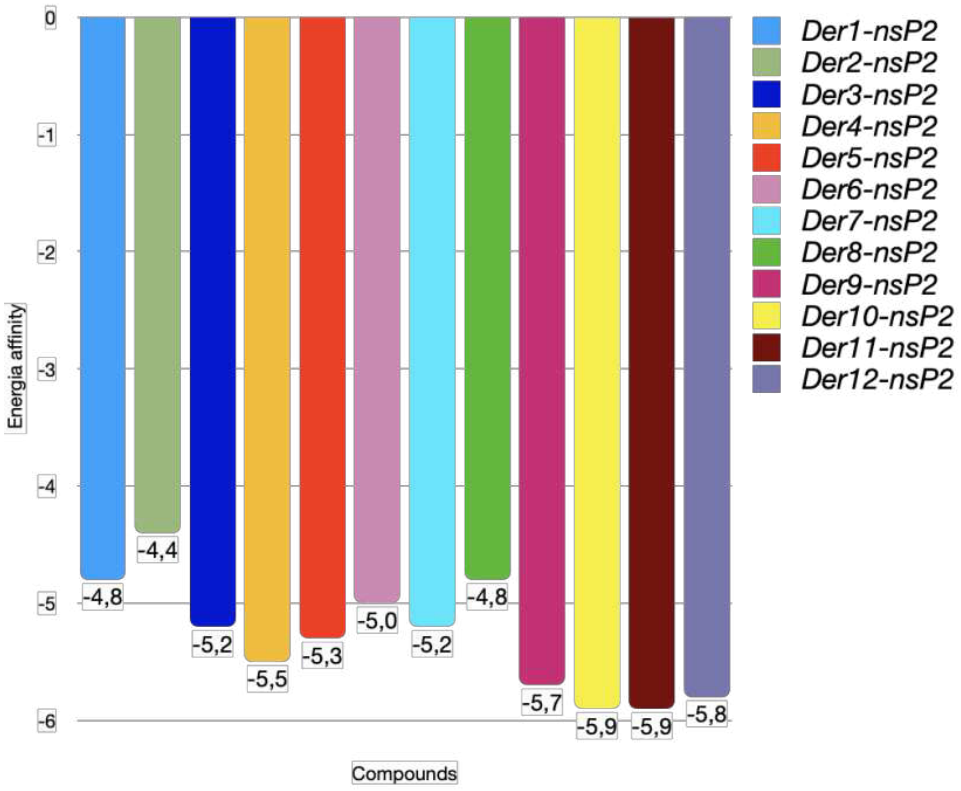
The nsP2 and nsP3 enzymes perform important roles in the viral replication, as the nsP2 exhibits phosphatase helicase and proteolytic activities, and the nsP3 synthesize the reverse if the negative RNA strip (POWERS, 2018). Therefore, the inhibition of those enzymes may be a possible treatment for Chikungunya.

The Figure 10 presents the energy affinity of the compounds with the nsP2. The majority of the complexes showed weak interactions, indicating. The best compounds being the Der9 to Der12, as they have similar structure, which differentiates the stereoisomerism in the carbon 7 (R or S) and the length of the alkyl in the carbon 2 (methyl or propyl). The weakest interaction of these four is the Der9 (- 5.7 kcal/mol), which has a methyl and R carbon, in contrast, Der11 (-5.9 kcal/mol) has a S carbon, indicating that the receptor may be sensitive to stereoisomers. Der12 (-5.8 kcal/mol) and Der11 (-5.9 kcal/mol) presents the same stereoisomerism, but the former has a propyl in carbon 2 and the latter a methyl, the small difference may indicate a small pocket in the receptor site or to achieve the lowest energy affinity a long molecule may be preferable.

It is important to note that the nsP2 is a well-studied target for molecular docking of CHIKV, in the works of Basseto *et al.* (BASSETTO *et al.*, 2013) and Das *et al.* (DAS *et al.*, 2016) long compounds were obtained with rich π-stacking sites with potential inhibitory of nsP2, however in the *in vitro* assays the first compound of Basseto was unable to inhibit, but some of their derivatives and the best compound of Das were able to, besides the importance of biological assays to reinforce computational simulations, small molecules with few π-

stacking sites may not be suited for the inhibition of nsP2, a possible explanation for the performance of the compounds presented.

Figure 10 – Energy affinity through molecular docking for the nsP2.

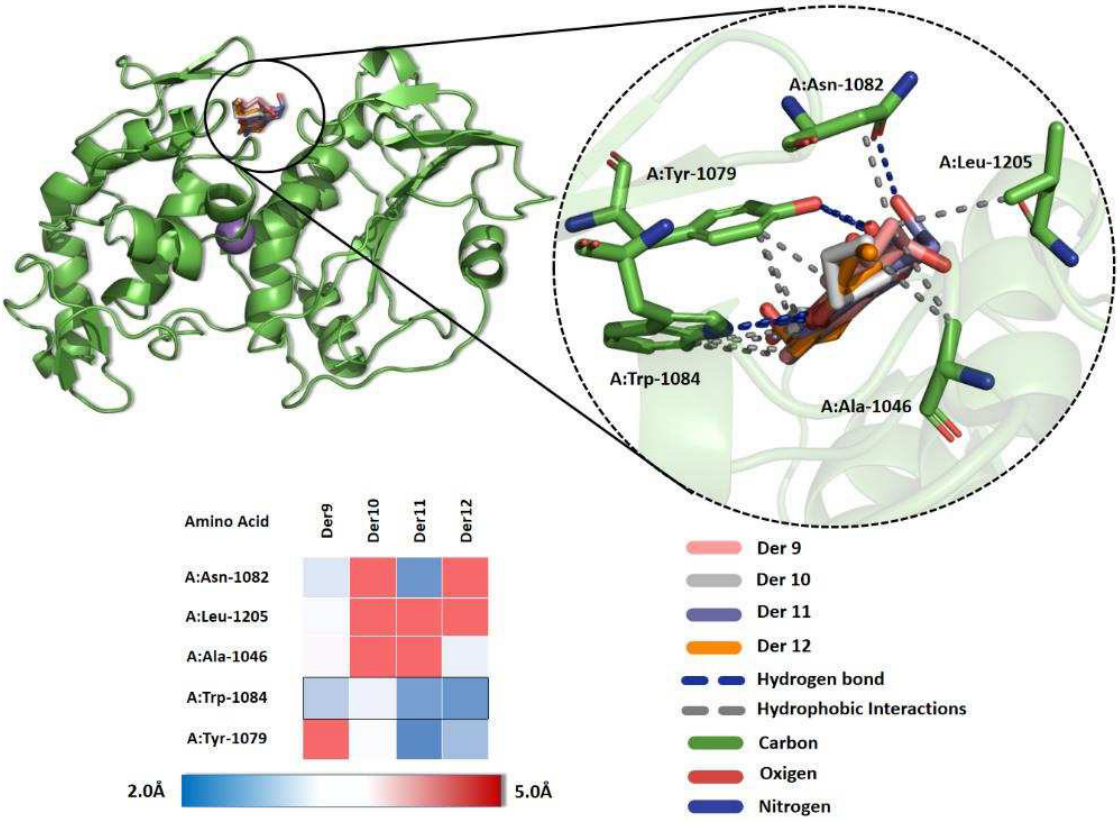


Source: Author.

A further evaluation of the interaction modes between the ligands and residues in the target, Figure 11 and Table 2, shows the influence of the stereoisomerism and alkyl group in the carbon 2. Der9 (R and methyl group) has the lowest energy affinity although it has 5 interactions, which they are predominantly hydrophobic and weak (distance > 3 Å). Der10 (R and propyl group) has 4 interactions, one of which is moderately strong (A:Trp1084, distance = 2.14 Å), responsible for it binding. Der11 (S and methyl group) has 6 interactions, 2 weaks (hydrophobic), 2 moderates (hydrogen bond) and 2 strong bonds (hydrogen bonds, A:Tyr-1079, distances: 1.85 Å; 1.96 Å). Der12 (S and propyl group) has 3 weak hydrophobic interactions and 2 moderate hydrogen bond interactions, reason that has slight higher binding energy with the ligand than Der9.

It is important to note the influence of the stereoisomerism and alkyl group in the residues interactions, as shown in the Table 2, an S isomery and methyl group has greater tendency to output the most interactions.

Figure 11 – Energy affinity through molecular docking for the nsP2.



Source: Author.

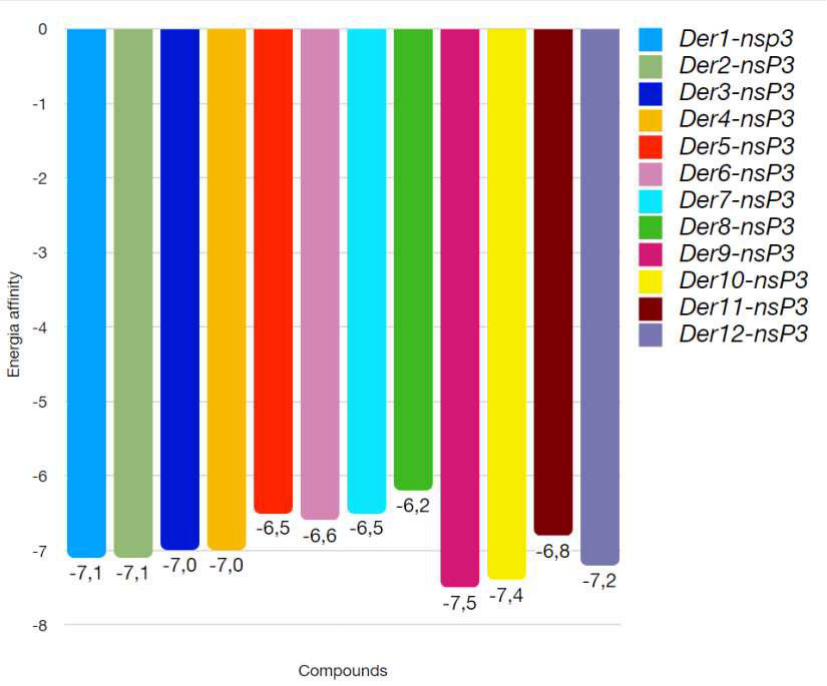
Table 2 – Interactions between the Der9 to Der12 with the nsP2 enzyme residues.

Ligand	RMSD (Å)	Receptor	Interaction	Distance (Å)
Der 9	1.821	A:Trp-1084	Hydrofobic	3.43
		A:Leu-1205	Hydrofobic	3.78
		A:Asn-1082	Hydrofobic	3.44
		A:Ala-1046	Hydrofobic	3.89
		A:Trp-1084	H-Bond	3.78
Der 10	1.865	A:Tyr-1079	Hydrofobic	3.83
		A:Trp-1084	Hydrofobic	3.65
		A:Tyr-1079	H-Bond	3.39
		A:Trp-1084	H-Bond	2.14
Der 11	1.416	A:Tyr-1079	Hydrofobic	3.87
		A:Trp-1084	Hydrofobic	3.86
		A:Tyr-1079	H-Bond	1.85
		A:Tyr-1079	H-Bond	1.96
		A:Asn-1082	H-Bond	2.08
		A:Trp-1084	H-Bond	2.21
Der 12	1.625	A:Ala-1046	Hydrofobic	3.60
		A:Tyr-1079	Hydrofobic	3.99
		A:Trp-1084	Hydrofobic	3.61
		A:Tyr-1079	H-Bond	2.73
		A:Trp-1084	H-Bond	2.06

Source: Author.

In the Figure 12, the energy affinity of the ligands and nsP3 is shown. The same criteria for selection were used for selecting the best ligands (energy affinity close to or lower than -6.0 kcal/mol). All the ligands has lower energy affinity with nsP3 than nsP2, as expected when comparing with the compounds (pyrimidone derivatives) tested by Zhang (ZHANG *et al.*, 2021c) through docking simulation, which were able to inhibit the nsP3 and they have some similarities with the chromones and phenol derivatives.

Figure 12 – Energy affinity through molecular docking for the nsP3.

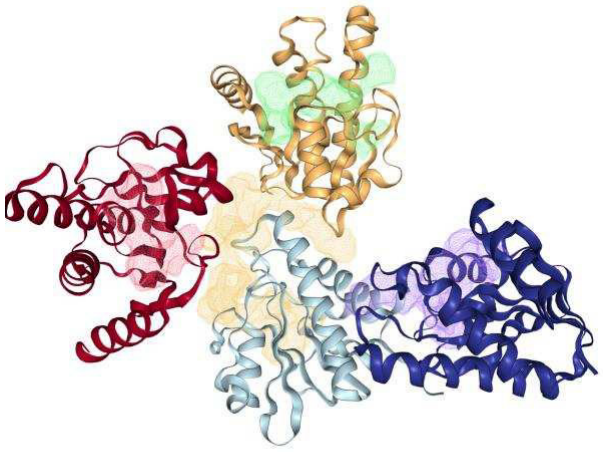


Source: Author.

The Figure 13 represents the possible binding sites of the enzyme 3GPG, DoGSiteScorer is a method based in grid that uses a Gaussian Difference (Difference of Gaussian - DoG) Filter for the detection of possible binding subpockets that is able to fit a sphere-like object, based on the density, a cluster of subpockets forms a pocket. Subsequently, physico-chemical descriptors are calculated to determine the volume, surface area, lipophilicity, that are used to calculate the simple score of each pocket(VOLKAMER *et al.*, 2012b).

In the Figure 13 its clearly stand out the central pocket (yellow), that pervades all the chains, with a simple score of 0.63 due to the high volume (1924.48 $Å^3$) and surface area (2328.39 $Å^2$). The second highest score, 0.61, is in the D chain (blue), the lilac pocket, with 949.4 $Å^3$ of volume and 1044.95 $Å^2$ of surface area. The third (green), 0.57, is in the B chain (orange), with 866.67 $Å^3$ and 1077.46 $Å^2$ and the last (red), 0.45, in the A chain (red) has 653.74 $Å^3$ and 748.74 $Å^2$. Although there is a higher probability of interaction the yellow pocket, the size of the ligands must be accounted for, so in the case of a small ligand that can fit in a small pocket may have a suitable interaction with its surrounding, while a large pocket size relative to a small ligand may indicate a greater distance of interaction, in contrast to the nsP2 simulation.

Figure 13 – Energy affinity through molecular docking for the nsP3.



Source: Author.

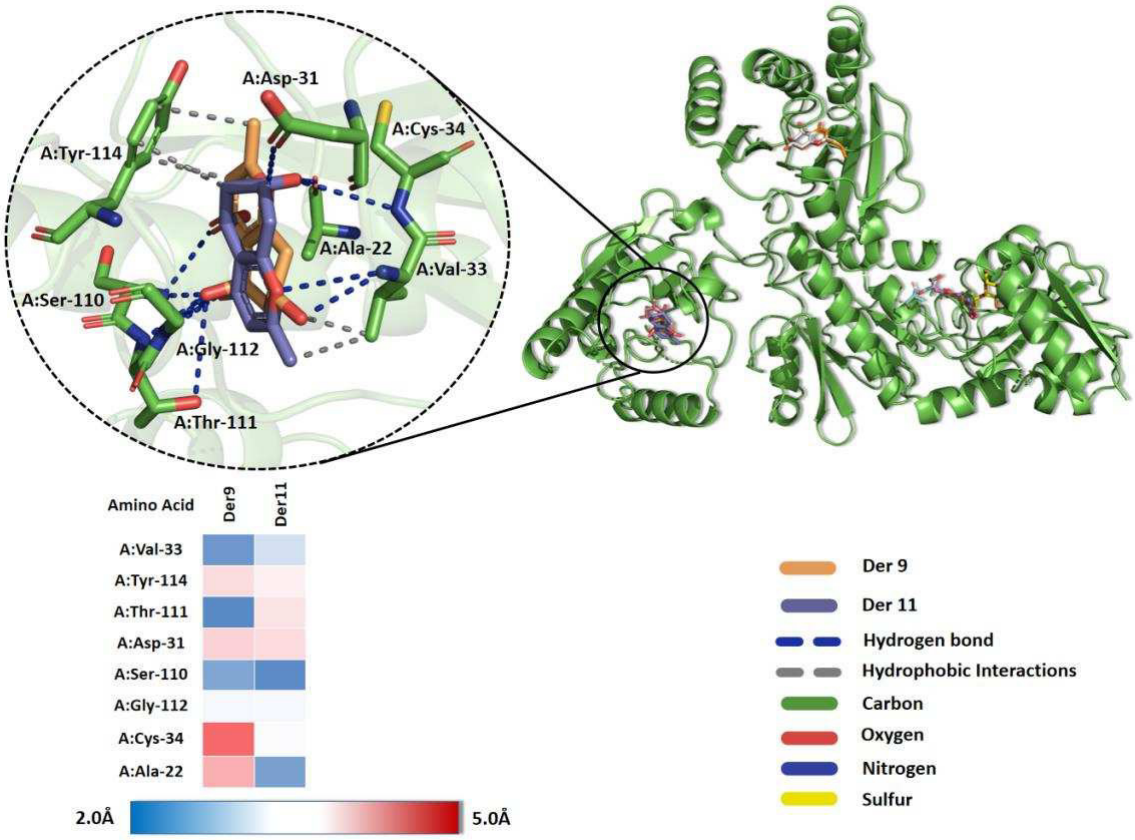
The interaction between the ligands Der9 and Der10, the two best ligands (-7.5 kcal/mol and -7.4 kcal/mol), and the nsP3 residues are presented in the Figure 14 and Table 3, which contains only the residues with the best interaction, with the exception of Der11, included due to the similarity and for comparison to the nsP2 enzyme. The main residues for the replication process of CHIKV are Asp 10, Ile 11, Asn 24, Asp 31, Gly 32, Val 33, Cys 34, Ser 11, Tyr 114, Val 133 and Arg 144, as described in the work of Puranik *et al.* (PURANIK *et al.*, 2019). The Der9 and Der11 docking results corroborate with the residues interaction, which 5 and 3 critical interactions, respectively.

As mentioned before, the main protein chains of nsP3 involved in interactions are the smallest, and their interactions have high potential of inhibition, it is an important contrast when compared to nsP2, while the ligands can fit in the yellow and lilac pockets, the distances are relatively long, decreasing the binding affinity. Also, a big pocket to a small ligand may increase instability of the binding in time, due to weak interactions with target, possible interactions with surroundings and freedom of movement, as in a small pocket site the opposite can be true, the ligand has higher probability to be entrapped because of the strong interactions and less degrees of freedom. Both cases may be evaluated by a Molecular Dynamics simulation.

Although Der8 has the lowest affinity for the target, -6,2 kcal/mol, it is interesting to note that, in the original work of the compounds, presented *in vitro* inhibitory action against ZIKV and still is inside the -6 kcal/mol against CHIKV, making it a candidate for an *in silico* simulation targeting ZIKV proteins or further studies *in vitro* to discover or propose a

mechanism of inhibition and the *in vitro* test of Der8 and CHIKV to develop as a potential drug to treat zika and chikungunya, as they share the vector and some symptoms. For an even further study, both simulations and *in vitro* assays should be made for dengue.

Figure 14 – 3D interactions of the complex of enzyme nsP3-ligands (Der9 and Der11), represented in orange (Der9) and lilac (Der11).



Source: Author.

Table 3 – Main Interactions between the Der9 to Der12 with the nsP3 enzyme residues. Only critical role in viral replication of CHIKV are presented.

Ligand	RMSD (Å)	Amino Acid	Interaction	Distance (Å)
Der9	1.821	A:Val-33	Hydrofobic	3.77
		A:Tyr-114	Hydrofobic	3.39
		A:Asp-31	H-Bond	3.52
		A:Val-33	H-Bond	2.28
		A:Ser-110	H-Bond	2.37
Der10	1.865	B:Asn-24	Hydrofobic	3.82
		B:Val-33	Hydrofobic	3.73
		B:Tyr-114	Hydrofobic	3.47
		B:Ser-110	H-Bond	2.98
Der11	1.416	A:Tyr-114	Hydrofobic	3.72
		A:Val-33	H-Bond	2.74
		A:Ser-110	H-Bond	2.20
Der12	1.574	B:Asn-24	Hydrofobic	3.81
		B:Val-33	Hydrofobic	3.68
		B:Tyr-114	Hydrofobic	3.62
		B:Ser-110	H-Bond	1.93
		B:Tyr-114	π Stacking	4.91

Source: Author.

3.3.2. ADMET

As described by Wager and coworkers (WAGER *et al.*, 2016a), more polar (40 < TPSA ≤ 90 Å²), less lipophilic (logP ≤ 3), less basic and larger than drugs available from data sets of commercially available, including from Pfizer (WAGER *et al.*, 2010), are less likely to be toxic in the nervous system *in vivo*. From those sets, some of the drugs are propofol (a phenol derivative), caffeine, clonazepam, modafinil, alprazolam, quetiapine, fluoxetin, escitalopram

and zolpidem, which are all permitted in Brazil under prescription or used as general anhestesic (BRASIL, 1998), with the exception of caffeine, from coffee drinks and anti-inflammatory drugs.

In the oral bioavailability radar in Figure 15a, it is possible to observe the physicochemical properties of the compounds Der1 to Der12 applied to the druglikeness criteria of Pfizer and GSK. The properties clearly show compounds of polarity within the ideal limit predicted by the 2016 Pfizer rule (WAGER *et al.*, 2016b) due to low logP and high TPSA values, due to the solvent extraction of the fermented material of *Daldinia sp*. with ethyl acetate (HPLC grade), as should be expected. The high hydroxylated, ether, phenolic and chromone groups also contribute to a high number of hydrogen bonding sites (HBA and HBD), relatively to the size and MW of those molecules.

In the sense of the druglikeness in Table 4, most of the molecules presents high QED values (above 0.6, in a range of 0.0 to 1.0), the main cause are the low molecular size ($MW < 300.12$ g.mol^{-1}) and low lipophilicity ($logP < 2$), with the exception of Der4, owing this to the most different molecular structure (highest TPSA, lowest number of logP and second highest MW), indicating that the Der1 to Der3 and Der5 to Der12 are similar to already known and well characterized drugs from the databases used by ADMET consensual prediction. In the MCE-18, it is possible to observe that scores ≥ 40 suggest that the compounds Der4, Der9, Der10, Der11 and Der12 present an Fsp3 of at least 50% ($Fsp3 \geq 0.5$), which basically follows the trends observed in medicinal chemistry, where it is worth highlighting the compounds Der10 and Der12 as leading compounds in occupying the ideal physicochemical space (Figure 15a) (IVANENKOV; ZAGRIBELNYY; ALADINSKIY, 2019). The physicochemical space of all the derivatives, as obtained by ADMETlab 2.0, is presented in the Appendix D.

Table 4 – Physicochemical properties of Der1 to Der12, with the Pfizer, Golden Triangle and GSK criteria for druglikness

Prop.	**Derivatives**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
logP	1.12	1.12	1.31	-0.61	1.83	0.61	0.69	0.77	-0.28	0.55	-0.37	0.50
logD	1.01	0.94	1.13	-0.65	2.13	1.02	1.09	0.94	-0.62	0.05	-0.36	0.57
MW	224.1	310.1	224.1	300.1	222.1	**168.0**	**168.0**	**138.0**	**196.0**	224.1	**196.0**	224.1
HBA	4	6	4	7	3	3	3	2	4	4	4	4
HBD	2	2	2	5	2	2	2	2	2	2	2	2
TPSA	66.76	93.06	66.76	119.6	49.69	49.69	49.69	40.46	70.67	70.67	70.67	70.67
RB	6	10	6	5	6	3	3	2	0	2	0	2
NR	1	1	1	2	1	1	1	1	2	2	2	2
SC	1	2	1	5	1	0	0	0	2	2	2	2
	Druglikeness criteria											
Pfizer rule*	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
GSK filter	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
GT rule	(–)	(–)	(–)	(–)	(–)	(+)	(+)	(+)	(+)	(–)	(+)	(–)
QED	0.73	0.70	0.73	0.46	0.71	0.70	0.70	0.63	0.62	0.78	0.62	0.78
Fsp^3	0.41	0.50	0.41	0.57	0.38	0.33	0.33	0.25	0.50	0.58	0.50	0.58
MCE-18	14.0	18.0	14.0	44.90	14.0	6.0	6.0	5.0	42.4	41.05	42.4	41.05

Source: Author.

Note: The prediction was made using the ADMETlab 2.0 platform, where the + tokens indicate an unfavorable druglikeness attribute; in bold, the descriptor that has at least one alert from the rules. *Pfizer's rule relates logP and TPSA attributes to the physical-chemical space of the ligands: low logP and high TPSA (logP < 3 and TPSA > 75 Å²), low toxic risk; high logP and low TPSA (logP > 3 and TPSA < 75 Å²), toxic risk.

The Table 5 brings the pharmacokinectics descriptors of the derivatives, which the results expressed in terms of percentage are considered excellent or moderate when between 30-70% (- token) and poor above 70% (+ token). A high Madin-Darby Canine Kidney cells passive permeability (P_{app} MDCK) is related to high oral bioavailability, with standard threshold value of $2x10^{-6}$ cm/s, below that would be estimated to have a poor oral absorption. The results are expected because of the size and $logD_{7.4}$, as described in the work of Johnson and coworkers (JOHNSON; DRESS; EDWARDS, 2009). It is worth highlighting that the compounds with the best druglikeness scores, that is, Der10 and Der12, also obtained the best fit to the golden triangle rule (GT rule), especially due to the lowest logD values of positive order (0.05 and 0.57, respectively), within a range of MW 200-500 g/mol (Figure 15b), with P_{app} MDCK values on the order of $1.6x10^{-5}$ and $2.3x10^{-5}$ cm/s, respectively.

As empirical data and literature reports, low molecular weight compounds usually are easily metabolized, nonetheless, some reactive metabolites can be originated in second

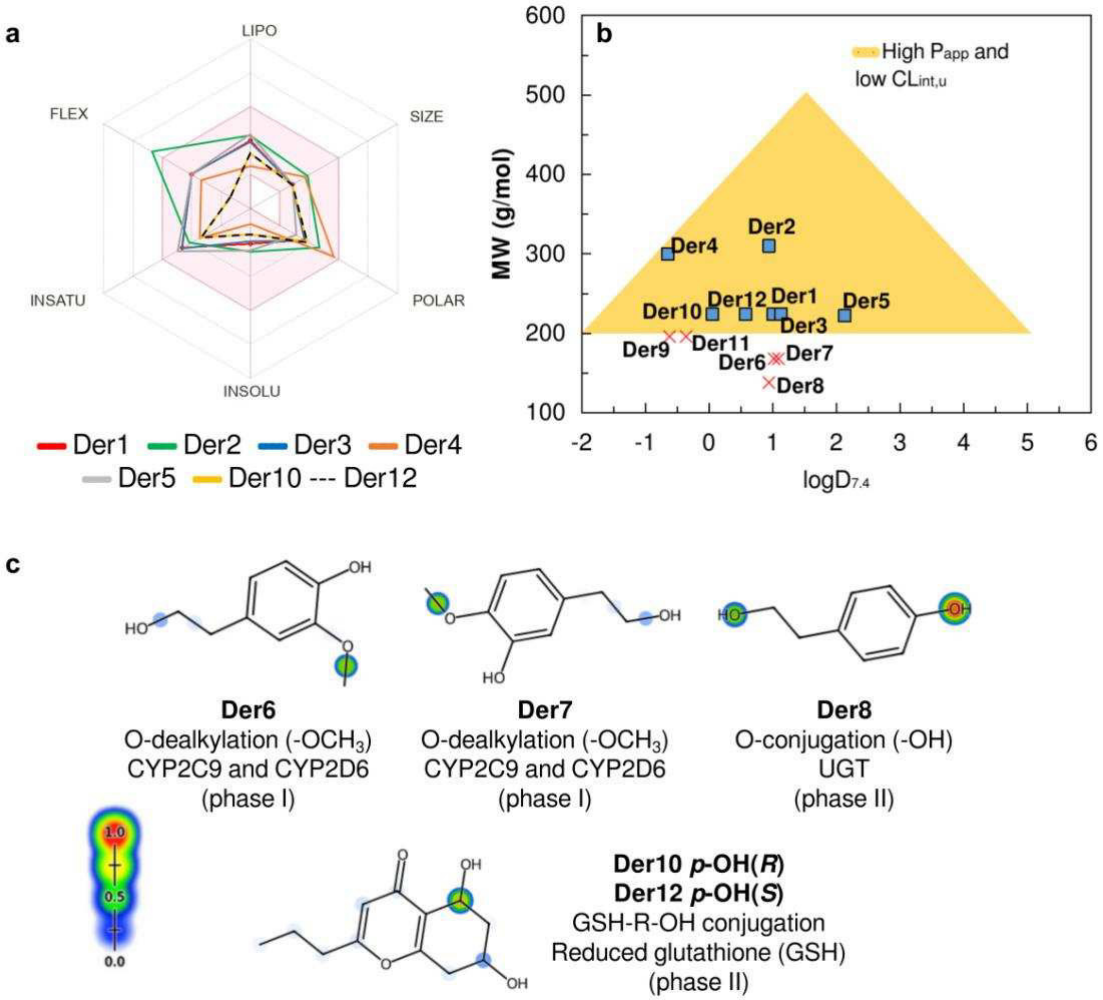
phase metabolism, especially in aromatic structures (HUGHES; MILLER; SWAMIDASS, 2015). With the GT rule plot, it was possible to observe that compounds Der6 to Der8 are more displaced from the ideal physicochemical space for good metabolic stability, due to MW < 200 g/mol (Figure 15b), being compounds more susceptible to metabolism of phase I (by CYP450) and phase II (by UGT). With the prediction of the site of metabolism, it was possible to observe that the compounds Der6 and Der7 have MW < 200 g/mol and are O-dealkylated by the CYP450 isoforms 2C9 and 2D6 due to the presence of a highly specific -OCH3 group, while the absence of this functional group increases the susceptibility of Der8 to undergo O-conjugation (-OH groups) by the UGT in phase II of metabolism. Despite the presence of metabolism sites, the compounds do not pose a risk of human hepatotoxicity (HT) or Ames mutagenicity due to the formation of reactive secondary metabolites (Table 5).

On the other hand, it is important to highlight that the compounds Der10 and Der12, despite the stereoisomeric difference in the p-OH position, can undergo conjugation with reduced glutathione (GSH) of the GSH-R-OH type in their o-OH groups, constituting a metabolite capable of covalently interacting with proteins and DNA (Figure 15c), presenting a structural warning regarding the HT model (Table 5).

The low probability of being a P-glycoprotein (P-gp) substrate of Der1 to Der8 indicates that these molecules can be easily absorbed in the intestine, in conform with elevated human intestine absorption (HIA), with HIA values estimated to be at least 70% (Table 5). In contrast to Der9 to Der12, which the P-gp may favors the excretion. It should clearly reflects in the clearance rate and half-life, although if safe for use as an efficient drug to treat chikungunya, novel methods of drug delivery can be used to bypass this, as example, in the form of a delayed release drugs, functionalization or core and shell nanoparticles with biocompatible molecules or materials (SEIDENSTUECKER *et al.*, 2017).

The plasma protein binding (PPB) indicates the interaction of the molecules with the proteins in blood serum, therefore an important descriptor for distribution and excretion, related to oral bioavailability. Due to the lipophilic character of these proteins and high polarity of the derivatives, the PPB is low (< 90%), showing that they can be transported with ease along the body. It's related with the volume distribution (VD), that corroborates with the PPB. Also, the capability to cross the blood-brain barrier (BBB) to act in the CNS, is low, due to the low lipophilicity(WAGER *et al.*, 2010), similar to PPB. Should be noted that ADMETlab 2.0 express BBB in terms of cm/s, were as the logBB > -1 is classified as BBB+ and logBB ≤ -1 is classified as BBB-, the result expressed in the Table 5 is the probability of BBB+.

Figure 15 – (**a**) Oral bioavailability radar with physicochemical properties of the most favorable ligands in relation to the druglikeness criteria: LIPO (logP < 5), SIZE (200 < MW < 500 g.mol^{-1}), POLAR (20 < TPSA < 120 Å^2), INSOLU (logD < 4), INSATU (Fsp3 > 0.5) and FLEX (RB < 10). (**b**) Alignment between MW and logD to estimate the ideal physicochemical space for P_{app} and $CL_{int,u}$ descriptors. (**c**) Site of metabolism prediction for the Der6 to Der8, Der10 and Der12 compounds.



Source: Author.

Table 5 – Predicted ADMET properties for Der1 to Der12 compounds.

Prop.	Derivatives											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Absorption												
P_{app}* ($x10^{-5}$)	2.1	2.5	2.1	3.7	1.4	1.5	1.5	1.6	11	1.6	13	2.3
P-gp%	(–) 5.4	(–) 0.1	(–) 2.7	(–) 4.5	(–) 0.6	(–) 3.9	(–) 3.9	(–) 0.3	(+) 99.7	(+) 99.9	(+) 99.7	(+) 99.8
HIA%	73.0	71.3	73.2	72.0	73.1	73.2	74.5	73.6	73.2	72.9	73.2	72.9
Distribution												
VD*	0.77	0.82	0.67	0.74	1.86	1.24	1.27	3.17	1.01	1.02	0.94	0.81
PPB%	56.6	55.3	58.3	38.9	49.1	42.8	42.5	38.9	44.2	46.8	44.2	46.8
BBB%	24.4	27.1	24.8	22.2	31.3	33.3	33.2	28.9	30.2	33.0	30.2	33.0
CYP450 metabolism												
2C9	(+)	(–)	(+)	(–)	(–)	(+)	(+)	(+)	(+)	(+)	(–)	(+)
2D6	(–)	(–)	(–)	(–)	(+)	(+)	(+)	(–)	(–)	(–)	(–)	(–)
3A4	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
Excretion												
$CL_{int,u}$*	15.1	12.2	14.8	2.5	12.2	12.6	12.8	13.4	3.3	3.4	4.7	4.7
Organic Toxicity												
HT%	(–) 13.1	(–) 18.6	(–) 7.2	(–) 04.5	(–) 27	(–) 9.4	(–) 12	(–) 4.8	(+) 84.4	(+) 95.3	(+) 73.3	(+) 91.9
Ames %	(–) 9.4	(–) 3.4	(–) 12.2	(–) 30.2	(–) 12.5	(–) 17.7	(–) 15.9	(–) 10.3	(–) 18.3	(–) 33.2	(–) 5.9	(–) 19.4

Source: Author.

When compared to literature, the ADMET results of Der8 agrees with the reports, as in the rapid oral absorption and excreted in about 8h, through kidney, phase II UGT and the ortho-sulphate (SULT) metabolic pathways. The sulphated product of Der8/Tyrosol (Tyr 4-O-sulphate) may protect cells from oxidized cholesterol and enhances the functionalities of the high density lipoprotein (HDL cholesterol) through its antioxidant property. Also, an important care when to prescribe Der8 for chikungunya is that the diagnostic must be correct, because in a case of a false negative for hemorrhagic dengue (concurrent disease that shares symptoms and vectors), the reported reduction of the blood viscosity can lead to a worsening clinical condition (MARKOVIĆ *et al.*, 2019).

3.4 Conclusion

The novel natural phenolic and chromone derivatives presented mild potential to inhibit the nsP2, but outstanding potential to inhibit the nsP3 enzyme of CHIKV. The docking simulation showed binding affinity below -6.2 to -7.5 $kcal.mol^{-1}$ to the nsP3 target, with strong interactions with residues crucial to its replication cycle. While the ADMET results shows that

the best docking molecules may have some drawbacks, the Der8 are one of the best candidates, for having a binding affinity of -6.2 kcal.mol^{-1}, high QED with good absorption, distribution, toxicity and excretion and moderate metabolism (CYP2C9), that can be address through a design of novel drugs, like virtual screening and *de novo* design, although the metabolites generated less likely to be toxic. But it has many advantages over the others, as it is already effective *in vitro* against ZIKV, a disease that shares the same vector and some symptoms, is an anti-inflammatory and viable for treatment of some cholesterol related conditions. As a natural compound with potential to treat chikungunya and zika, a step further in the sense of *in vitro* assays and *in silico* for dengue virus, should be made to treat endemic diseases.

CHAPTER IV[2]

4. *DE NOVO* DESIGN OF BIOACTIVE PHENOL AND CHROMONE DERIVATIVES FOR INHIBITORS OF SPIKE GLYCOPROTEIN OF SARS-COV-2 *IN SILICO*

Abstract

This work presents the synthesis of twelve phenol and chromone derivatives, prepared by the analogs, and the possibility of conducting an *in silico* study of its derivatives as a therapeutic alternative to combat the SARS-CoV-2, pathogen responsible for COVID-19 pandemic, using its S glycoprotein as a macromolecular target. After the initial screening for the ranking of the products, it was chosen which structure presented the best energy bond with the target. As a result, derivative 4 was submitted to a molecular growth study using artificial intelligence, where 8,436 initial structures were obtained that passed through the interaction filters and similarity to the active glycoprotein pocket through the MolAICal computational package. Thus, 557 Hits with active configuration were generated, which is very promising compared to the BLA reference link for inhibiting the biological target. Molecular dynamics also simulated these compounds to verify their stability within the active protein site to seek new therapeutic propositions to fight against the pandemic. The Hit 48 and 250 are the most active compounds against SARS-CoV-2. In summary, the results show that the Hit 250 would be more active than the natural compound, which could be further developed for further testing against SARS-CoV-2. The study employs the *De Novo* approach to design new drugs, combining artificial intelligence and molecular dynamics simulations to create efficient molecular structures. This research aims to contribute to the development of effective therapeutic strategies against the pandemic.

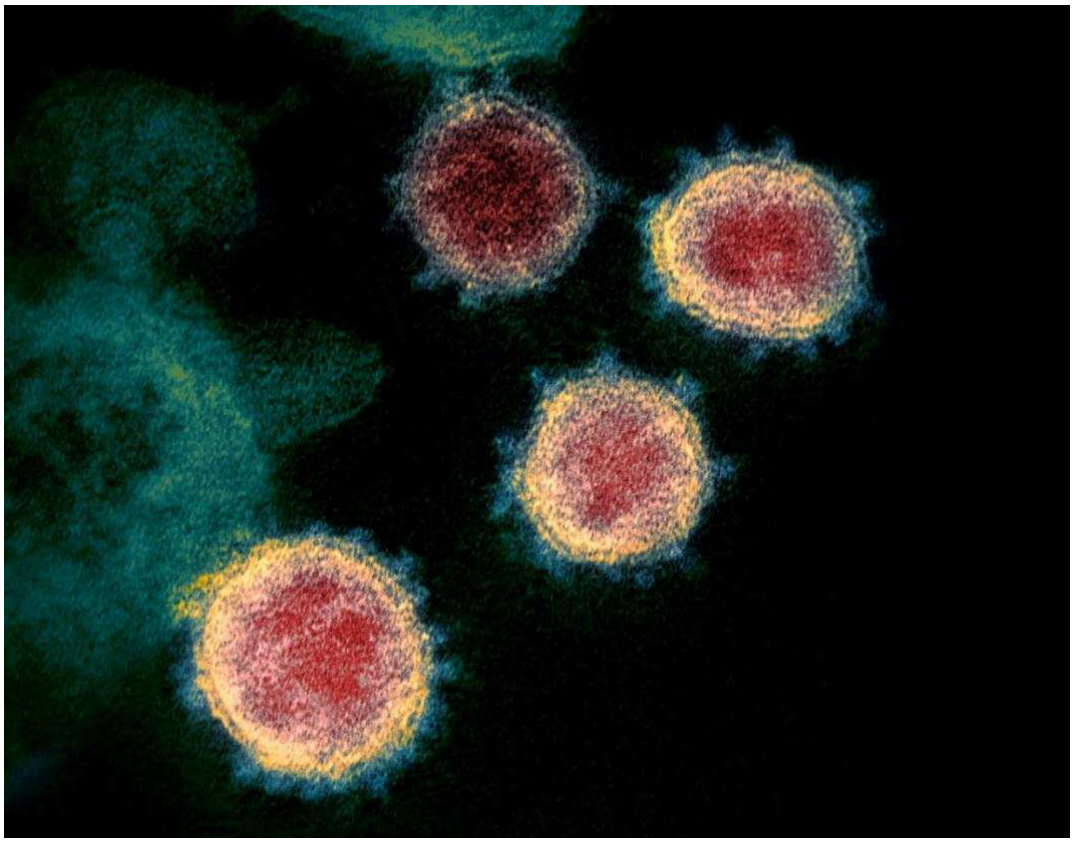
Keywords: main protease; COVID-19; molecular docking; pandemic; deep learning.

2. This chapter is based on the results of the paper entitled "*De novo* Design of bioactive phenol and chromone derivatives for Inhibitors of Spike glycoprotein of SARS-CoV-2 *in silico*", published in the 3 Biotech.

4.1 Introduction

The worldwide outbreak of COVID-19 caused by the novel SARS-CoV-2 virus (Figure 16) has generated a significant health problem (ROSA *et al.*, 2021) WHO declared this outbreak a pandemic, and China has been the most affected country. It is one of the most challenging problems of the 21st century and has changed the lives of people all over the world. SARS-CoV-2 is a novel coronavirus that is responsible for the outbreak of COVID-19, a disease that resembles SARS or MERS coronaviruses (ZHANG *et al.*, 2021b). It is classified as a new type of epidemic pneumonia with a high mortality rate (GE *et al.*, 2021). The disease is transmitted from person to person, and its symptoms are fever, coughing, and pneumonia (THANH TUNG *et al.*, 2020).

Figure 16 – Transmission electronic microscopy, digitally colored, of SARS-CoV-2 isolated getting out of the cellular surface in a culture medium.



Source: *National Institute of Allergy and Infectious Diseases – Rocky Mountain Laboratories* (2020). Image licensed under *Creative Commons* 2.0 (CC BY 2.0), <https://creativecommons.org/licenses/by/2.0/deed.pt>, assigned to NIAID-RML. No alteration has been made, original version:<https://www.flickr.com/photos/niaid/49534865371/>. All the rights reserved to the original authors.

In response, researchers and doctors rushed to find a suitable treatment for COVID-19, repurposing drugs like hydroxychloroquine, ivermectin, and remdesivir, to name a few. Many of those drugs showed no benefit and in some cases even harmful effects (FERREIRA *et al.*, 2021; SHIRAZI *et al.*, 2022). Some promising drugs are natural molecules and their derivatives, which are generally cheaper and more available than synthetic drugs (SINGH; BHARDWAJ; PUROHIT, 2022), like limonoids (OLIVEIRA *et al.*, 2021), tangerentin (DA ROCHA *et al.*, 2021), resveratrol, emodin, naringenin (CHAKRAVARTI *et al.*, 2021) that may interact with the ACE2, S protein or Mpro, the main targets for COVID-19 treatment.

In this sense, Zhang *et al.* (2021) found phenols and chromones derivated from *Daldinia sp.* had antiviral and antibacterial properties. In his work, the molecules (Figure 2) and their derivatives assays showed anti-ZIKV (zika virus) and anti-influenza activities (ZHANG *et al.*, 2021a).

As the need for rapid screening of drugs and testing, amid a pandemic, computational tools can be an alternative to traditional drug development, as it's effective and cheaper. Among the diverse approaches, Deep learning is a prevalent form of artificial intelligence that has been successfully applied in medical diagnostics, cell image analysis, organic synthesis, drug classification, and others(KERMANY *et al.*, 2018; MIAO *et al.*, 2019; MOEN *et al.*, 2019; SEGLER; PREUSS; WALLER, 2018).

There are some options when looking for a drug planning tool using the deep learning model and classical algorithm that can perform the growth of fragments in the initial seed in the active pocket of receptors via the genetic algorithm. However, the MolAICal computational package is a representative, free, easy-to-install software for planning new drugs that can grow ligands fragment by fragment in the receiving bag. Another highlight is the Vinardo score, which MolAICal uses to evaluate the protein bag's binding affinity of growth ligands. In addition, MolAICal can also filter growth ligands according to synthetic accessibility (SA), Lipinski's rule of five, and Pan-assay interference compounds (PAINS) when new drug design tasks are being performed.

Based on the characteristics and merits of deep learning and classical programming, the MolAICal package is programmed to design 3-D drugs in the specific cavity of the protein (BAI *et al.*, 2021). The MolAICal package contains two modules written in the JAVA language.

Within this perspective, the present work proposes the search for new molecules based on phenol and chromones extracted from *Daldinia sp.*, that exhibited some *in vitro* inhibitory antiviral properties, as efficient inhibitors of SARS-CoV-2 through De Novo design,

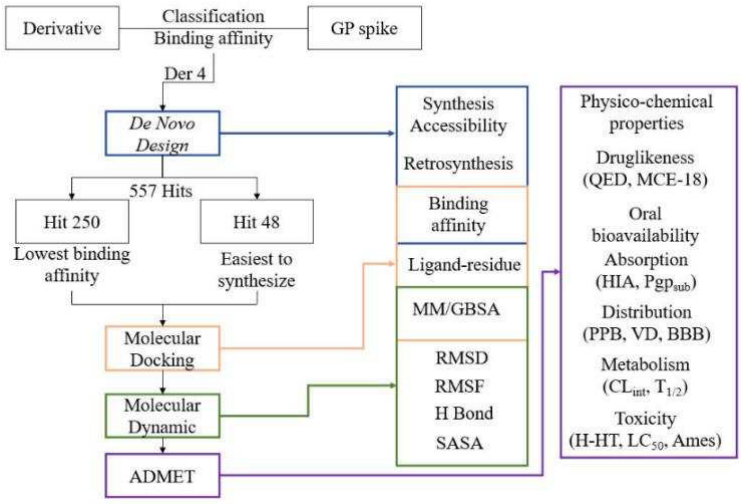
aided by MolAICal artificial intelligence and additional study of molecular dynamics to generate new compounds that scores the highest binding affinity to the Spike glycoprotein and synthetic accessibility, with further investigation about the drug behavior with an absorption, distribution, metabolism, excretion and toxicity (ADMET) to provide their feasability of these new molecules as therapeutic agents.

4.2 Methodology

4.2.1 In silico study

An overall view of this work is simplified and represented in the Figure 17. First, an evaluation of the binding affinity of the derivatives and BLA (Biliverdin IX Alpha) with the Spike glycoprotein, through molecular docking, is realized to rank and select the lowest binding affinity from the compounds. In the second step, the selected molecule was used in the *De Novo Design*, in MolAICal, obtaining some potential drugs, from which the lowest binding affinity and highest synthetic accessibility are chosen as the best potential drugs and starting points to the furter steps. In the sequence, study and evalution of molecular docking (binding affinity, interaction with protein residues and MM/GBSA), molecular dynamics (RMSD, RMSF, H-bond, SASA) and ADMET (druglikeness, MCE-18, pharmacokinetic prediction, metabolism and oral toxicity) is performed to assess the potential to an *in vitro* and clinical trials of these molecules, using the BLA as a reference molecule.

Figure 17 – General workflow.

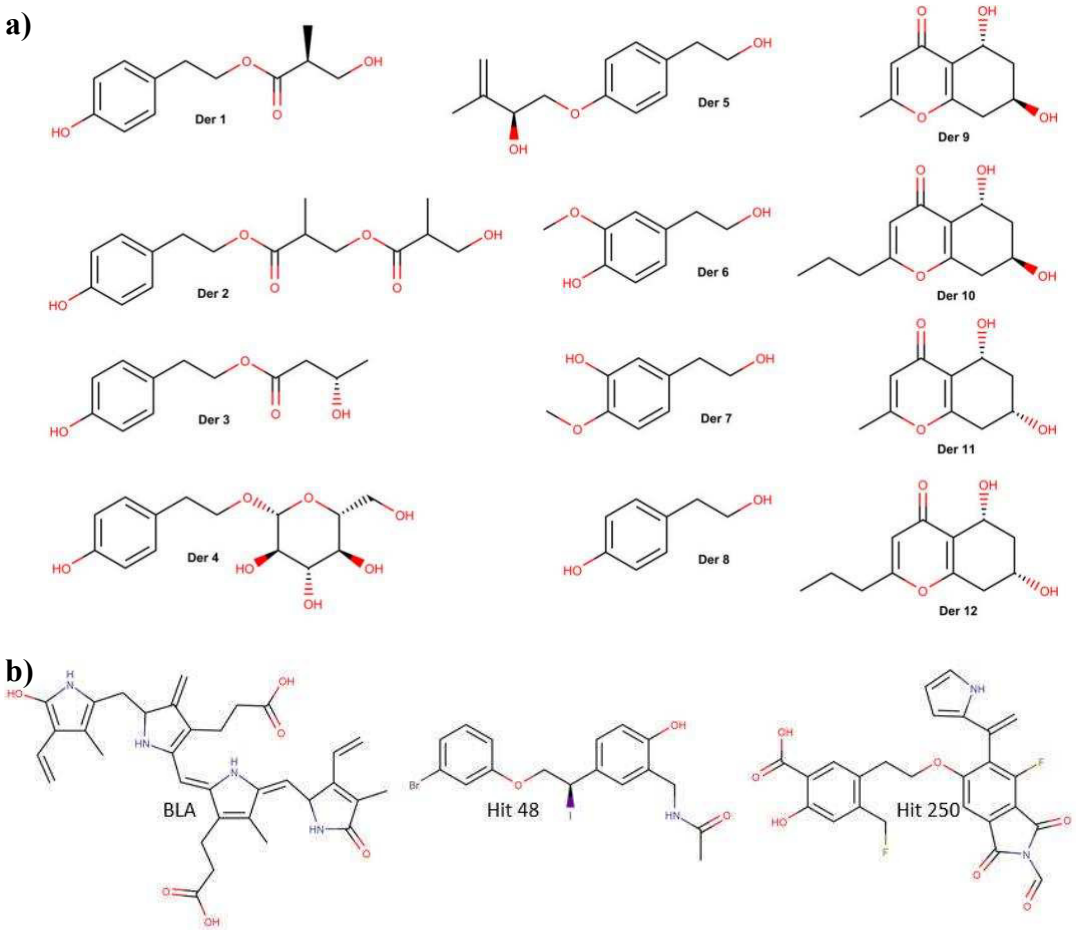


Source: Author.

4.2.2. Preparation of binders and proteins

The derivatives (**1**-**12**), from the work of Zhang *et al.* (ZHANG *et al.*, 2021a) (Figure 18), BLA, and the Hits 48 and 250 were created in Chem3D software (AHMADI *et al.*, 2005). The structures obtained in 3D were submitted subsequently to auto-optimization settings which was applied the force field MMFF94S (WAHL *et al.*, 2019), to generate bioactive conformations by minimization of randomly generated conformers, with algorithm Steepest Descent algorithm (PETROVA; SOLOV'EV, 1997), and Step per Update 4 (SUTTON; MAHMOOD; WHITE, 2016) by software AVOGADRO® (HANWELL *et al.*, 2012). All files with ligands were converted to corresponding formats (.mol2 and .pdbqt) with the addition of ionization and tautomeric states at pH 7.4 by using OpenBabel ver. 3.0.0 software (O'BOYLE *et al.*, 2011).

Figure 18 – Chromone and phenols derivatives, from *Daldinia sp*. (a). BLA, Hit 48 and Hit 250 (b).

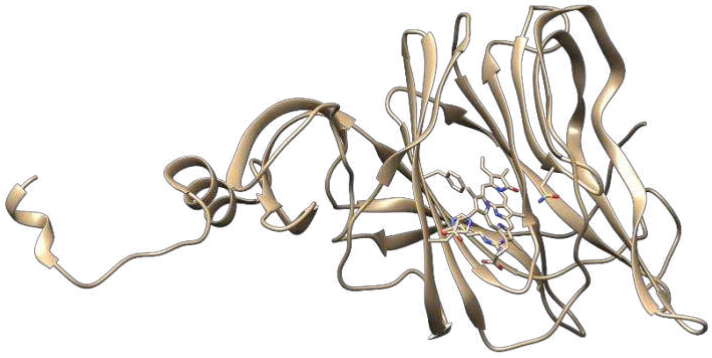


Source: Author.

4.2.3. Protein Structural Preparation

The receptor under study was the spike glycoprotein (glycoprotein S or E2) of SARS-CoV-2, obtained from the protein database repository code (PDB) ID 7B62 (ROSA *et al.*, 2021), whose crystalline structure was obtained by X-ray diffraction. To validate the simulations, the redocking technique was performed on the co-crystallized ligand, biliverdin ix alpha (BLA), which was in the original file of the co-crystallized protein. In addition, the interfering residues, water molecules, and synthetic inhibitors were removed. Polar hydrogens were added to binders and protein separately. The used software was AutoDock Tools (MORRIS *et al.*, 2009a, 2009b). The BLA and GP spike, without water, interfering residues and not showing the hydrogens are presented in Figure 19, made in UCSF Chimera 1.17.1 (PETTERSEN *et al.*, 2004).

Figure 19 – BLA + GP spike protein structure. No interfering residues and water.



Source: Author.

4.2.4. Deep Learning Model and De Novo Drug Design

MolAICal contains the deep learning generator model of the drug, which was trained from 21,064 FDA-approved drug fragments. The 90 fragments generated by MolAICal and another 30 primary fragments were mixed for fragment growth in the cavity of the spike glycoprotein.

4.2.5. Grid coordinates

The x, y, and z coordinates of the center of the cavity box of the glycoprotein were

set to 21.404, 14.571, and -18.006 Å, respectively. The cavity box lengths of the protein will be set to 30.0 Å along the x, y, and z directions. The fittest molecules were extracted for the subsequent evolved growth of 10% of the generated molecular populations. The 140 best molecules of generated molecular populations will be developed as the mother molecules. Thus, over 60 molecules were randomly selected from the generated molecular populations to increase the diversity and novelty of growth ligands. The maximum population was set at 3000.

4.2.6. Fibonacci Points, Lipinski filter, and Interference

Fibonacci's 361 points are generated for the search for fragment disturbance, using the golden angle to distribute the points of the subsequent fragments from the initial growth fragment in the Spike glycoprotein pocket. Them when the fragments grow and forms a ligand, a genetic algorithm is applied to optimize molecular conformation of the ligand. Crossover and mutation operators were set to 1.0 and 0.5, respectively.

According to Lipinski's rule of five, a set of rules of the physico-chemical descriptors that encompasses most drugs used for druglikeness and ADMET, values of crystal binders in the glycoprotein was be defined for the values of XLOGP (5.0), hydrogen acceptors (10), hydrogen donors (5), molecular weight (500), and rotary bonds (10).

Pan-assay interference compounds (PAINS) are compounds that may not have a therapeutic effect *in vivo* despite showing *in silico* fitting scores (BAELL; WALTERS, 2014), these compounds that usually are false positives, in the case of genetic algorithm may induce a false convergence of the optimal solution, were filtered out of unwanted growth binders.

After the genetic algorithm produced a generation, MolAICal uses a Vinardo score, to select the fittest ligands for the next generation. For this, it considers steric interactions, hydrophobicity and H-bonds to evaluate the affinity between ligand and protein pocket.

4.2.7. Synthesis Accessibility

MolAICal has the Ambit-SA library, which allows to evaluate how easy a compound can be synthesize. This library utilizes 4 scores: molecular, stereochemical, fused and bridged systems complexity, which ranges from 0 to 100 (hardest to easiest). The synthetic accessibility score of the growth ligands was saved in the statistical results file at the end of the simulation. A total of 30 cycle generations were carried out for the entire drug design process.

A total of six parallel drug design processes were carried out at protein. The spike glycoprotein of the generated binders was saved between 480 and 785 from the molecular weight.

To further investigate the viability of the molecules, a retrosynthesis evaluation of the Hit 48 and Hit 250 were made, to show the real viability to obtain through organic synthesis mechanistic route, their inputs and costs. For this, it was used the Manifold, an online platform that allows to input a SMILES string or draw a chemical structure to develop a synthetic route based on database of reagents that already exist and are commercially available internationally. The routes can be classified by the number of steps, cost, database (PubChem, SureChembl), stores (Sigma, Enamine, Chemspace, Emolecules, Mcule, Moplport, Wuxi) and delivery time (1 to 12 weeks), for each step of the route the name of the mechanism and reaction, which the literature and patents that are based on for the majority of the mechanisms proposed. Futhermore, Manifold also presents the physicochemical descriptors and medicinal chemistry alerts of the final product, with each alert to a fragment of the product theres a citation or database (for instance, ZINC) that indicates the probable problem in each fragment.

4.2.8. Processing

A total of 30 CPU multicores were executed in parallel for the entire molecular growth process. Drug as the whole design process combined with deep learning model and classical programming was carried out automatically by MolAICal's designed package.
Molecular generation against the target protein was performed using the MolAiCal computational package (BAI *et al.*, 2021) on a 10th Generation Intel®™ Core Intel CPU, up to 32 GB RAM, and an NVIDIA® GeForce® GTX 1660 Ti GPU, with a scanning time set to 8h. Ten clusters were used to generate the structures. An average of approximately 8,436 molecules were developed for the target protein during the experimental period.

4.2.9. Molecular docking and dynamics general filter

For this study, it is curious to point out that the estimated Gibbs free energy (ΔG) of binding is dependent on the semi-empirical free energy force field AMBER (EBERHARDT *et al.*, 2021) which composes the Autodock Vina algorithm. While stability analysis of ligand-receptor complex formation is possible through stability analysis using MM/GBSA calculations.

4.2.10. Molecular Docking

The code used was AutoDock Vina, with its Lamarckian genetic algorithm (AG) in combination with grid-based affinity energy (Trott & Olson, 2010), with the anchor region according to the synthetic binding found co-crystallized in the protein (BLA). The Spike glycoprotein was obtained from the RCSB Protein Data Bank (PDB ID: 7B62) (ROSA *et al.*, 2021). Its structure was archived in the Protein Database with a resolution of 2.16 Å, determined from X-ray diffraction, classified as viral protein. The Lipinski's rule of five (Benet et al, 2016), RMSD of up to 2.0 Å (HEVENER *et al.*, 2009), and affinity energy less than -7.0 kcal/mol were used as an exclusion factor. The most favorable ones were represented by the lowest free binding energy (ΔG) (GURUNG *et al.*, 2016). Discovery Studio (BIOVIA, 2015) conducted interaction 3D/2D visualization analysis studies, and Poseview was added (FRICKER; GASTREICH; RAREY, 2004; STIERAND; MAASS; RAREY, 2006).

4.2.11. Molecular dynamics

Molecular dynamics (MD) simulations were performed with the program NAMD (PHILLIPS *et al.*, 2005). The best conformations obtained in molecular docking were in the water solvated case in the TIP3P model (KATO *et al.*, 2021), and in the CHARMM36-mar2019 force field (HUANG *et al.*, 2017). The preparation of the system was carried out in two steps. In the first step, the ligands were parameterized on the Charmm-Gui server (JO *et al.*, 2013), and then they were submitted to the CGenFF server for parameter identification for CHARMM36 (VANOMMESLAEGHE *et al.*, 2010).

In the second step, the protein was prepared in the NAMD program. 1 Na^+ ion per ligand was added to neutralize the total charge of the system. The latter was subjected to energy minimization by the Steepest Descent method. Then, the system was subjected to NVT and NPT equilibrations under conditions described by Langevin (FARAGO, 2019). The production simulations to study the system were performed for 100 ns. N3 was used as a standard reference drug to analyze the interactions between the ligand and the protein.

The quality of the structures obtained in MDs was evaluated using the following parameters with NAMD: Potential Energy (kcal/mol) (DIEZ *et al.*, 2014); Protein-Ligand Interaction Energy (kcal/mol); Root Mean Square Deviation (RMSD, Å) of protein, ligands and distances between them; Root Mean Square Fluctuation (RMSF, Å), minimum distances

between proteins and ligands observed in MD (ARSHIA *et al.*, 2021). Hydrogen bonds were evaluated with Visual Molecular Dynamics (VMD) (HUMPHREY; DALKE; SCHULTEN, 1996). The graphs will be generated using the Qtrace program (LIMA *et al.*, 2012; PHILLIPS *et al.*, 2005)

4.2.12. MM/GBSA calculations

MM/GBSA was calculated by MolAICal (BAI *et al.*, 2021) on the basis of the MD log file of NAMD software (PHILLIPS *et al.*, 2005). The MM/GBSA is estimated by equations **(1)**, **(2)**, and **(3)**.

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{MM}} + \Delta G_{\text{sol}} - T\Delta S \quad \textbf{(1)}$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{internal}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdw}} \quad \textbf{(2)}$$

$$\Delta G_{\text{sol}} = \Delta G_{\text{GB}} + \Delta G_{\text{SA}} \quad \textbf{(3)}$$

Where ΔE_{MM}, ΔG_{sol}, and $T\Delta S$ represent the gas phase MM energy, solvation-free energy (sum of polar contribution ΔG_{GB} and nonpolar contribution ΔG_{SA}), and conformational entropy, respectively. ΔE_{MM} contains van der Waals energy ΔE_{vdw}, electrostatic $\Delta E_{\text{ele,}}$ and $\Delta E_{\text{internal}}$ of bond, angle, and dihedral energies.

Molecular dynamics simulations are an effective tool for understanding the relationships between the structure and function of macromolecules. Thus, this means that the information obtained from the dynamic properties of macromolecules is detailed enough to challenge the conventional paradigm of structural bioinformatics, which focuses on studying unique structures, and instead allows the analysis of conformational sets. The entropic contribution can be assessed based on MD trajectories by performing MD simulations.

However, this contribution is usually ignored and, when considered, is mostly configurational rather than thermal. Configurational entropy can be estimated using trajectories based on the variance-covariance matrix of atomic positional fluctuations. A quasi-harmonic method can be used, in which the variance-covariance matrix is calculated for all atoms of the complex. In the quasi-harmonic process, the mass-weighted variance-covariance matrix is calculated from the DM trajectories using Cartesian coordinates. The global translations and rotations of the solute molecule are removed using the slightest squares adjustments of mass-weighted coordinates.

The GB method (with a, b, and c set to 0.8, 0, and 2.91, respectively, and with the default modified Bondi radii) was used to calculate the polar solvation energy, and the non-polar solvation energy was calculated using the solvent accessible surface area, according to equation **(4)**.

$$\Delta G_{\text{np}} = \gamma \, \text{SASA} + b \qquad \textbf{(4)}$$

The nonpolar component of desolvation was estimated by using the LCPO algorithm, with γ being 11.948 kcal/mol/Å^2 and b 12.862 kcal/mol. Entropy was calculated by a standard mode analysis of the calculated harmonic frequencies at the MM level. In addition to water, to increase the accuracy waste with more than 8 Å of the binders (GENHEDEN; RYDE, 2010).

In the MM/GBSA calculations, the polar component of desolvation was calculated by the modified GB model (GB$^{\text{OBC1}}$, igb = 2 in Amber18) developed by Onufriev *et al.* (ONUFRIEV; BASHFORD; CASE, 2000), the exterior (solvent) dielectric constant was set to 80 as default.

4.2.13. Statistical analysis

The results were expressed as standard error ± of each experiment. After confirming the normality of distribution and data variance homogeneity, the groups' differences were submitted to variance analysis (unidirectional ANOVA), followed by the Tukey test (BLAND; ALTMAN, 1995). All analyses were performed using Origin 8.5, with a statistical significance of 5% ($p < 0.05$).

4.2.14. In silico ADMET study

This predictive study of druglikeness properties and pharmacokinetic descriptors of absorption, distribution, metabolism, excretion, and toxicity (ADMET) was adapted from the methodologies of Lima *et al.* and Rocha *et al.* (LIMA *et al.*, 2022; ROCHA *et al.*, 2022), where different services available online constitute a consensus prediction between empirical decisions and numerical descriptors of *in vivo* and *in vitro* tests deposited in databases. Initially, the two-dimensional structural representation of the compounds was converted into a simplified

molecular-input line-entry system (SMILES) and submitted to the ADMETlab 2.0 server for quantitative estimation of druglikeness (QED) and for the similarity test with compounds registered in patents of the Medicinal Chemistry Evolution algorithm, 2018 (MCE-18), presented in equations **(5)** and **(6)** respectively.

$$QED = exp\left(\frac{1}{n}\sum_{i=1}^{n} ln\, d_i\right) \quad \textbf{(5)}$$

$$MCE = \left(AR + NAR + Chiral + Spiro + \frac{sp^3 + Cyc - Acyc}{1 + sp^3}\right) Q^1 \quad \textbf{(6)}$$

where QED is defined by the sum of the physical-chemical properties (n = 8) that are within the ideality limit (d_i), which include: molecular weight (MW), partition coefficient (logP), H-bond donors (HBD), and H-bond acceptors (HBA), Topological Polar Surface Area (TPSA), number of rotatable bonds (nRot), number of aromatic rings (nAR) and reactive molecular fragments (BICKERTON *et al.*, 2012). MCE-18 relates the distribution of sp^3 hybridization atoms between cyclic and acyclic structures, which include aromatic (AR) and non-aromatic (NAR) rings, chiral centers, and spiro-cyclic groups, where the final score expresses the degree of similarity of the compounds with substances registered in patents in recent years, where MCE-18 values > 45 show a better fit in this spectrum (IVANENKOV; ZAGRIBELNYY; ALADINSKIY, 2019). The results were compared to the druglikeness from the Pfizer Rule (optimal: logP ≤ 3, and TPSA > 75 Å²) (HUGHES *et al.*, 2008), GSK Filter (optimal: logP ≤ 4 and MW ≤ 400 g/mol) (GLEESON, 2008) and Golden Triangle rule (optimal: -2 < logD ≤ 5 and 200 < MW ≤ 500 g/mol) (JOHNSON; DRESS; EDWARDS, 2009).

And then, the SMILES code of the ligands was reported to PreADMET and ADMETlab 2.0 servers for estimation of passive permeability by the Madin-Darby Canine Kidney cells model (P_{app} MDCK), P-glycoprotein substrate (Pgp), human intestinal absorption (HIA), volume of distribution (VD), plasma protein binding (PPB) and blood-brain barrier permeability (BBB) as indicative of activity in the central nervous system (CNS). Finally, metabolism sites and reactive structural fragments were detected from the consensual structural reading between XenoSite and Stoptox servers and related to excretion descriptors, including intrinsic clearance rate ($CL_{int,u}$) and half-life ($T_{1/2}$), organ toxicity descriptors, which include human hepatotoxicity (H-HT) and Ames Mutagenicity (Xu *et al.*, 2012), as well as acute toxicity, such as median lethal dose (LD_{50}) in rats and median lethal concentration (LC_{50}) in

Minnow, performed on ProTox-II and pkCSM.

4.2.15. Prediction of NMR spectra

To further characterize the novel theoric molecues and to facilitate the work after obtaining Hit 48 and Hit 250, the hydrogen-1 nuclear magnetic resonance (^{1}H NMR) and carbon-13 nuclear magnetic resonance (^{13}C NMR) were predicted using MarvinSketch 22.22, the results are presented in Appendix E. Should be noted that the IR spectra of both compounds were tried in C6H6 IR predictor, with no success due to no response of the server.

4.3 Results and discussion

4.3.1. In silico study

Based on the virtual screening performed by AutoDock Vina, it was possible to verify the affinity energies (kcal/mol) and correlated Root Mean Square Deviation (RMSD) between binder receptor against the Spike glycoprotein, especially derivative 4, with affinity energy of -6.8 kcal/mol (RMSD 1.088 Å), evidencing moderate competitiveness, compared to the reference linker BLA with -8.1 kcal/mol (TUMSKIY; TUMSKAIA, 2021; WU *et al.*, 2022) (Table 6 and Figure 20).

All simulations performed (docking and re-docking) presented RMSD values lower than 2 Å, highlighting the *best pose* of the BLA-glycoprotein complex, which presented RMSD in the order of 2.0 Å. From the best pose choices based on the RMSD, the binding affinity of the complexes for the ligands was evaluated, where again, the complex can be highlighted, which presented energy in the order of -8.1 kcal/mol.

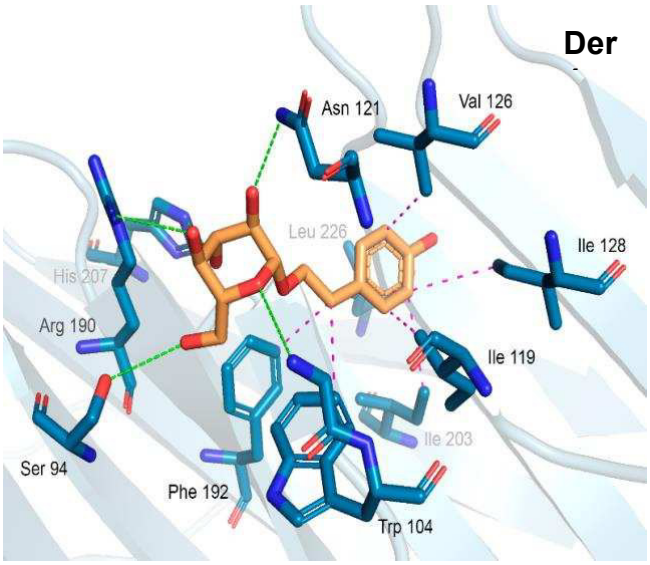
Thus, to assess the stability of the complex (proteins/ligand), the binding energy was used as a parameter, which has ideality parameters values below -6.0 kcal/mol (SHITYAKOV; FÖRSTER, 2014). Then, compound 4, which interactions with GP spike residues are presented in Figure 20 was used as a starter for drug design *De novo* to potentiate binding capacity and bring new bioactive structures.

Table 6 – Classification of the compounds accordingly to its binding Affinity in the molecular docking.

Compounds	Binding Affinity
	AutoDock vina (kcal/mol)
BLA	-8.1
Der 4	-6.8
Der 12	-6.5
Der 2	-6.4
Der 1	-6.3
Der 9	-6.3
Der 3	-6.2
Der 5	-6.2
Der 10	-6.2
Der 11	-6.1
Der 6	-5.2
Der 7	-5.2
Der 8	-5.1

Source: Author

Figure 20 – 3D interactions of Der4 and the residues of GP spike.



Source: Author.

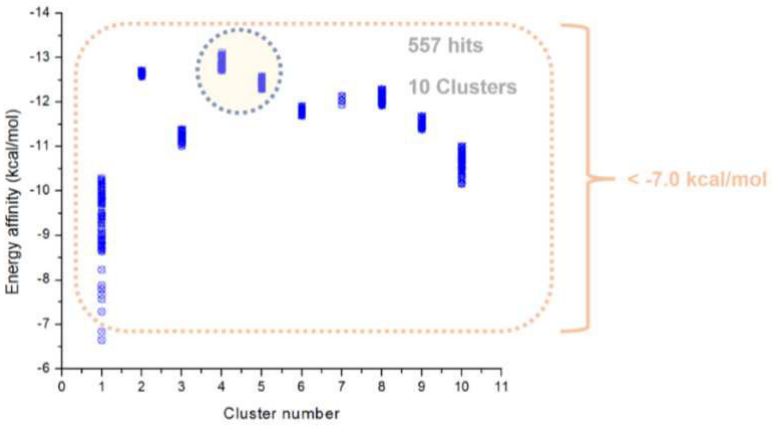
4.3.2. Production of Hits

After the new drug design method, using MolAICal artificial intelligence software, 557 hits (the active substance in the system) were obtained, from the initial phenol derivative growth structure (derivative 4), in Figure 20.

Up to 10 clusters were processed, all of which obtained structures with affinity

energy variation ranging from -13.0 to -7.0 kcal/mol, which obtained more favorable energies. As show in Figure 21.

Figure 21 – Graph of clusters by binding affinity.



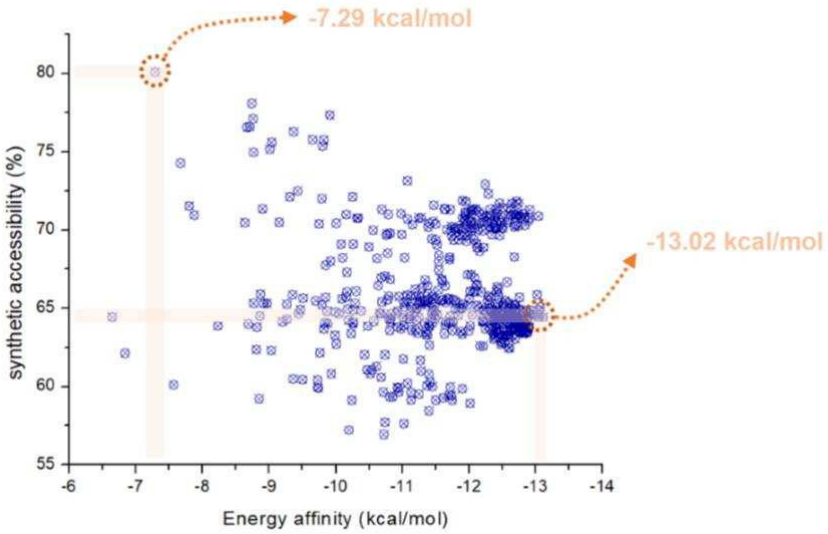
Source: Author.

4.3.3. Synthetic accessibility

In the case of a molecule projected by *De novo* design, the experimental validation of its activity requires the synthesis of the compound. An approach to estimate the ease of synthesis of a ligand is called synthetic accessibility (SA), which is used to generate drug-like molecules and is necessary for many areas in the drug discovery process (JAIN; AGRAWAL, 2004; WANG *et al.*, 2022). The evaluation of the SA of a lead candidate is a task that plays a role in the discovery of the lead, regardless of the method by which the lead candidate is identified (SCOTTI *et al.*, 2012). The more complex the synthesis of the leading candidate, the more time and resources are needed to explore this specific area of the chemical space.

When chemical structures are built during the *De novo* drug design process, it cannot be taken for granted that such compounds' chemical synthesis is feasible. The synthetic accessibility pattern of the study's hits presented a very characteristic behavior of literature, where when the best fits protein, the more complex the synthesis becomes (ERTL; SCHUFFENHAUER, 2009). However, it is possible to get hits with ease of 75-80%, with an energy affinity of up to –7.29 to -13.02 kcal/mol, as shown in Figure 22.

Figure 22 – Synthetic accessibility graph by binding affinity.



Source: Author.

From the obtained results, the best molecules, accordingly to binding affinity (extreme right of the graph) and synthetic accessibility (top of the graph) were selected under the threshold of binding affinity of -7.0 kcal/mol or bellow. Therefore, the Hit 48 (highest synthetic accessibility, near 80% and -7.29 kcal/mol) and Hit 250 (64% and -13.02 kcal/mol).

Curiously, Hit 250 does not have a stereocenter, an important factor considered by Ambit-AS, the algorithm used by MolAICal to evaluate the synthetic accessibility, whilst Hit 48 has only one stereocenter, in the carbon bonded to iodine. It can be argued that the difference between then is more related to the size of the molecule, characteristic relatively higher in Hit 250 than in Hit 48. Besides that, can be supposed, based in the graph of the Figure 10, that the difference of stereoisomers does not have significant influence in the binding energy.

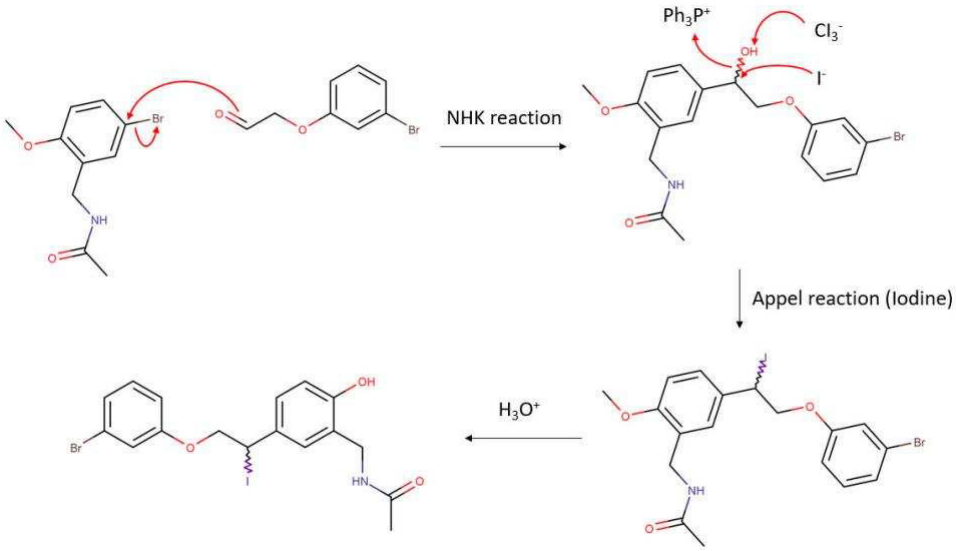
The obtained results are consistent with the literature, because it is expected that for a molecule that better fit in the target site of a protein, more specific it is (ERTL; SCHUFFENHAUER, 2009). However, its relatively easy obtain Hits between 67.5% and 72.5% with binding affinity between -11.5 kcal/mol and -13.0 kcal/mol, as can be seen in the clusters of Hits in Figure 22.

4.3.3.1 Evaluation of synthetic accessibility

After obtained the synthetic accessibility, then a mechanistic proposal synthetic route through softwares or webservers of retrosynthesis, like Manifold. Between the many

possible routes, it is demonstrated in Figure 23, one of the possible routes for Hit 48, starting from 2 reagents readily available in big international chemical suppliers or in the search engine database of those suppliers. The reaction occurs in 3 steps, starting from N-[(5-bromo-2-methoxyphenyl)methyl]acetamide (PubChem CID: 49172932) and 2-(3-bromophenyl)acetaldehyde (PubChem CID: 20387554), that reacts in a Nozaki-Hiyama-Kishi reaction mechanism (reaction NHK), in which a nickel and chromium catalyst, from a aldehyde and a phenyl halide forms an alcohol as a product. In sequence, an Appel reaction for the insertion of iodine, using triphenylphosphine (PubChem CID: 11776) and carbon tetraiodide (PubChem CID: 10487), that converts the alcohol to a halide. In final, converts the methoxide in hydroxide utilizing and strong acid (for instance, H_2SO_4).

Figure 23 – Mechanistic proposal synthetic route by Manifold for Hit 48.



Source: Author.

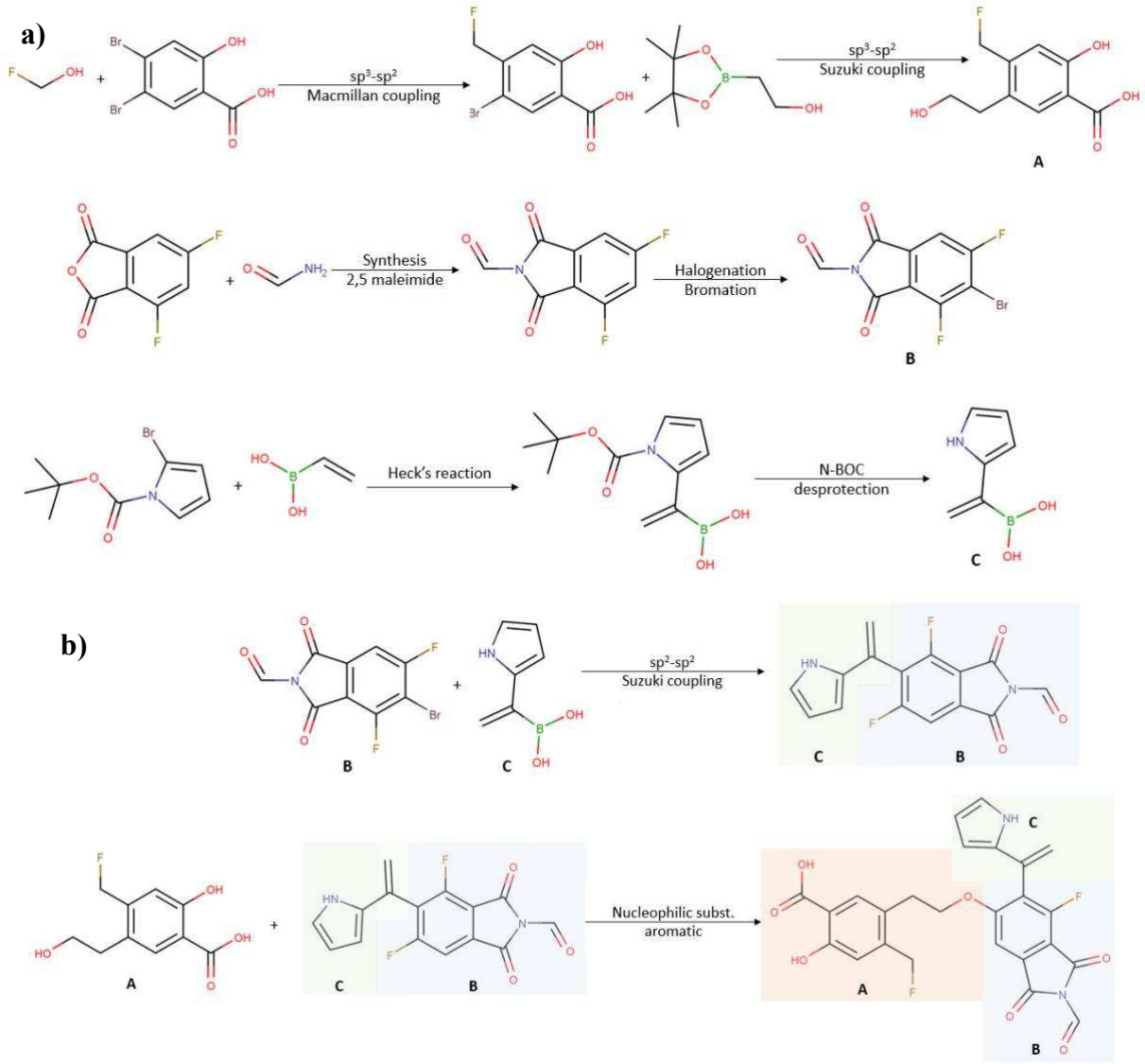
Despite the result from MolAICal/Ambit-SA, the reaction suggested by Manifold shows an economic challenge even for the Hit 48, what requires an route expansion from more common and cheaper reagents, in other words, the synthesis of the starting molecules of the proposed route by common reagents of organic laboratories and industry, to make the production feasible to attend to a large scale demand. Nonetheless, this possibility incurs in more time consuming steps and high probability of less yield, for optimization, scalable pilot plants and industry simulation should be made.

In contrast, the Hit 250 is in line with the synthesis accessibility factor, with routes varying between 8 and 10 steps, with little divergence between starting reagents. An example

of the synthetic routes of Hit 250 is show in Figure 24. In this route, the intermediate molecules are called A (light orange), B (light blue) and C (light green), and in their reactions, those fragments were highlighted for better visualization.

It is noted an excess of coupling steps that needs metallic catalysts, in these specific reaction, palladium catalysts (Suzuki coupling and Heck's reaction), which is known for, among many interesting uses in chemistry, as an expensive metal, that would further raise costs of this synthesis. Besides that, the precursors of A and B may poison the catalyst, lowering its efficiency, in an already extensive process, therefore with a low final yield and life cost (ERHARDT *et al.*, 2008).

Figure 24 – Mechanistic proposal synthetic route by Manifold for Hit 250. In a) the precursors, in b) the production of Hit 250.



Source: Author.

4.3.4. Affinity energy

The 100 closest best results in affinity energy and RMSD were selected from the analogs produced, triggering the ligands to be grouped into clusters of similarity. The results of the generation experiments showed that the molecules obtained later had a better-fit score. The consequences for spike glycoprotein demonstrate this trend well because the clusters generated were formed in proportion to time. Thus, this is possible because larger values generate more diverse molecules, but the convergence of mooring scores becomes poorer. According to the methodology, cluster 1 initially presented two structures that did not interact very well with the protein, giving energies up to > -7.0 kcal/mol. From Cluster 2 up to Cluster 10, there was an increase in affinity in the results, presenting energies below -9.5 kcal/mol, where the most stable structure was Hit 250, with a binding energy of -13.02 kcal/mol, representative of Cluster 4, taking into account, that the designs presented greater structural complexity, which decreased their SA. Cluster 1, in particular, presented a model that best adapted to synthetic accessibility 80%, in this case, the Hit 48, with an affinity energy of -7.29 kcal/mol. These results are presented in Figure 10.

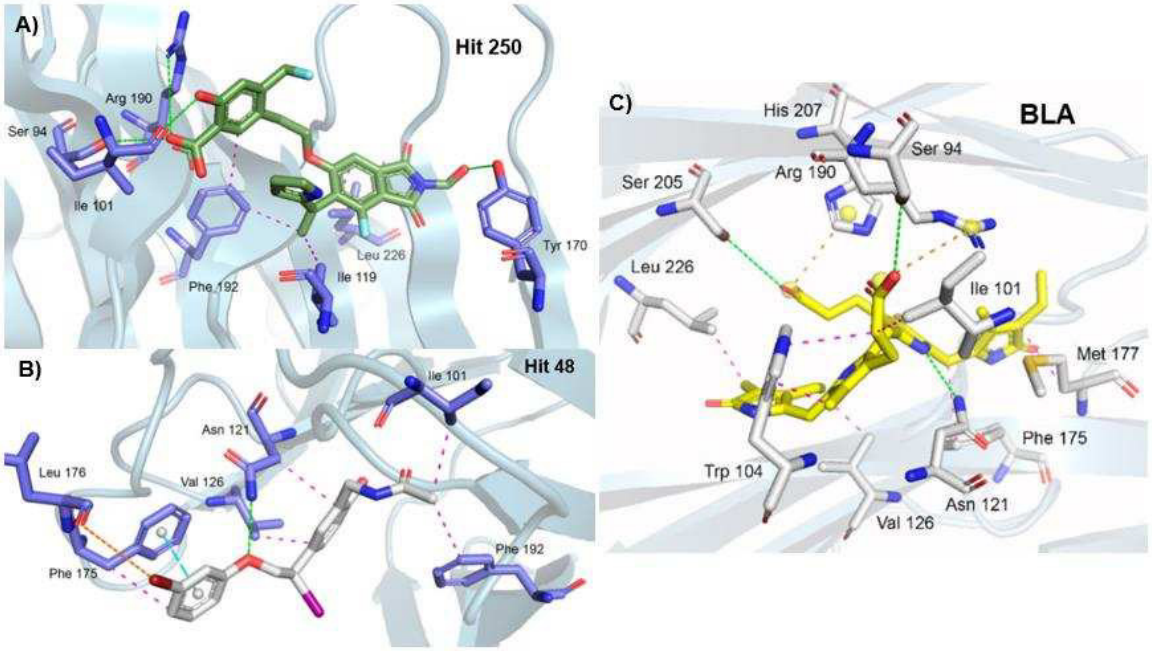
4.3.5. Interaction with protein residues

In a series of docking simulations performed by Singh *et al.* (SINGH; BHARDWAJ; PUROHIT, 2022; SINGH; PUROHIT, 2023a, 2023b), it is possible to observe the strong influence of compounds consisting of at least two rigid rings that have ether (R-O-R) and carbonyl (R-C=O) functional oxygenated groups, whether ester or ketone, on the selective modulation of Spike glycoprotein of SARS -CoV-2.

Hit 250 provided in molecular docking an RMSD of 1.3 Å, with an affinity energy of -13.02 kcal/mol, interacting in the same region as the native BLA linker, as shown in Figure 25a. With strong interactions, it presented four hydrogen bonds in the residues Ser 94 (2.74 Å), Ile 101 (2.10 Å), Tyr 170 (2.28 Å), and Arg 190 (3.04 Å), with a strong contribution from its oxygenated H-bond donor groups, accompanied by hydrophobic interactions pi-alkyl and alkyl with residues Ile 119 (3.77 Å), Phe 192 (3.94/3.78 Å), and Leu 226 (3.79 Å), with a strong contribution from its benzene and the heterocyclic rigid rings (Figure 25a). The luminant represented by Hit 48, in Figure 25b, presented the polar interaction represented by a strong hydrogen bond in the residue Asn 121 (2.51Å), where the ether group (R-O-R) is the

nucleophilic acceptor. In addition, a pi-Stacking interaction was observed with the residue Phe 175 (3.71 Å) and a halogen bond between the leu 176 residue (3.96 Å) and the bromine atom of the ligand. The coronaviral spike is the dominant viral antigen and the target of neutralizing antibodies. Finally, the linker used as a reference standard, the BLA, in the redocking study showed strongly three hydrogen bonds with the residues Ser 94 (2.54 Å), Asn 121 (2.29 Å), and Ser 205 (2.56 Å). With hydrophobic interactions in the residues Ile 101, Trp 104, Val 126, Phe 175, Met 177, and Leu 226, shown in Figure 25c. Previous studies have shown that substitutions of Spike residues closely involved in ligand binding as His 207, Arg 190, and Asn 121, have Influenced in inhibition mechanism of the protein (KIM *et al.*, 2021; KUMAR *et al.*, 2020; ROSA *et al.*, 2021; WAGENER *et al.*, 2020). Additional data are presented in more detail in Table 7.

Figure 25 – 3D interactions of Hit 250 (a), Hit 48 (b) and BLA (c).



Source: Author.

Table 7 – Interactions between the residues of GP spike and BLA, Hit 250, Hit 48 and Der 4

GP spike residues	Molecular Docking results			
	Ligand distance (Å)			
	BLA	Hit 250	Hit 48	Der 4
Ser 94	2,54 (HB)	2,74 (HB)	-	2,69 (HB)
Ile 101	3,84 (HI)	2,10 (HB)	3,54 (HI)	
Gly 103	-	-	-	2,92 (HB)
Trp 104	3,87 (HI) 2,85 (HI)	-	-	3,62 (HI)
Ile 119	-	3,77 (HI)	-	3,58 (HI)
Asn 121	2,29 (HB)	-	3,65 (HI) 2,51 (HB)	2,16 (HB)
Val 126	3,27 (HI)	-	3,41 (HI)	3,78 (HI)
Ile 128	-	-	-	3,88 (HI)
Tyr 170	-	2,28 (HB)	-	-
Phe 175	3,79 (HI)	-	3,71 (PS)	-
Leu 176	-	-	3,96 (HaB)	-
Met 177	3,78 (HI)	-	-	-
Arg 190	5,05 (SB)	3,04 (HB)	-	2,79 (HB)
Phe 192	-	3,94 (HI) 3,78 (HI)	3,77 (HI)	3,73 (HI)
Ile 203	-	-	-	3,84 (HI)
Ser 205	2,56 (HB)	-	-	-
His 207	4,36 (SB)	-	-	2,73 (HB)
Leu 226	3,30 (HI)	3,79 (HI)	-	3,63 (HI)

Source: Author.
Note: HI (hydrophobic interactions: alkyl e π-alkyl) HB (hydrogen bond), SB (saline bridge), HaB (halogen bond) and PS (π stacking)

4.3.6. MM/GBSA calculations

After balancing the production dynamics, the sampling of the steps was performed from 5 to 5, following the sampling interval of 10ns of the methodology for estimating the free energy variation using multiple trajectories. MM/GBSA calculations were performed in an implicit solvent field simulating a 0.15M saline solution.

Although formally, the calculation of the free energy variation in this technique goes through the analysis of entropy from the normal modes of the system equations **(7)**, **(8)**, and **(9)**.

$$\Delta G_{interaction}^{(vac)} = \left(E_{complex}^{MM} - E_{target}^{MM}\right) - \left(E_{complex}^{MM} - E_{ligand}^{MM}\right) + T\Delta S_{NORMODS} \quad \textbf{(7)}$$

$$\Delta G_{interaction}^{(vac)} = E_{complex}^{MM} - E_{target}^{MM} - E_{complex}^{MM} + E_{ligand}^{MM} + T\Delta S_{NORMODS} \quad \textbf{(8)}$$

$$\Delta G_{interaction}^{(vac)} = E_{ligand}^{MM} - E_{target}^{MM} + T\Delta S_{NORMODS} \quad \textbf{(9)}$$

Calculations of normal modes are pretty time-consuming and computationally costly. This type ultimately makes virtual screening calculations, which are the focus of this study, not entirely impossible. However, there is a more important reason. It has been shown that entropy calculations decrease the correlation of predicted affinity values with experimental values when the analysis is done with a few microstates sampled from the trajectories (HOU *et al.*, 2011; RASTELLI *et al.*, 2009). Because calculation time limitations are essential, including these calculations in the procedure is not encouraged.

MM/GBSA energies are considered a way to estimate free energy for *in silico* study of ligands in protein complexes (GENHEDEN; RYDE, 2015). They are typically based on MD simulations and bring accuracy between empirical punctuation and strict alchemical disturbance (CHEN *et al.*, 2018). As in conformational entropy, it is tough to obtain a concurrent value. Mainly, if ligands do not have any binding-induced structural changes in MD simulations, conformational entropy is generally ignored to calculate by standard mode analysis (WANG *et al.*, 2017). MolAICal, therefore, provided a quick way to evaluate bonding free energy without ligand entropy based on the three-trajectory approach. Where, once again, the native ligand BLA/S-glycoprotein complex proved to be the best result, which continued to be the most stable in the study system, based on its free energy, with -28.79 kcal/mol concerning the other ligand under study, Hit 48, which presented a free energy -20.91 kcal/mol, and Hit 250, with free energy estimative -15.50 kcal/mol. Therefore, the interaction energy decomposition technique revealed the contribution of the ligand-receptor complex and its final energy in Table 8.

Table 8 – Predicted free energy of BLA, Hit 48 and Hit 250 with GP spike

Complex	**ΔE (electrostatic) + ΔG (sol)**	**ΔE (VDW)**	**ΔG binding (kcal.mol^{-1})**	**Standard Deviation**
BLA/ GP spike	24,01	-52,80	-28,79	± 0,0262
Hit 48/ GP spike	15,51	-36,42	-20,91	± 0,0180
Hit 250/ GP spike	31,20	-45,69	-14,50	± 0,0260

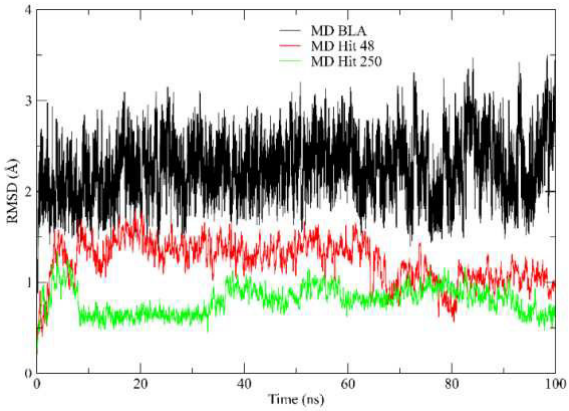
Source: Author.

4.3.7. Molecular dynamics

4.3.7.1. RMSD analysis

After the analysis of the energy values, other important parameters to investigate the quality of the molecular dynamics are Root Mean Square Deviation (RMSD) of protein (backbone) to Hit 250, Hit 48, and BLA. The RMSD values obtained by the protein backbone along the MDs show all values between 0.76 Å and 2.10 Å. In the MD with the Hit 250 ligand, the profile is closer to 0.8 Å. Soon after the MD of Hit 48 demonstrated similar behavior to the previous one, but only reached a more stable configuration when it arrived at 82 ns, with an average RMSD of 1.0 Å. Already finalizing, the RMSD values of MD with the reference linker BLA showed a situation of suitability to the most favorable system during the trajectory of 100ns, with an average value of 1.9 Å, as shown in Figure 26.

Figure 26 –RMSD values of GP spike with BLA (black), Hit 48 (red) and Hit 250 (green).

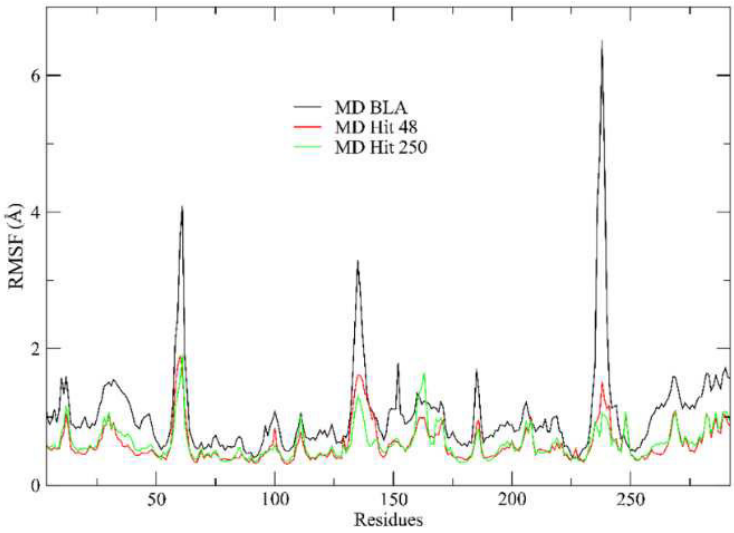


Source: Author.

4.3.7.2. RMSF analysis

The Root Mean Square Fluctuation (RMSF) is a parameter related to the flexibility of individual protein residues, serving to qualitatively assess the progression of molecular dynamics(DONG *et al.*, 2018). Considering Figure 27, a similar profile of RMSF values is observed regardless of the ligand in contact. However, it should be noted that the native leach er has obtained greater fluctuations in the residues Gln 52, Phe 135, and Gly 232, with values above 2.0 Å.

Figure 27 – RMSF values of GP spike with BLA (black), Hit 48 (red) and Hit 250 (green).

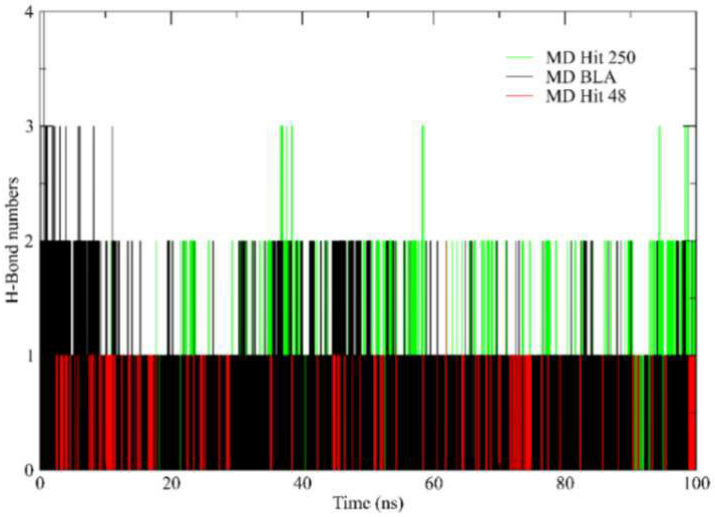


Source: Author.

4.3.7.3. H-bonds

The number of hydrogen bonds (H-Bond) found during MDs, considering the maximum value of 3.3 Å, are shown in Figure 28 and Table 9. About MD with reference ligand, BLA presents up to 4 hydrogen bonds per frame and several frames with three hydrogen bonds. The Hit 250 has only ten frames with three hydrogen bonds and several frames with two hydrogen bonds during molecular dynamics. And finally, MD with Hit 48 presents a smaller number of frames with two hydrogen bonds and several frames with one hydrogen bond. Therefore, it can be inferred that so much of the BLA as Hit 250 tend to interact more with Spike glycoprotein, making it a possible efficient antiviral against this virus. The interaction tendency between glycoprotein and Hit 250 can be confirmed by maintaining the binding site along the MD, evidenced by the hydrogen bonds detected in the three systems.

Figure 28 – *Frames* of the number of hydrogen bonds between GP spike with BLA (black), Hit 48 (red) and Hit 250 (green).



Source: Author.

Table 9 – Residues of the GP spike that showed H-bond along the MDs. In bold, critical role in the replication cycle residues.

System	H-Bond
MD with BLA	Ser 94, Glu 96, Asn 99, Ile 101, **Asn 121**, Tyr 170, Ser 172, Gln 173, Asp 178, **Arg 190**, and **His 207**
MD with Hit 250	Ser 94, Glu 96, Ile 101, **Asn 121**, Gln 173, and **Arg 190**
MD with Hit 48	Gly 103, Ser 205, **His 207**, and Leu 226

Source: Author.

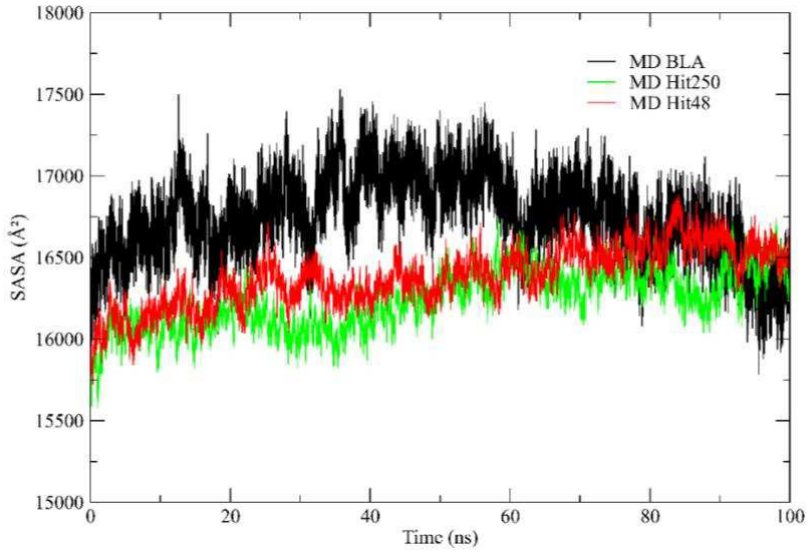
In the residues identified with H-bonds along the molecular dynamics, in Table 9, the recurrence of residues **Asn 121, Arg 190,** and **His 207** is observed, thus demonstrating an interaction potential of both Hit 48 and Hit 250, like the interaction between S-glycoprotein and BLA.

4.3.7.4. Solvent Accessible Surface Area

Solvent Accessible Surface Area (SASA) is defined as the surface area of a protein that interacts with its solvent molecules (MAZOLA *et al.*, 2015). Average SASA values for free BLA, Hit 250, and Hit 48 complexes were monitored during 100 ns MD simulations. The traces for the SASA in Figure 29 show a steep increase within 10 ns indicating structural relaxation. The average SASA values for free BLA, Hit 250, and Hit 48 complexes were found to be 16,611 $Å^2$, 16,249 $Å^2$, and 16,223 $Å^2$, respectively. There was no major change observed in the SASA

values due to ligand binding. After this time, the values fluctuate around a constant value. We thus assume that the simulation times of 100 ns were sufficient for sampling equilibrated systems. The highest SASA is found for the S-glycoprotein molecules with the stabilizing monovalent ions. The run without monovalent ions shows a large fluctuation, whereas the systems with higher ion concentrations have smaller areas and may be shrinking under the influence of the surface charge, yielding more compact protein structures. Further inspection of the data demonstrates that the fluctuation or 'breathing' of the relaxed surface is mainly due to a fluctuation of the SASA of the flexible C-terminal area.

Figure 29 – SASA values of GP spike BLA (black), Hit 48 (red) and Hit 250 (green).



Source: Author.

4.3.8. In silico ADMET study

4.3.8.1. Evaluation of druglikeness

For Wager and coworkers (WAGER *et al.*, 2016b), low lipophilic compounds (logP < 3) that are larger and more polar than commercially available CNS active substances (TPSA > 75 Å²) reside in a physicochemical space where *in vivo* toxicity is unlikely. Compounds with high lipophilicity and low polarity tend to be more toxic than safe, in addition to showing unfavorable pharmacodynamic interactions against biological targets (HUGHES *et al.*, 2008).

In the druglikeness radar of Figure 30, it is possible to notice that the three

compounds, that is, Hit 48 (Figure 30a), Hit 250 (Figure 30b), and the BLA ligand (Figure 30c) move outside an ideal spectrum of mediated lipophilicity by logP, with values greater than 3.0 (Table 10). Compound Hit 48 showed low topological polarity (TPSA = 58.56 Å²) which, when combined with high lipophilicity, classified it as possibly CNS permeant toxicant, according to the Pfizer filter that combines these two attributes (Table 10).

Table 10 – Physicochemical Properties and estimation of druglikeness. The * notation indicates that Pfizer's rule relates logP and TPSA attributes to the physical-chemical space of the ligands: low logP and high TPSA (logP < 3 and TPSA > 75 Å²), low toxic risk; high logP and low TPSA (logP > 3 and TPSA < 75 Å²), toxic risk. Note: in bold, properties favorable to binders.

Property	Hit 48	Hit 250	BLA
Physicochemical Properties			
logP	3.84	3.98	5.46
logD	3.14	1.24	2.08
MW	488.94 g/mol	496.11 g/mol	582.25 g/mol
HBA	4	9	10
HBD	2	3	6
TPSA	58.56 Å²	137.0 Å²	171.63 Å²
nRot	7	9	11
nRing	2	4	4
MaxRing	6	9	5
nHet	6	11	10
fChar	0	0	0
nRig	13	26	29
nStereo	1	0	0
Medicinal Chemistry			
Pfizer rule*	2 Alerts; logP>3 and TPSA<75 Å²; (–)	**1 alert; logP>3; (+)**	**1 alert; logP>3; (+)**
GSK filter	1 Alert; MW>400 g/mol; (–)	1 Alert; MW>400 g/mol; (–)	2 Alerts; logP>4 and MW>400 g/mol; (–)
Golden Triangle	0 alert; (+)	**0 alert; (+)**	1 alert; MW>500 g/mol; (–)
QED	0.468	0.305	0.202
Fsp^3	0.23	0.12	0.21
MCE-18	26.0	**54.0**	**60.0**

Source: Author.

It is curious to note that both Hit 48 and Hit 250 failed the GSK filter druglikeness criteria for having MW > 400 g/mol, an indication that these substances may have limited pharmacokinetics, such attributes include solubility, absorption and stability metabolism (GLEESON, 2008). However, it was possible to observe that Hit 250 passed the safety and pharmacodynamic criteria of the Pfizer rule, as it occupies a physical-chemical space where compounds with TPSA reside within the ideality spectrum (Figure 30b).

4.3.8.2. Evaluation of MCE-18

In recent years, the molecules claimed for patents follow a physicochemical trend that deviates from the medicinal chemistry spectrum of the commonly used "rule of five". This chemical singularity focuses on how the fraction of sp^3 hydrous carbons is distributed in aliphatic structures, chiral centers, and aromatic and non-aromatic cyclic structures. In this perspective, molecules registered in patents have been shown to be slightly more lipophilic and more polar than commercially available therapeutics (IVANENKOV; ZAGRIBELNYY; ALADINSKIY, 2019; WAGER *et al.*, 2010).

In this test, it was possible to observe that QED values lower than 0.5 (on a scale ranging from 0.0 to 1.0) are directly related to the large molecular size of the ligands (MW > 400 g/mol), and are reduced as that the TPSA increases to 171.63 Å² (BLA), depending on the number of HBA atoms (Table 10). However, the structural complexity involving the Hit 250 and BLA ligands, especially due to the total of 4 aromatic (or heteroaromatic) rings, including the total of 9 atoms in the 2,3-dihydro-1H-isoindole-1,3-dione of the BLA complexed ligand, which yielded the ligands an MCE-18 score of 54.0 and 60.0, respectively. This finding suggests that the ligands present an excellent degree of similarity with the structural complexity of the compounds registered in patents in recent years (Table 10).

4.3.8.3. Predicted pharmacokinetic descriptors

The oral bioavailability of a drug concerns the alignment between the pharmacological portion absorbed as a function of a low rate of hepatic clearance. Pharmacological databases, such as Pfizer, Inc., estimate that a ligand exhibits high passive permeability when its *in vitro* P_{app} MDCK value is greater than $10x10^{-6}$ cm/s, which results in high oral bioavailability as it the clearance rate decreases (VAN DE WATERBEEMD; GIFFORD, 2003; WAGER *et al.*, 2010). For Johnson *et al.* (HUGHES *et al.*, 2008), these descriptors are closely related to the buffer lipophilicity (logD) at pH 7.4, limited to small compounds that are not very lipophilic, that is, that occupy a physical-chemical space formed by $-2 < \text{logD at pH } 7.4 \leq 5$ and $200 < \text{MW} \leq 500$ g/mol.

In the graph in Figure 30d, it is possible to observe that the three ligands are outside the ideality spectrum for good intestinal permeability. This empirical decision corroborates the estimated P_{app} MDCK descriptors, where values equal to and less than $4.7x10^{-8}$ cm/s suggest a

low passive permeability (Table 11). However, the substances showed low susceptibility to being Pgp substrates, as an indication of good intestinal absorption, with HIA values > 90% for compounds Hit 48 and Hit 250 (Table 11).

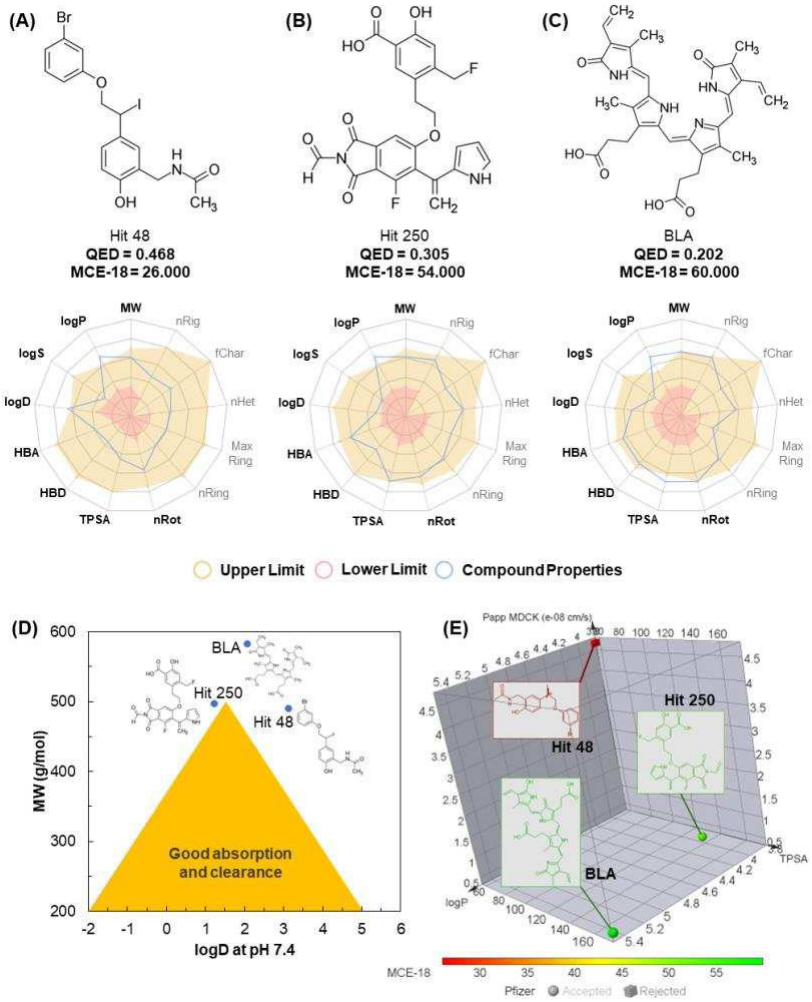
In addition, it is possible to note the contribution of the high lipophilicity in the distribution of the compounds in the blood plasma and in the CNS. Compounds of greater lipophilicity can bind strongly with serum proteins and have their tissue distribution affected (DYABINA *et al.*, 2016; PIRES; KAMINSKAS; ASCHER, 2018). In this study, it was possible to observe that the compounds presented PPB < 90%, which allows a considerable distribution in biological tissues. At the same time, the low polarity of Hit 48 makes it more susceptible to distribution in the CNS, corroborating the permeability coefficient in the BBB in the order of 2.969, which represents a ratio of the concentration of the compound in the brain by its distribution in the blood ($C_{[Brain]}/C_{[Blood]}$) (Table 11).

Table 11 – Pharmacokinetic descriptors of the ligands.

Property	**Hit 48**	**Hit 205**	**BLA**
P_{app} MDCK	$4{,}70.E^{-8}$ cm.s^{-1}	$4{,}35.E^{-9}$ cm.s^{-1}	$4{,}34.E^{-9}$ cm.s^{-1}
Pgp-sub	(–)	(–)	(–)
HIA	95,92%	91,51%	86,78%
VD	0,76 L.kg^{-1}	0,68 L.kg^{-1}	0,41 L.kg^{-1}
PPB	89,67%	86,55%	86,43%
BBB ($C_{[Brain]}/C_{[Blood]}$)	2,969	0,029	0,098
H-HT	(–) 0,72	(–) 0,66	(–) 0,75
Mutagen	(–) 0,69	(–) 0,71	(–) 0,71
LC_{50} *Pimephales p.*	0,02 mM	0,32 mM	1,86 mM

Source: Author.

Figure 30 – Relationship between structure and druglikeness (QSAR) of Hit 48 (a), Hit 250 (b) and BLA (c), prediction of balance between absorption and clearance (d) and pharmacokinetic physical-chemical space (e).



Source: Author.

4.3.8.4. Metabolism and oral acute toxicity

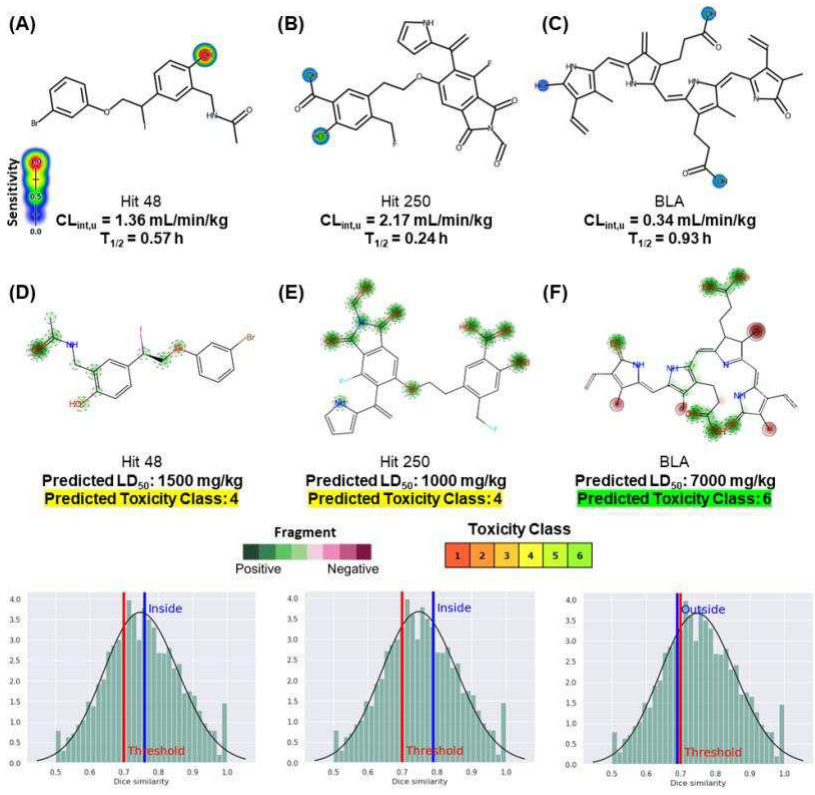
Predicting the sites of metabolism allows us to estimate the effects of drug biotransformation on hepatic clearance and adverse effects on the human liver. Empirical analysis suggests that compounds with MW around 500 g/mol are metabolically unstable that is, they have structural fragments susceptible to biotransformation, forming secondary metabolites that are more water-soluble and more favorable to excretion. However, some biotransformations can form chemically reactive intermediates, such as epoxidation mediated by aromatic hydroxylation(HUGHES; MILLER; SWAMIDASS, 2015; JOHNSON; DRESS; EDWARDS, 2009).

In this predictive test, the fragments are identified from a data library that relates the degree of sensitivity of the functional groups and structural fragments to be biotransformed in the human liver microsome system with the degree of specificity of these in the molecular structure (ZHENG *et al.*, 2009). Here, it was possible to observe, mainly, that the aromatic centers of the ligands do not pose a risk of hydroxylation, reducing the risk of these substances forming reactive secondary metabolites (Figure 31a-c), which implies a low risk of human hepatotoxicity and mutagenicity (Table 6). Hit 48 has a phase II metabolism site in its phenolic hydroxyl, sensitive to conjugation reactions via UGT (UDP-glucuronosyltransferase), indicating that the substance is more resistant to phase I metabolism, with an order of $CL_{int,u}$ estimated at 1.36 mL/min/kg which may be indicative of good oral bioavailability (Figure 16a). However, this metabolism pathway seeks to optimize the excretion pathway.

A low rate of hepatic clearance implies a longer half-life ($T_{1/2}$) for pharmacological action (VAN DE WATERBEEMD; GIFFORD, 2003). This is observed when comparing the metabolism pathways of Hit 48 (Figure 31a) and Hit 250 (Figure 31b), where the higher incidence of metabolism sites induces a shorter $T_{1/2}$ time to Hit 250, estimated at 0.24 h, depending on its highest clearance order compared to Hit 48, with an estimated $CL_{int,u}$ value of 2.17 mL/min/kg (Figure 31b). In the probability maps of Figures 31d-f, it is possible to observe that the metabolism sites are within the positive contributions that reduce the acute toxicity of the Hit 48 ligands (Figure 31d) and Hit 250 (Figure 31e), where the predicted LD_{50} values of 1500 mg/kg and 1000 mg/kg indicate that they are compounds of toxicity class 4 (DIAZA *et al.*, 2015), which are compounds that require control of the administered oral dose. These compounds showed an order of similarity greater than 70% (inside the threshold) with the compounds deposited in the PubMed database (BORBA *et al.*, 2022).

In addition, it is worth mentioning that Hit 250 and BLA binders showed the best LC_{50} values for Fathead Minnow, where values of 0.32 mM and 1.86 mM (in logarithmic scale), respectively, suggest that the minimum effective dose does not pose a risk to a fish population tested, as an indication of the safety of oral administration in humans (Table 10).

Figure 31 – Metabolism site prediction of Hit 48 (a), Hit 250 (b) and BLA (c) and fragment-based acute toxicity prediction of Hit 48 (d) Hit 250 (e) and BLA (f).



Source: Author.

One of the limitations of *De Novo* drug design has been that they cannot identify a perfect compound for synthesis since some of the potential hits generated have complexity that compromises the realization of their synthesis. However, in return, they can identify high-quality ideas for future *in vitro* and *in vivo* assays. Considering the imperfections of automated chemical synthesis planning and reaction pathway design, combining AI-driven generative molecular design models with advanced synthesis and retrosynthesis algorithms could offer ample future opportunities for new molecular discoveries.

4.4. Conclusion

It was carried out through the innovative computer-aided drug design *De novo*, researching new drug candidates to treat Sars-Cov-2, more precisely, S-glycoprotein as a target. Therefore, an important role was played in developing new anti-Covid-19 drugs, which was of significant importance because, amid the limitation of resources, it accelerated the drug development process, reducing the time and additional costs of traditional screening.

Some new derivatives of phenols and chromones were designed and studied through molecular docking, where 557 hits were generated. The mode of binding of the proposed compounds with the target protein was evaluated, and the data from docking studies explained that some newly designed analogs had a significantly high affinity for the target protein compared to BLA as a reference linker.

The compound with the highest affinity value was Hit 250, which proved to be the most potent inhibitor in this *in silico* study series with a binding energy of -13.02 kcal/mol. Still, its synthetic viability was close to 50%, besides showing lower stability in the molecular dynamics analysis studies of RMSD, RMSF, and SASA. While another drew attention was Hit 48, with -7.29 kcal/mol of affinity energy, presenting better synthetic viability, close to 80%, and better stability in the study of molecular dynamics, compared to the reference drug BLA, with the binding affinity of -8.1 kcal/mol. In the ADMET tests, Hit 250 showed greater similarity with species registered in patents and stands out concerning Hit 48, for occupying a physical-chemical space with a low toxic incidence in vivo, due to its high polarity. Therefore, it is suggested that these compounds can be used in clinical trials to test their effectiveness for social benefits as a standard for future projects, optimization, and research in producing more effective analogs.

CHAPTER V

5. CONCLUSION

Computational chemistry is one of the first and major steps in drug development, the last decade it has increased substantially because of the technology, softwares, servers, and novel approaches, to increase the chances of success and provide better starting point in preclinical trials. Thus, this worked showed that the molecules extracted from *Daldinia sp.* has interesting bioactivities that should be explored further as it is, to date, understudied despite almost global distribution and with *in silico* and *in vitro* desirable properties. Therefore, it is an economic, sustainable and viable source of high valor bioactive molecules.

In the first work, the chromone and phenol derivatives from *Daldinia*, as potential drugs capable of inhibiting CHIKV, are safe for oral use with the exception of chromones. From the phenol derivatives, the Der 8 is highlighted as the best candidate, as it is able to inhibit ZIKV and readily available (Tyrosol).

In the second work, the Hits 48 and 250 showed great potential to treat SARS-CoV-2, despite the difficulty of synthesis of both, and good oral safety. These molecules may also be potential candidates for inhibition of future SARS viruses.

Therefore, this work is expected to provide a fresh perspective on Daldinia natural products and make significant contributions to the literature on drug discovery and in silico design, as presented in Appendix F. The results demonstrate the advantages of using the computational method in the first phase of drug development. As a next step, future research should focus on prospecting bioactive natural products from Daldinia, optimizing the extraction, purification, and production of derivatives, conducting organic synthesis of Hit 48 and Hit 250, and performing preclinical tests of CHIKV and SARS-CoV-2 (including its variants).

REFERENCES

ABRAHAM, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, [*s.l*], v. 1–2, p. 19–25, 2015.

AHMADI, M. *et al.* Chem3D 15.0 user guide. Macromolecules, [*s.l*], v. 24, n. 2, p. 1–61, 2005.

ALENCAR, W. L. M. *et al.* Interactions of Co, Cu, and non-metal phthalocyanines with external structures of SARS-CoV-2 using docking and molecular dynamics. **Scientific Reports**, [*s.l*], v. 12, n. 1, p. 1–20, 2022.

ALLOUCHE, A. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. **Journal of computational chemistry**, [*s.l*], v. 32, p. 174–182, 2012.

ARAGÃO, F. F. DE; TOBIAS, A. F. Pharmacological treatment of pain in pregnancy. **Brazilian Journal of Pain**, [*s.l*], v. 2, n. 4, p. 374–380, 2019.

ARSHIA, A. H. *et al.* De Novo design of novel protease inhibitor candidates in the treatment of SARS-CoV-2 using deep learning, docking, and molecular dynamic simulations. **Comput Biol Med**. [*s.l*], v. 139, 2021.

BAELL, J.; WALTERS, M. A. Chemistry: Chemical con artists foil drug discovery. **Nature**, [*s.l*], v. 513, n. 7519, p. 481–483, 2014.

BAI, Q. *et al.* MolAICal: A soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. **Briefings in Bioinformatics**, [*s.l*], v. 22, n. 3, p. 1–12, 2021.

BAI, Q. *et al.* Application advances of deep learning methods for de novo drug design and molecular dynamics simulation. **Wiley Interdisciplinary Reviews**: **Computational Molecular Science**, [*s.l*], v. 12, n. 3, p. 1–19, 2022.

BANERJEE, P. *et al.* ProTox-II: A webserver for the prediction of toxicity of chemicals. **Nucleic Acids Research**, [*s.l*], v. 46, n. W1, p. W257–W263, 2018.

BAR-ON, Y. M. *et al.* SARS-CoV-2 (COVID-19) by the numbers. **eLife**, [*s.l*], v. 9, 2020.

BARTHOLOMEEUSEN, K. *et al.* Chikungunya fever. **Nature Reviews Disease Primers**, [*s.l*], v. 9, n. 1, p. 1–21, 2023.

BASSETTO, M. *et al.* Computer-aided identification, design and synthesis of a novel series of compounds with selective antiviral activity against chikungunya virus. **Antiviral Research**, [*s.l*], v. 98, n. 1, p. 12–18, 2013.

BATTISTI, V.; URBAN, E.; LANGER, T. Antivirals against the chikungunya virus. **Viruses**, [*s.l*], v. 13, n. 7, 2021.

BICKERTON, G. R. *et al.* Quantifying the chemical beauty of drugs. **Nature Chemistry**, [*s.l*], v. 4, n. 2, p. 90–98, 2012.

BIOVIA, D.S. **Discovery Studio Modeling Environment**. Versão 4.5. San Diego: Dassault Systèmes. 2015.

BLAND, J. M.; ALTMAN, D. G. **Tukey Multiple Comparison test. British Medical Journal**, [*s.l*], v. 310, 1995.

BORBA, J. V. B. *et al.* STopTox: An in Silico Alternative to Animal Testing for Acute Systemic and Topical Toxicity. **Environmental Health Perspectives**, [*s.l*], v. 130, n. 2, p. 1–12, 2022.

BORN, J. *et al.* PaccmannRL: Designing Anticancer Drugs From Transcriptomic Data via Reinforcement Learning. **Lecture Notes in Computer Science** (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), [*s.l*], v. 12074 LNBI, p. 231–233, 2020.

BORN, J. *et al.* Data-driven molecular design for discovery and synthesis of novel ligands: A case study on SARS-CoV-2. **Machine Learning: Science and Technology**, [*s.l*], v. 2, n. 2, p. 0–15, 2021a.

BORN, J. *et al.* PaccMannRL: De novo generation of hit-like anticancer molecules from transcriptomic data via reinforcement learning. **iScience**, [*s.l*], v. 24, n. 4, 2021b.

BORRE, F. *et al.* Impact of the COVID-19 Pandemic on Infectious Diseases in Brazil: A Case Study on Dengue Infections. **Epidemiologia**, [*s.l*], v. 3, n. 1, p. 97–115, 2022.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária. **Portaria nº 344**, de 12 de maio de 1998. Aprova o Regulamento Técnico sobre substâncias e medicamentos sujeitos a controle especial. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF, 15 mai. 1998. Disponível em: https://antigo.anvisa.gov.br/documents/10181/2718376/(31)PRT_SVS_344_1998_COMP.pdf/0075d46b-4214-4363-a190-0ac168c140a0. Acesso em 17 set. 2022.

BRITO, S. B. P. *et al.* Pandemia da COVID-19: o maior desafio do século XXI TT - COVID-19 pandemic: the biggest challenge for the 21st century. **Vigil. sanit. debate**, [*s.l*], v. 8, n. 2, p. 54–63, 2020.

CARVALHO, D. O. *et al.* Suppression of a field population of Aedes aegypti in Brazil by sustained release of transgenic male mosquitoes. **PLoS Neglected Tropical Diseases**, [*s.l*], v. 9, n. 7, p. 1–15, 2015.

CAVALCANTI, L. P. DE G. *et al.* Competência de peixes como predadores de larvas de Aedes aegypti, em condições de laboratório. **Revista de Saúde Pública**, [*s.l*], v. 41, n. 4, p. 638–644, 2007.

CDC. Centers for Disease Control and Prevention. Introduction to Epidemiology. *In:* CDC. **Principles of Epidemiology in Public Health Practice**. Atlanta: Departamento of Health & Human Services. 3rd ed. 2012, p.1-93. Disponível em: https://archive.cdc.gov/www_cdc_gov/csels/dsepd/ss1978/SS1978.pdf. Acesso em 14 ago. 2023.

CDC. **COVID-19 Treatments and Medications**. Georgia, 2023. Disponível em: <https://www.cdc.gov/coronavirus/2019-ncov/your-health/treatments-for-severe-illness.html>.

CHAKRAVARTI, R. *et al.* A review on potential of natural products in the management of COVID-19. **RSC Advances**, [*s.l*], v. 11, n. 27, p. 16711–16735, 2021.

CHEMAXON. Marvin 22.22, 2022. Disponível em: <http://www.chemaxon.com>
CHEN, F. *et al.* Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. **Rna**, [*s.l*], v. 24, n. 9, p. 1183–1194, 2018.

CHENG, V. C. C. *et al.* Viral Infections, an Overview with a Focus on Prevention of Transmission. **International Encyclopedia of Public Health**, [*s.l*], v.7, n.2, p. 368–377, out. 2016.

CHERRY, J. D. The chronology of the 2002-2003 SARS mini pandemic. **Paediatric Respiratory Reviews**, [*s.l*], v. 5, n. 4, p. 262–269, 2004.

CHERRY, J. D.; KROGSTAD, P. SARS: The first pandemic of the 21st century. **Pediatric Research**, [*s.l*], v. 56, n. 1, p. 1–5, 2004.

CONTRERAS, S.; IFTEKHAR, E. N.; PRIESEMANN, V. From emergency response to long-term management: the many faces of the endemic state of COVID-19. **The Lancet Regional Health - Europe**, [*s.l*], v. 30, n. May, p. 100664, 2023.

DA FONSECA, A. M. *et al.* Naphthoquinones biflorin and bis-biflorin (Capraria biflora) as possible inhibitors of the fungus Candida auris polymerase: molecular docking, molecular dynamics, MM/GBSA calculations and in silico drug-likeness study. **Journal of Biomolecular Structure and Dynamics**, [*s.l*], v. 41, n. 21, p. 11564–11577, 2022.

DA ROCHA, M. N. *et al.* Virtual Screening of Citrus Flavonoid Tangeretin: A Promising Pharmacological Tool for the Treatment and Prevention of Zika fever and COVID-19. **Journal of Computational Biophysics and Chemistry**, [*s.l*], v. 20, n. 3, p. 283–304, 2021.

DA SILVA NETO, S. R. *et al.* Arboviral disease record data - Dengue and Chikungunya, Brazil, 2013–2020. **Scientific Data**, [*s.l*], v. 9, n. 1, p. 1–11, 2022.

DAINA, A.; MICHIELIN, O.; ZOETE, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. **Scientific Reports**, [*s.l*], v. 7, n. January, p. 1–13, 2017.

DAS, P. K. *et al.* Design and validation of novel chikungunya virus protease inhibitors. **Antimicrobial Agents and Chemotherapy**, [*s.l*], v. 60, n. 12, p. 7382–7395, 2016.

DE ABREU, F. V. S. *et al.* Haemagogus leucocelaenus and Haemagogus janthinomys are the primary vectors in the major yellow fever outbreak in Brazil, 2016–2018. **Emerging Microbes and Infections**, [*s.l*], v. 8, n. 1, p. 218–231, 2019.

DIAS, D. A.; URBAN, S.; ROESSNER, U. A Historical overview of natural products in drug discovery. **Metabolites**, [*s.l*], v. 2, n. 2, p. 303–336, 2012.

DIAZA, R. G. *et al.* Comparison of in silico tools for evaluating rat oral acute toxicity. **SAR and QSAR in environmental research**, [*s.l*], v. 26, n. 1, p. 1–27, 2015.

DIEZ, M. *et al.* Insights into mechanism kinematics for protein motion simulation. **BMC Bioinformatics**, [*s.l*], v. 15, n. 1, p. 1–14, 2014.

DONG, Y. W. *et al.* Structural flexibility and protein adaptation to temperature: Molecular dynamics analysis of malate dehydrogenases of marine molluscs. **Proceedings of the National Academy of Sciences of the United States of America**, [*s.l*], v. 115, n. 6, p. 1274–1279, 2018.

DOYTCHINOVA, I. Drug Design—Past, Present, Future. **Molecules**. [*s.l*], v. 27, n. 5, 2022. Disponível em https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8911833/pdf/molecules-27-01496.pdf. Acesso 20 set. 2023.

DU, L.; KING, J. B.; CICHEWICZ, R. H. Chlorinated polyketide obtained from a Daldinia sp. treated with the epigenetic modifier suberoylanilide hydroxamic acid. **Journal of Natural Products**, [*s.l*], v. 77, n. 11, p. 2454–2458, 2014.

DYABINA, A. S. *et al.* Prediction of blood-brain barrier permeability of organic compounds. Doklady. **Biochemistry and biophysics**, [*s.l*], v. 470, n. 1, p. 371–374, set. 2016.

EBERHARDT, J. *et al.* AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. **Journal of Chemical Information and Modeling**, [*s.l*], v. 61, n. 8, p. 3891–3898, 23 ago. 2021.

ERHARDT, S. *et al.* Mechanisms of catalyst poisoning in palladium-catalyzed cyanation of haloarenes. Remarkably facile C-N bond activation in the [(Ph3P) 4Pd]/[Bu4N]+ CN- system. **Journal of the American Chemical Society**, [*s.l*], v. 130, n. 14, p. 4828–4845, 2008.

ERTL, P.; SCHUFFENHAUER, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. **Journal of Cheminformatics**, [*s.l*], v. 1, n. 1, p. 8, 2009.

EVANS, B. R. *et al.* Transgenic Aedes aegypti Mosquitoes Transfer Genes into a Natural Population. **Scientific Reports**, [*s.l*], v. 9, n. 1, p. 1–6, 2019.

FARAGO, O. Langevin thermostat for robust configurational and kinetic sampling. **Physica A: Statistical Mechanics and its Applications**, [*s.l*], v. 534, p. 122210, 15 nov. 2019.

FERREIRA, L. L. G.; DE MORAES, J.; ANDRICOPULO, A. D. Approaches to advance drug discovery for neglected tropical diseases. **Drug Discovery Today**, [*s.l*], v. 27, n. 8, p. 2278–2287, 2022.

FERREIRA, R. M. *et al.* Outcomes associated with Hydroxychloroquine and Ivermectin in hospitalized patients with COVID-19: A single-center experience. **Revista da Associação Medica Brasileira**, [*s.l*], v. 67, n. 10, p. 1466–1471, 2021.

FIGUEIREDO, L. T. M. Human urban arboviruses can infect wild animals and jump to sylvatic maintenance cycles in South America. **Frontiers in Cellular and Infection Microbiology**, [*s.l*], v. 9, n. JUL, p. 1–6, 2019.

FLORESTA, G. *et al.* Artificial Intelligence Technologies for COVID-19 De Novo Drug Design. **International Journal of Molecular Sciences**, [*s.l*], v. 23, n. 6, 2022.

FRICKER, P. C.; GASTREICH, M.; RAREY, M. Automated Drawing of Structural Molecular Formulas under Constraints. **Journal of Chemical Information and Computer Sciences**, [*s.l*], v. 44, n. 3, p. 1065–1078, mai. 2004.

FUHRMANN, J. *et al.* A New Lamarckian Genetic Algorithm for Flexible Ligand-Receptor Docking. **Journal of computational chemistry**, [*s.l*], v. 31, n. January 2010, p. 1911–1918, 2010.

GAUTRET, P. *et al.* Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. **International Journal of Antimicrobial Agents**, [*s.l*], v. 56, n. 1, p. 105949, 2020.

GE, Y. *et al.* An integrative drug repositioning framework discovered a potential therapeutic agent targeting COVID-19. **Signal Transduction and Targeted Therapy**, [*s.l*], v. 6, n. 1, 2021.

GENHEDEN, S. *et al.* AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning. **Journal of Cheminformatics**, [*s.l*], v. 12, n. 1, p. 1–9, 2020.

GENHEDEN, S.; RYDE, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. **Expert opinion on drug discovery**, [*s.l*], v. 10, n. 5, p. 449–461, maio 2015.

GENHEDEN, S.; RYDE, U. L. F. How to obtain statistically converged MM/GBSA results. **Journal of Computational Chemistry**, [*s.l*], v. 31, n. 4, p. 837–846, 2010.

GLEESON, M. P. Generation of a Set of Simple, Interpretable ADMET Rules of Thumb. **Journal of Medicinal Chemistry**, [*s.l*], v. 51, n. 4, p. 817–834, 1 fev. 2008.

GOLDSMITH, C. S. *et al.* Ultrastructural Characterization of SARS Coronavirus. **Emerging Infectious Diseases**, [*s.l*], v. 10, n. 2, p. 320–326, 2004.

GOTE, V. *et al.* A Comprehensive Review of mRNA Vaccines. **International Journal of**

Molecular Sciences, [*s.l*], v. 24, n. 3, 2023.

GRAVAGNA, K. *et al.* Global assessment of national mandatory vaccination policies and consequences of non-compliance. **Vaccine**, [*s.l*], v. 38, n. 49, p. 7865–7873, 2020.

GURUNG, A. B. *et al.* Molecular docking of the anticancer bioactive compound proceraside with macromolecules involved in the cell cycle and DNA replication. **Genetics and molecular research**: **GMR**, [*s.l*], v. 15, n. 2, maio 2016.

HAILEAMLAK, A. Pandemics Will be More Frequent. **Ethiopian Journal of Health Sciences**, [*s.l*], v. 32, n. 2, p. 228, 2022.

HANWELL, M. D. *et al.* Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. **Journal of Cheminformatics**, [*s.l*], v. 4, n. 1, p. 17, 2012.

HEVENER, K. E. *et al.* Validation of Molecular Docking Programs for Virtual Screening against Dihydropteroate Synthase. **Journal of Chemical Information and Modeling**, [*s.l*], v. 49, n. 2, p. 444–460, 23 fev. 2009.

HOU, T. *et al.* Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. **Journal of Chemical Information and Modeling**, [*s.l*], v. 51, n. 1, p. 69–82, 24 jan. 2011.

HUANG, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. **Nature Methods**, [*s.l*], v. 14, n. 1, p. 71–73, 2017.

HUANG, J. T. *et al.* The cost of anti-Asian racism during the COVID-19 pandemic. **Nature Human Behaviour**, [*s.l*], v. 7, n. 5, p. 682–695, 2023.

HUGHES, J. D. *et al.* Physiochemical drug properties associated with in vivo toxicological outcomes. **Bioorganic & Medicinal Chemistry Letters**, [*s.l*], v. 18, n. 17, p. 4872–4875, set. 2008.

HUGHES, T. B.; MILLER, G. P.; SWAMIDASS, S. J. Modeling Epoxidation of Drug-like Molecules with a Deep Machine Learning Network. **ACS Central Science**, [*s.l*], v. 1, n. 4, p. 168–180, 22 jul. 2015.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: visual molecular dynamics. **Journal of Molecular Graphics**, [*s.l*], v. 14, n. 1, p. 27- 28,33-38, fev. 1996.

HUSSAIN, W.; AMIR, A.; RASOOL, N. Computer-aided study of selective flavonoids against chikungunya virus replication using molecular docking and DFT-based approach. **Structural Chemistry**, [*s.l*], v. 31, n. 4, p. 1363–1374, 2020.

INDU, P. *et al.* Exploring potential anti-chikungunya virus activity of phytocompounds: Computational docking and in vitro studies. **Journal of King Saud University - Science**, [*s.l*], v. 34, n. 6, p. 102157, 2022.

IVANENKOV, Y. A.; ZAGRIBELNYY, B. A.; ALADINSKIY, V. A. Are We Opening the

Door to a New Era of Medicinal Chemistry or Being Collapsed to a Chemical Singularity? **Journal of Medicinal Chemistry**, [*s.l*], v. 62, n. 22, p. 10026–10043, 27 nov. 2019.

JAIN, S. K.; AGRAWAL, A. De novo Drug Design: An overview. **Indian Journal of Pharmaceutical Sciences**, [*s.l*], v. 66, n. 6, p. 721–728, 2004.

JO, S. *et al.* CHARMM-GUI Ligand Binder for absolute binding free energy calculations and its application. **Journal of Chemical Information And Modeling**, [*s.l*], v. 53, n. 1, p. 267–277, jan. 2013.

JOHNSON, T. W.; DRESS, K. R.; EDWARDS, M. Using the Golden Triangle to optimize clearance and oral absorption. **Bioorganic & Medicinal Chemistry Letters**, [*s.l*], v. 19, n. 19, p. 5560–5564, 2009.

KATO, K. *et al.* Molecular dynamics simulations for the protein–ligand complex structures obtained by computational docking studies using implicit or explicit solvents. **Chemical Physics Letters**, [*s.l*], v. 781, p. 139022, 2021.

KERMANY, D. S. *et al.* Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. **Cell**, [*s.l*], v. 172, n. 5, p. 1122- 1131.e9, fev. 2018.

KERRIDGE, W. L. I.; SILVA, C. S. D.; UPSHUR, R. The Fragility of Scientific Rigour and Integrity in " Sped up Science ": Research Misconduct , Bias , and Hype and in the COVID - 19 Pandemic. **Journal of Bioethical Inquiry**, [*s.l*], n. 0123456789, 2023.

KIM, D.-H. *et al.* Hemin as a novel candidate for treating COVID-19 via heme oxygenase-1 induction. **Scientific Reports**, [*s.l*], v. 11, n. 1, p. 21462, 2021.

KOVAČEVIĆ, S. Z. *et al.* Prediction of In-silico ADME properties of 1,2-o-isopropylidene aldohexose derivatives. **Iranian Journal of Pharmaceutical Research**, [*s.l*], v. 13, n. 3, p. 899–908, 2014.

KÜHNE, T. D. *et al.* CP2K: An electronic structure and molecular dynamics software package -Quickstep: Efficient and accurate electronic structure calculations. **Journal of Chemical Physics**, [*s.l*], v. 152, n. 19, 2020.

KUMAR, S., NYODU, R., MAURYA, V.K., SAXENA, S.K. Morphology, Genome Organization, Replication, and Pathogenesis of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). In: Saxena, S. (eds) **Coronavirus Disease 2019 (COVID-19**). Medical Virology: From Pathogenesis to Disease Control. Singapore: Springer, 2020.

L DELANO, W. Pymol: An open-source molecular graphics tool. **{CCP4} Newsletter on Protein Crystallography**, v. 40, p. 1–8, 2002.

LI, J.; *et al.* Natural Products from the genus *Daldinia* and Their Bioactivities. [*s.l*], **Medicine Research**, [*s.l*], 2021, v. 5, n. 3, p . 210005.

LI, Y. H.; CHEN, S. P. Evolutionary history of Ebola virus. **Epidemiology and Infection**, [*s.l*], v. 142, n. 6, p. 1138–1145, 2014.

LIMA, A. H. *et al.* Molecular modeling of T. rangeli, T. brucei gambiense, and T. evansi sialidases in complex with the DANA inhibitor. **Chemical biology & drug design**, [*s.l*], v. 80, n. 1, p. 114–120, jul. 2012.

LIMA, J. DOS R. *et al.* Diterpene Sonderianin isolated from Croton blanchetianus exhibits acetylcholinesterase inhibitory action and anxiolytic effect in adult zebrafish (Danio rerio) by 5-HT system. **Journal of Biomolecular Structure and Dynamics**, [*s.l*], v. 40, n. 24, p. 13625–13640, 2022.

LIMA, J. P. O. *et al.* De novo design of bioactive phenol and chromone derivatives for inhibitors of Spike glycoprotein of SARS-CoV-2 in silico. **3 Biotech**, [*s.l*], v. 13, n. 9, p. 1–19, 2023a.

LIMA, J. P. O. *et al.* De novo design of bioactive phenol and chromone derivatives for inhibitors of Spike glycoprotein of SARS-CoV-2 in silico. **3 Biotech**, [*s.l*], v. 13, n. 9, p. 301, 2023b.

LIMA, R. C. Distanciamento e isolamento sociais pela COVID-19 no Brasil: Impactos na saúde mental. **Physis**, [*s.l*], v. 30, n. 2, p. 1–10, 2020.

LIN, L. P. *et al.* Carbon-nitrogen bond formation to construct novel polyketide-indole hybrids from the indole-3-carbinol exposed culture of Daldinia eschscholzii. **Synthetic and Systems Biotechnology**, [*s.l*], v. 7, n. 2, p. 750–755, 2022.

LIPINSKI, C. A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3. **Advanced Drug Delivery Reviews**, [*s.l*], v. 46, n. 1, p. 3–26, 2001.

LIU, H. *et al.* Salidroside promotes peripheral nerve regeneration based on tissue engineering strategy using Schwann cells and PLGA: In vitro and in vivo. **Scientific Reports**, [*s.l*], v. 7, n. July 2016, p. 1–11, 2017.

LOAIZA-CANO, V. *et al.* Antiviral role of phenolic compounds against dengue virus: A review. **Biomolecules**, [*s.l*], v. 11, n. 1, p. 1–28, 2021.

LU, S. M. *et al.* Two new ketene derivatives from the endophytic fungus Daldinia eschscholtzii J11. **Phytochemistry Letters**, [*s.l*], v. 58, n. October, p. 81–85, 2023.

MALET, H. *et al.* The Crystal Structures of Chikungunya and Venezuelan Equine Encephalitis Virus nsP3 Macro Domains Define a Conserved Adenosine Binding Pocket. **Journal of Virology**, [*s.l*], v. 83, n. 13, p. 6534–6545, 2009.

MARKOV, P. V. *et al.* The evolution of SARS-CoV-2. **Nature Reviews Microbiology**, [*s.l*], v. 21, n. 6, p. 361–379, 2023.

MARKOVIĆ, A. K. *et al.* Hydroxytyrosol, tyrosol and derivatives and their potential effects on human health. **Molecules**, Basel, v. 24, n. 10, 2019.

MARTINI, M. *et al.* The Spanish Influenza Pandemic: A lesson from history 100 years after 1918. **Journal of Preventive Medicine and Hygiene**, [*s.l*], v. 60, n. 1, p. E64–E67, 2019.

MATLOCK, M. K.; HUGHES, T. B.; SWAMIDASS, S. J. XenoSite server: A web-available site of metabolism prediction tool. **Bioinformatics**, Oxford, v. 31, n. 7, p. 1136–1137, 2015.

MAZOLA, Y. *et al.* A comparative molecular dynamics study of thermophilic and mesophilic β-fructosidase enzymes. **Journal of molecular modeling**, [*s.l*], v. 21, n. 9, p. 228, set. 2015.

MCCOLLUM, A. M. *et al.* Poxvirus viability and signatures in historical relics. **Emerging Infectious Diseases**, [*s.l*], v. 20, n. 2, p. 177–184, 2014.

MEHRANFAR, A.; IZADYAR, M. Theoretical Design of Functionalized Gold Nanoparticles as Antiviral Agents against Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). **Journal of Physical Chemistry Letters**, [*s.l*], v. 11, n. 24, p. 10284–10289, 2020.

MEYER, H.; EHMANN, R.; SMITH, L. Smallpox in the Post-Eradication Era. **Viruses**. Basel, v. 12, 2, jan 2020. p. 1–11.

MIAO, R. *et al.* Improved Classification of Blood-Brain-Barrier Drugs Using Deep Learning. **Scientific Reports**, London, v. 9, n. 1, p. 8802, 2019.

MOEN, E. *et al.* Deep learning for cellular image analysis. **Nature Methods**, New York, v. 16, n. 12, p. 1233–1246, 2019.

MOLYNEUX, D. H. *et al.* The history of the neglected tropical disease movement. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, [*s.l*], v. 115, n. 2, p. 169–175, 2021.

MORRIS, G. M. *et al.* Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, [*s.l*], v. 30, n. 16, p. 2785–2791, 2009a.

MORRIS, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **Journal of computational chemistry**, [*s.l*], v. 30, n. 16, p. 2785–2791, dez. 2009b.

NEESE, F. *et al.* The ORCA quantum chemistry program package. **Journal of Chemical Physics**, [*s.l*], v. 152, n. 22, 2020.

NGUYEN, D. D. *et al.* Rigidity Strengthening: A Mechanism for Protein-Ligand Binding. **Journal of Chemical Information and Modeling**, [*s.l*], v. 57, n. 7, p. 1715–1721, 2017.

NOCCHI, S. R. *et al.* Development and evaluation of topical formulations that contain hydroethanolic extract from Schinus terebinthifolia against HSV-1 infection. **Brazilian Journal of Pharmaceutical Sciences,** São Paulo, v. 58, p. 1–11, 2022.

NOGUEIRA, F.; BRITO, J. C. The costly hydroxychloroquine battle in Brazil. **Journal of**

Evidence-Based Healthcare, [*s.l*], v. 2, n. 1, p. 45, 2020.

NUNES DA ROCHA, M. *et al.* Predictive ADMET study of rhodanine-3-acetic acid chalcone derivatives. **Journal of the Indian Chemical Society**, [*s.l*], v. 99, n. 7, 2022.

O'BOYLE, N. M. *et al.* Open Babel: An open chemical toolbox. **Journal of Cheminformatics**, [*s.l*], v. 3, n. 1, p. 33, 2011.

OLIVEIRA, V. M. DE *et al.* Molecular docking identification for the efficacy of natural limonoids against COVID-19 virus main protease. **Journal of the Indian Chemical Society**, [*s.l*], n. September, p. 6, 2021.

ONUFRIEV, A.; BASHFORD, D.; CASE, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. **The Journal of Physical Chemistry B**, [*s.l*], v. 104, n. 15, p. 3712–3720, 1 abr. 2000.

PATINY, L. *et al.* The C6H6 NMR repository: An integral solution to control the flow of your data from the magnet to the public. **Magnetic Resonance in Chemistry**, [*s.l*], v. 56, n. 6, p. 520–528, 2018.

PETROVA, S. S.; SOLOV'EV, A. D. The Origin of the Method of Steepest Descent. **Historia Mathematica**, [*s.l*], v. 24, n. 4, p. 361–375, 1997.

PETTERSEN, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. **J Comput Chem**, [*s.l*], v. 25, 1 jan. 2004.

PHILLIPS, J. C. *et al.* Scalable molecular dynamics with NAMD. **Journal of Computational Chemistry**, [*s.l*], v. 26, n. 16, p. 1781–1802, dez. 2005.

PIRES, D. E. V; KAMINSKAS, L. M.; ASCHER, D. B. Prediction and Optimization of Pharmacokinetic and Toxicity Properties of the Ligand. **Methods in molecular biology** (Clifton, N.J.), [*s.l*], v. 1762, p. 271–284, 2018.

PLOTNIKOV, M. B.; PLOTNIKOVA, T. M. Tyrosol as a Neuroprotector: Strong Effects of a "Weak" Antioxidant. **Current Neuropharmacology**, [*s.l*], v. 19, n. 4, p. 434–448, 2020.

POWERS, A. M. Vaccine and therapeutic options to control chikungunya virus. **Clinical Microbiology Reviews**, [*s.l*], v. 31, n. 1, p. 1–29, 2018.

PURANIK, N. V. *et al.* Evaluation of the antiviral potential of halogenated dihydrorugosaflavonoids and molecular modeling with nsp3 protein of chikungunya virus (chikv). **ACS Omega**, [*s.l*], v. 4, n. 23, p. 20335–20345, 2019.

R. P. D. BANK. **RCSB PDB - 3TRK**: Structre of the Chikungunya virus nsP2 protease, [*s.d*] Disponível em: <https://www.rcsb.org/structure/3TRK>. Acesso em: 4 out. 2023.

RABI, F. A. *et al.* Sars-cov-2 and coronavirus disease 2019: What we know so far. **Pathogens**, [*s.l*], v. 9, n. 3, p. 1–14, 2020.

RAMOS, A. L. B. M. *et al.* A eficiência das ações de combate à dengue na atenção primária à saúde no Brasil / The effiency of actions to combat dengue in primary healthcare in Brazil. **Brazilian Journal of Health Review**, [*s.l*], v. 4, n. 3, p. 10575–10595, 2021.

RASTELLI, G. *et al.* Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. **Chemical Biology & Drug Design**, [*s.l*], v. 73, n. 3, p. 283–286, mar. 2009.

ROSA, A. *et al.* SARS-CoV-2 can recruit a heme metabolite to evade antibody immunity. **Science Advances**, [*s.l*], v. 7, n. 22, p. 1–15, 2021.

ROUGERON, V. *et al.* Chikungunya, a paradigm of neglected tropical disease that emerged to be a new health global risk. **Journal of Clinical Virology**, [*s.l*], v. 64, p. 144–152, 2015.

SADYBEKOV, A. V.; KATRITCH, V. Computational approaches streamlining drug discovery. **Nature**, [*s.l*], v. 616, n. 7958, p. 673–685, 2023.

SAKAI, T.; MORIMOTO, Y. The History of Infectious Diseases and Medicine. **Pathogens**, [*s.l*], v. 11, n. 10, 2022.

SANTOS, K. B. *et al.* Highly Flexible Ligand Docking: Benchmarking of the DockThor Program on the LEADS-PEP Protein-Peptide Data Set. **Journal of Chemical Information and Modeling**, [*s.l*], v. 60, n. 2, p. 667–683, 2020.

SARKAR, S.; GARDNER, L. Zika: The cost of neglect. **Palgrave Communications**, [*s.l*], v. 2, n., p. 1–6, set. 2016.

SCOTTI, L. *et al.* SAR, QSAR and docking of anticancer flavonoids and variants: a review. **Current topics in medicinal chemistry**, [*s.l*], v. 12, n. 24, p. 2785–2809, 2012.

SEGLER, M. H. S.; PREUSS, M.; WALLER, M. P. Planning chemical syntheses with deep neural networks and symbolic AI. **Nature**, [*s.l*], v. 555, n. 7698, p. 604–610, mar. 2018.

SEIDENSTUECKER, M. *et al.* Composite material consisting of microporous β-TCP ceramic and alginate for delayed release of antibiotics. **Acta Biomaterialia**, [*s.l*], v. 51, p. 433–446, mar. 2017.

SHANG, J. *et al.* Cell entry mechanisms of SARS-CoV-2. **Proceedings of the National Academy of Sciences**, [*s.l*], v. 117, n.21, p. 11727-11734, mai. 2020.

SHARP, P. M.; HAHN, B. H. Origins of HIV and the AIDS pandemic. **Cold Spring Harbor Perspectives in Medicine**, [*s.l*], v. 1, n. 1, p. 1–22, set. 2011.

SHIRAZI, F. M. *et al.* Repurposing the drug, ivermectin, in COVID-19: toxicological points of view. **European Journal of Medical Research**, [*s.l*], v. 27, n. 1, p. 1–11, fev. 2022.

SHITYAKOV, S.; FÖRSTER, C. In silico predictive model to determine vector-mediated transport properties for the blood-brain barrier choline transporter. **Advances and Applications in Bioinformatics and Chemistry**, [*s.l*], v. 7, n. 1, p. 23–36, set. 2014.

SILVA, S. J. R. DA; MAGALHÃES, J. J. F. DE; PENA, L. Simultaneous Circulation of DENV, CHIKV, ZIKV and SARS-CoV-2 in Brazil: an Inconvenient Truth. **One Health**, [*s.l*], v. 16, n.12, p. 2020–2022, jun. 2021.
SINGH, R.; BHARDWAJ, V. K.; PUROHIT, R. Inhibition of nonstructural protein 15 of SARS-CoV-2 by golden spice: A computational insight. **Cell Biochemistry and Function**, [*s.l*], v. 40, n. 8, p. 926–934, out., 2022.

SINGH, R.; PUROHIT, R. Multi-target approach against SARS-CoV-2 by stone apple molecules: A master key to drug design. **Phytotherapy Research**, [*s.l*], v. 38, n.1 p.7-10, fev, 2023a.

SINGH, R.; PUROHIT, R. Computational analysis of protein-ligand interaction by targeting a cell cycle restrainer. **Computer Methods and Programs in Biomedicine**, [*s.l*], v. 231, p. 107367, abr. 2023b.

SOBNGWI, E. *et al.* Doxycycline vs Hydroxychloroquine + Azithromycin in the Management of COVID-19 Patients: An Open-Label Randomized Clinical Trial in Sub-Saharan Africa (DOXYCOV). **Cureus**, [*s.l*], v. 15, n. 9, p. 1–9, set. 2023.

SPYROU, M. A. *et al.* The source of the Black Death in fourteenth-century central Eurasia. **Nature**, [*s.l*], v. 606, n. 7915, p. 718–724, jun. 2022.

STIERAND, K.; MAASS, P. C.; RAREY, M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. **Bioinformatics**. Oxford, v. 22, n. 14, p. 1710–1716, jul. 2006.

SUTTON, R. S.; MAHMOOD, A. R.; WHITE, M. An emphatic approach to the problem of off-policy temporal-difference learning. **Journal of Machine Learning Research**, [*s.l*], v. 17, mai. 2016.

THANH TUNG, B. *et al.* Screening Virtual ACE2 Enzyme Inhibitory Activity of Compounds for COVID-19 Treatment Based on Molecular Docking. **VNU Journal of Science: Medical and Pharmaceutical Science**s, [*s.l*], v. 36, n. 4, dez. 2020.

TIAN, H.; KETKAR, R.; TAO, P. ADMETboost: a web server for accurate ADMET prediction. **Journal of Molecular Modeling**, [*s.l*], v. 28, n. 12, p. 1–6, dez. 2022.

TORNG, W.; ALTMAN, R. B. Graph Convolutional Neural Networks for Predicting Drug-Target Interactions. **Journal of Chemical Information and Modeling**, [*s.l*], v. 59, p. 4141-4149, out. 2019.

TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, [*s.l*], v. 31, n. 2, p. 455–461, jan. 2010.

TUMSKIY, R. S.; TUMSKAIA, A. V. Multistep rational molecular design and combined docking for discovery of novel classes of inhibitors of SARS-CoV-2 main protease 3CLpro. **Chemical Physics Letters**, [*s.l*], v. 780, p. 138894, out. 2021.

VAN DE WATERBEEMD, H.; GIFFORD, E. ADMET in silico modelling: towards prediction paradise? **Nature Reviews Drug Discovery**, [*s.l*], v. 2, n. 3, p. 192–204, mar. 2003.

VANOMMESLAEGHE, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. **Journal of Computational Chemistry**, [*s.l*], v. 31, n. 4, p. 671–690, mar. 2010.

VICENTE, C. R. *et al.* Impact of concurrent epidemics of Dengue, Chikungunya, Zika, and COVID-19. **Revista da Sociedade Brasileira de Medicina Tropical**, [*s.l*], v. 54, p. e0837–2020, fev.2021.

VOLKAMER, A. *et al.* Analyzing the topology of active sites: On the prediction of pockets and subpockets. **Journal of Chemical Information and Modeling**, [*s.l*], v. 50, n. 11, p. 2041–2052, 2010.

VOLKAMER, A. *et al.* Combining global and local measures for structure-based druggability predictions. **Journal of Chemical Information and Modeling**, [*s.l*], v. 52, n. 2, p. 360–372, 2012a.

VOLKAMER, A. *et al.* Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment. **Bioinformatics**, Oxford, v. 28, n. 15, p. 2074–2075, 2012b.

VON DELFT, A. *et al.* Accelerating antiviral drug discovery: lessons from COVID-19. **Nature Reviews Drug Discovery**, [*s.l*], v. 22, n. 7, p. 585–603, 2023.

WAGENER, F. A. D. T. G. *et al.* Targeting the Heme-Heme Oxygenase System to Prevent Severe Complications Following COVID-19 Infections. **Antioxidants** , Basel, v. 9, n. 6, jun. 2020.

WAGER, T. T. *et al.* Defining Desirable Central Nervous System Drug Space through the Alignment of Molecular Properties, in Vitro ADME, and Safety Attributes. **ACS Chemical Neuroscience**, [*s.l*], v. 1, n. 6, p. 420–434, 16 jun. 2010.

WAGER, T. T. *et al.* Central Nervous System Multiparameter Optimization Desirability: Application in Drug Discovery. **ACS Chemical Neuroscience**, [*s.l*], v. 7, n. 6, p. 767–775, 2016a.

WAGER, T. T. *et al.* Central Nervous System Multiparameter Optimization Desirability: Application in Drug Discovery. **ACS Chemical Neuroscience**, [*s.l*], v. 7, n. 6, p. 767–775, 15 jun. 2016b.

WAHL, J. *et al.* Accuracy evaluation and addition of improved dihedral parameters for the MMFF94s. **Journal of Cheminformatics**, [*s.l*], v. 11, n. 1, p. 53, 2019.

WANG, C. *et al.* Recent Developments and Applications of the MMPBSA Method. **Frontiers in molecular biosciences**, [*s.l*], v. 4, p. 87, 2017.

WANG, M. *et al.* Deep learning approaches for de novo drug design: An overview. **Current Opinion in Structural Biology**, [*s.l*], v. 72, n. November 2021, p. 135–144, 2022.

WANG, M. Y. *et al.* SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. **Frontiers in Cellular and Infection Microbiology**, [*s.l*], v. 10, n. November, p. 1–17, 2020.

WERTHEIM, J. O. Viral Evolution: Mummy Virus Challenges Presumed History of Smallpox. **Current Biology**, [*s.l*], v. 27, n. 3, p. R119–R120, 2017.

WHO. **WHO COVID-19 Dashboard**. Disponível em: https://covid19.who.int. Acesso em 18 dez. 2023.

WHO. **Neglected tropical diseases**. [s.l], Disponível em: https://www.who.int/health-topics/neglected-tropical-diseases#tab=tab_1. Acesso em 30 nov. 2023.

WINKLER, A. The Approach of the Black Death in Switzerland and the Persecution of Jews, 1348 – 1349. **Swiss American Historical Society Review**, [*s.l*], v. 43, n. 3, p. 4–23, nov. 2007.

WISE, J. Covid-19: WHO declares end of global health emergency. **BMJ**, [*s.l*], v. 381, n. 2, May, p. 1041, 2023.

WU, J. *et al.* Identification of raloxifene as a novel α-glucosidase inhibitor using a systematic drug repurposing approach in combination with cross molecular docking-based virtual screening and experimental verification. **Carbohydrate research**, [*s.l*], v. 511, p. 108478, jan. 2022.

WU, N. *et al.* Long-term effectiveness of COVID-19 vaccines against infections, hospitalisations, and mortality in adults: findings from a rapid living systematic evidence synthesis and meta-analysis up to December, 2022. **The Lancet Respiratory Medicine**, [*s.l*], v. 11, n. 5, p. 439–452, 2023.

WUTTHIWONG, N. *et al.* Daldiniaeschsone a, a rare tricyclic polyketide having a chromone unit fused to a δ-lactone and its symmetrical biphenyl dimer, daldiniaeschsone b, from an endophytic fungus *Daldinia eschscholtzii* SDBR-CMUNKC745. **Journal of Fungi**, Basel, v. 7, n. 5, maio, 2021.

XIN, H. *et al.* Estimating the Latent Period of Coronavirus Disease 2019 (COVID-19). **Clinical Infectious Diseases**, [*s.l*], v. 74, n. 9, p. 1678–1681, 2022.

XIONG, G. *et al.* ADMETlab 2.0: An integrated online platform for accurate and comprehensive predictions of ADMET properties. **Nucleic Acids Research**, [*s.l*], v. 49, n. W1, p. W5–W14, 2021.

YAN, J. *et al.* α-Glucosidase inhibition by luteolin: Kinetics, interaction and molecular docking. **International Journal of Biological Macromolecules**, [*s.l*], v. 64, p. 213–223, 2014.

YUSUF, D. *et al.* An alternative method for the evaluation of docking performance: RSR vs RMSD. **Journal of Chemical Information and Modeling**, [*s.l*], v. 48, n. 7, p. 1411–1422, 2008.

YUYAMA, K. T. *et al.* isolated from the Brazilian Amazon : taxonomic features and mycelial growth conditions. **Acta Amazonica**, [*s.l*], v. 43, n. 92, p. 1–8, 2013.

ZARA, A. L. DE S. A. *et al.* Estratégias de controle do Aedes aegypti: uma revisão. Epidemiologia e servicos de saúde. **Revista do Sistema Único de Saúde do Brasil**, [*s.l*], v. 25, n. 2, p. 391–404, 2016.

ZHANG, D. *et al.* New phenol and chromone derivatives from the endolichenic fungus Daldinia species and their antiviral activities. **RSC Advances**, [*s.l*], v. 11, n. 36, p. 22489–22494, 2021a.

ZHANG, J. *et al.* Structural impact on SARS-CoV-2 spike protein by D614G substitution. **Science**, [*s.l*], v. 372, n. 6541, p. 525–530, 2021b.

ZHANG, S. *et al.* Pyrimidone inhibitors targeting Chikungunya Virus nsP3 macrodomain by fragment-based drug design. **PLoS ONE**, [*s.l*], v. 16, n. 1 January, p. 1–17, 2021c.

ZHENG, M. *et al.* Site of metabolism prediction for six biotransformations mediated by cytochromes P450. **Bioinformatics**. Oxford, v. 25, n. 10, p. 1251–1258, maio 2009.

APPENDIX A – WEBSERVERS AND SOFTWARES MAIN PAGES AND EASE OF ACCESS TO DATA

Name	Link
ADMETboost	<https://ai-druglab.smu.edu/admet>
ADMETlab 2.0	<https://admetmesh.scbdd.com/>
AiZynthFinder	<https://github.com/MolecularAI/aizynthfinder>
AutoDock Vina	<https://vina.scripps.edu/>
C6H6	<https://www.cheminfo.org/flavor/c6h6/index.html>
CP2K	<https://www.cp2k.org/>
Gabedit	<https://gabedit.sourceforge.net/>
GROMACS	<https://www.gromacs.org/>
Manifold	<https://app.postera.ai/>
MarvinSketch	<https://chemaxon.com/marvin>
MolAICal	<https://molaical.github.io/>
ORCA	<https://orcaforum.kofo.mpg.de/app.php/portal>
PaccMann^{rl}	<https://github.com/PaccMann/paccmann_rl>
pkCSM	<https://biosig.lab.uq.edu.au/pkcsm/>
PreADMET	<https://preadmet.webservice.bmdrc.org/>
ProTox-II	<https://tox-new.charite.de/protox_II/>
PyMOL	<https://pymol.org/2/>
STopTox	<https://stoptox.mml.unc.edu/>
SwissADME	<http://www.swissadme.ch/>
UCSF Chimera	<https://www.cgl.ucsf.edu/chimera/>
XenoSite	<https://xenosite.org/>

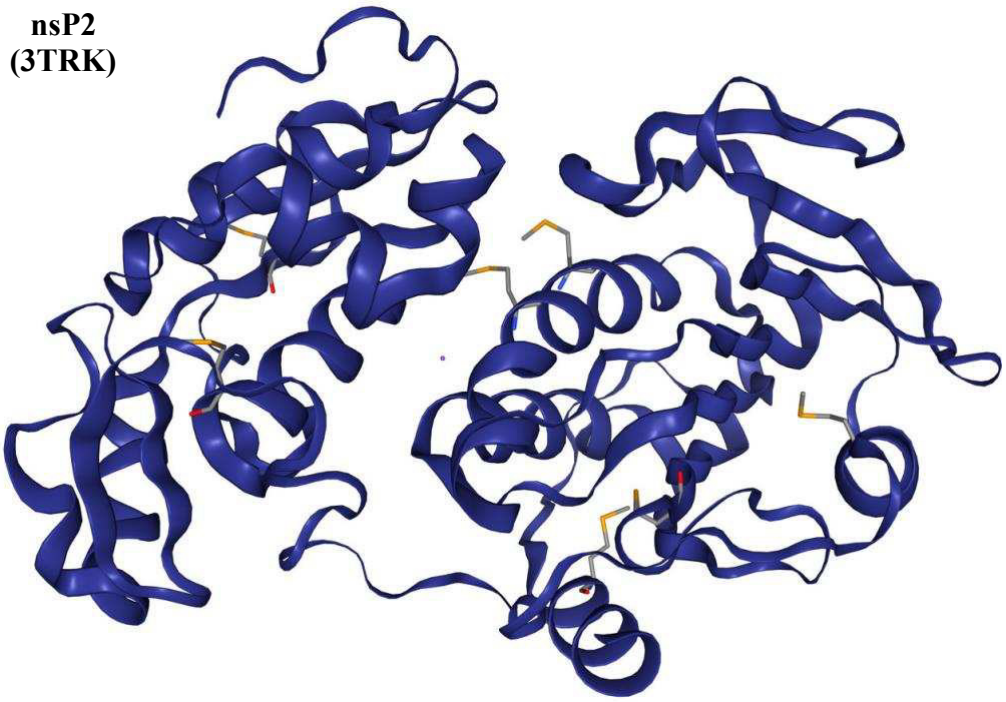
To favor the continued work and provide the data for scientific checking (after publication of related papers), the following QR code are presented to access the Github folder (<https://github.com/PetrusChem/In-silico-inhibition-of-SARS-CoV-2-and-CHIKV-by-phenol-and-chromone-derivatives>), containing data and shortcuts, for the print version.

APPENDIX B – SMILES STRINGS AND PUBCHEM LINKS

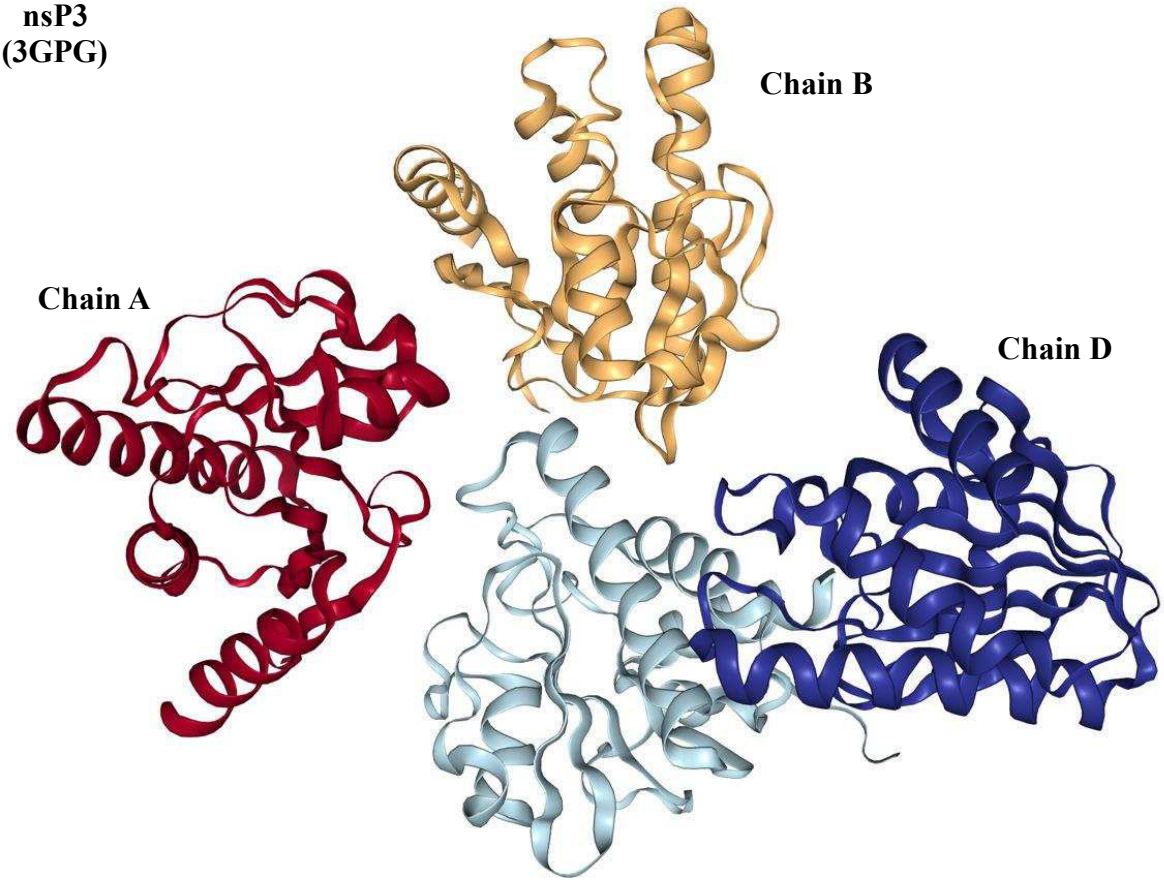
Compounds	SMILES	Analysis	PubChem link
Der1	C[C@@H](CO)C(=O)OCCC1=CC=C(O)C=C1	From the www woo Zhang *et al.* (ZHANG *et al.*, 2021a) <https://www.rsc.org/suppdata/d1/ra/d1ra03754d/d1ra03754d1.pdf>	–
Der2	CC(CO)C(=O)OCC(C)C(=O)OCCC1=CC=C(O)C=C1		–
Der3	C[C@H](O)CC(=O)OCCC1=CC=C(O)C=C1		–
Der4	OC[C@H]1O[C@@H](OCCC2=CC=C(O)C=C2)[C@H](O)[C@@H](O)[C@@H]1O		<https://pubchem.ncbi.nlm.nih.gov/compound/159278>
Der5	CC(=C)[C@H](O)COC1=CC=C(CCO)C=C1		<https://pubchem.ncbi.nlm.nih.gov/compound/50993978>
Der6	COC1=C(O)C=CC(CCO)=C1		<https://pubchem.ncbi.nlm.nih.gov/compound/16928>
Der7	COC1=C(O)C=C(CCO)C=C1		<https://pubchem.ncbi.nlm.nih.gov/compound/10034991>
Der8	OCCC1=CC=C(O)C=C1		<https://pubchem.ncbi.nlm.nih.gov/compound/10393>
Der9	CC1=CC(=O)C2=C(C[C@H](O)C[C@H]2O)O1		–
Der10	CCCC1=CC(=O)C2=C(C[C@H](O)C[C@H]2O)O1		–
Der11	CC1=CC(=O)C2=C(C[C@@H](O)C[C@H]2O)O1		<https://pubchem.ncbi.nlm.nih.gov/compound/156582518>
Der12	CCCC1=CC(=O)C2=C(C[C@@H](O)C[C@H]2O)O1		<https://pubchem.ncbi.nlm.nih.gov/compound/156582517>
Hit48	CC(=O)NCC1=CC(=CC=C1O)C(I)COC1=CC(Br)=CC=C1		–
Hit250	OC(=O)C1=CC(CCOC2=CC3=C(C(=O)N(C=O)C3=O)C(F)=C2C(=C)C2=CC=CN2)=C(CF)C=C1O		–
Prednisone	CC12CC(=O)C3C(C1CCC2(C(=O)CO)O)CCC4=CC(=O)C=CC34C		<https://pubchem.ncbi.nlm.nih.gov/compound/Prednisone>
Dipyrone	CC1=C(C(=O)N(N1C)C2=CC=CC=C2)N(C)CS(=O)(=O)[O-].[Na+]		<https://pubchem.ncbi.nlm.nih.gov/compound/522325>
Paracetamol	CC(=O)NC1=CC=C(C=C1)O		<https://pubchem.ncbi.nlm.nih.gov/compound/Acetaminophen>

APPENDIX C – TARGETS NSP2 AND NSP3 3D STRUCTURES

nsP2
(3TRK)

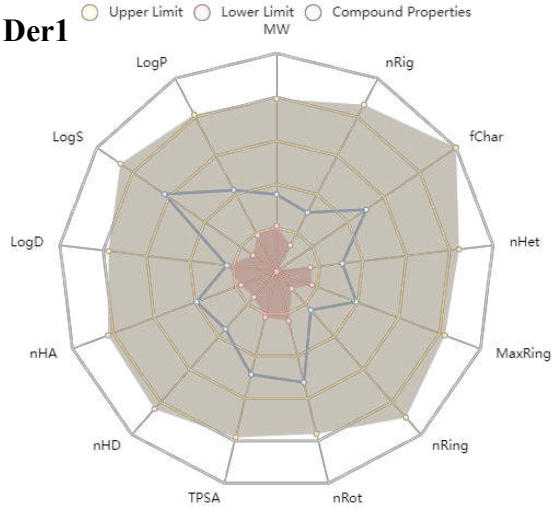


nsP3
(3GPG)

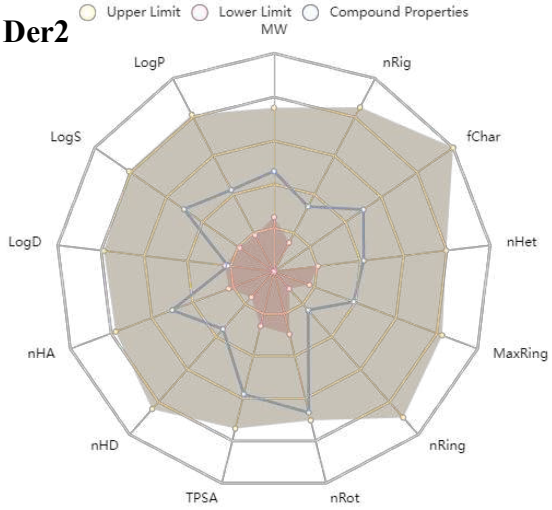


APPENDIX D – RELATIONSHIP BETWEEN STRUCTURE AND DRUGLIKENESS OF DER1 TO DER12

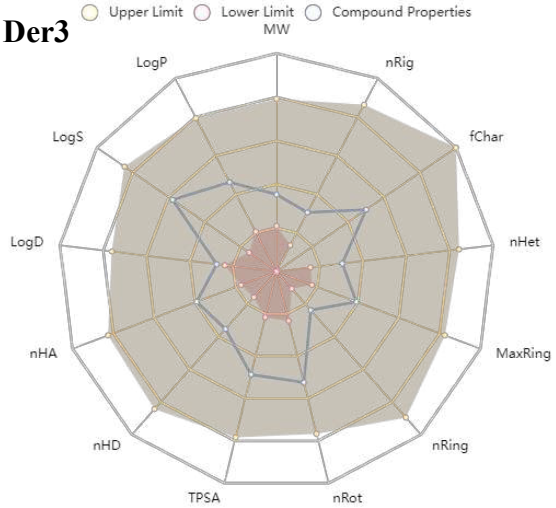
Der1



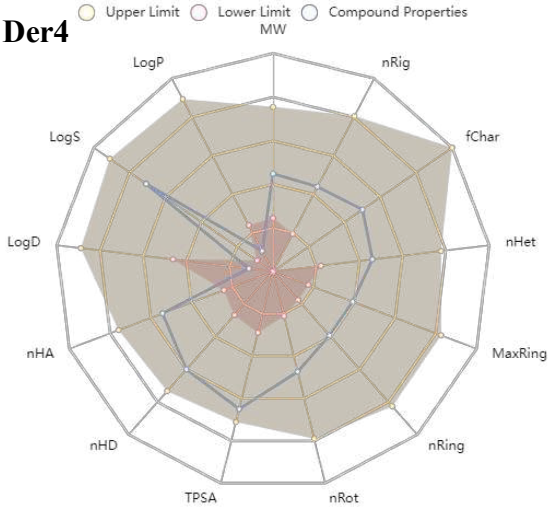
Der2



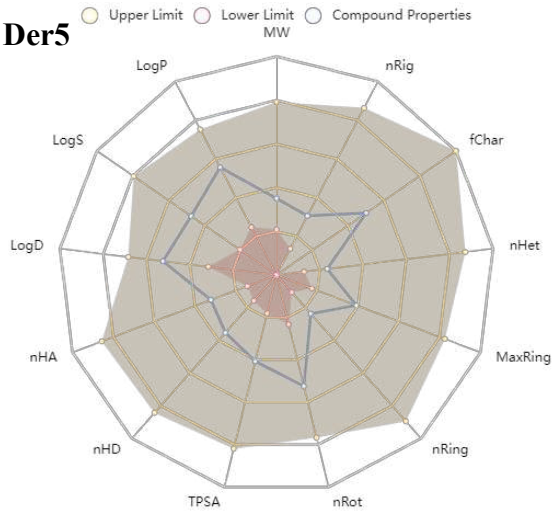
Der3



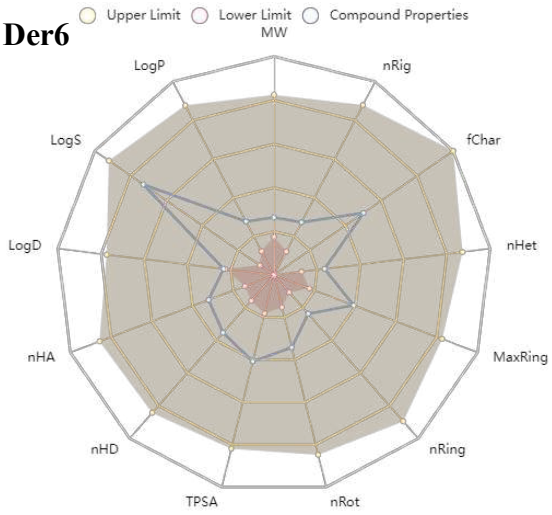
Der4



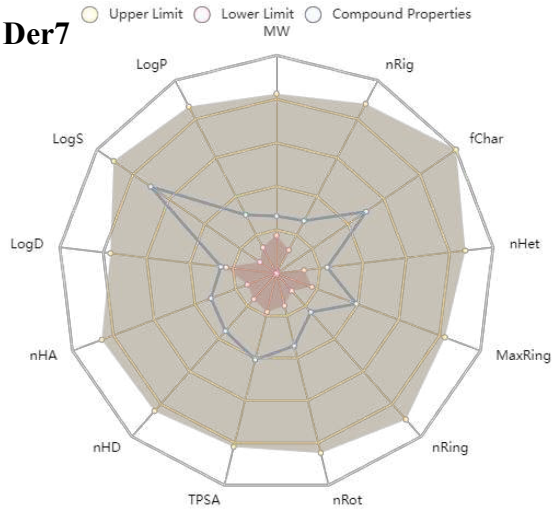
Der5



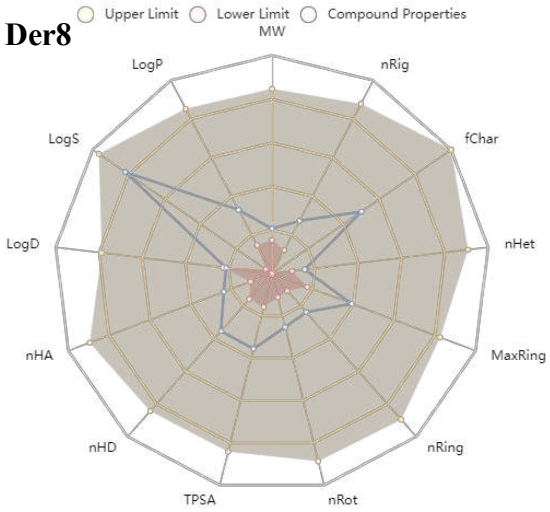
Der6



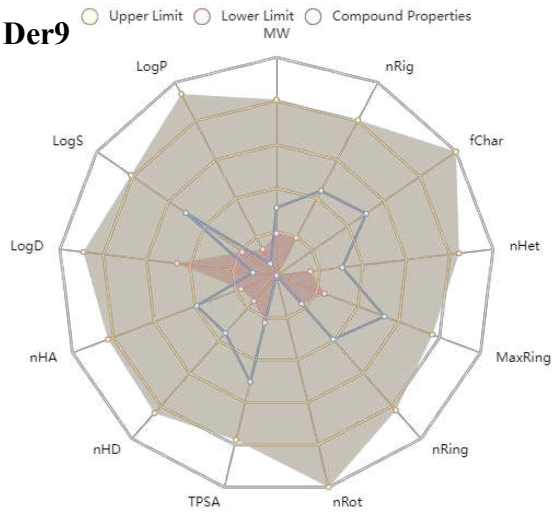
Der7



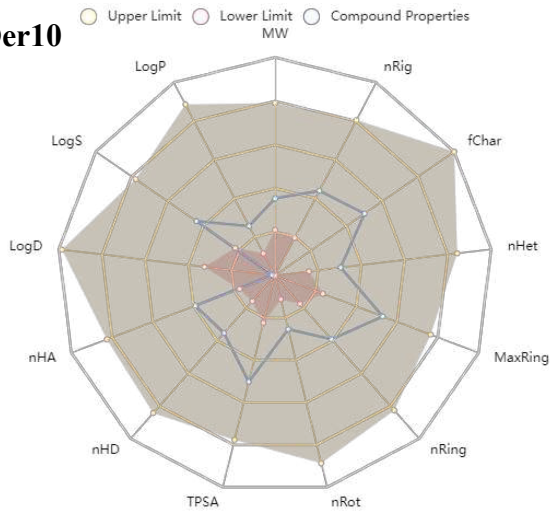
Der8



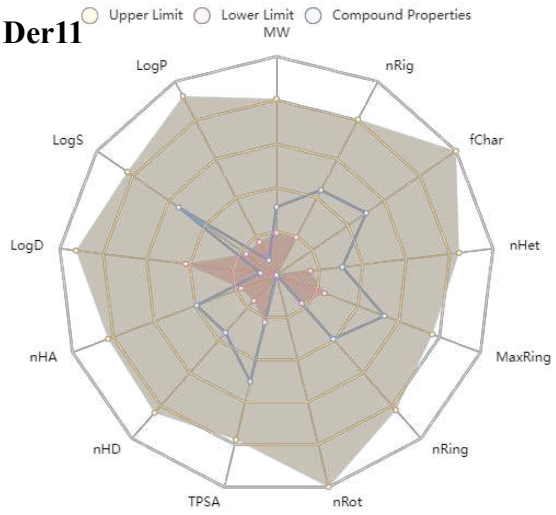
Der9



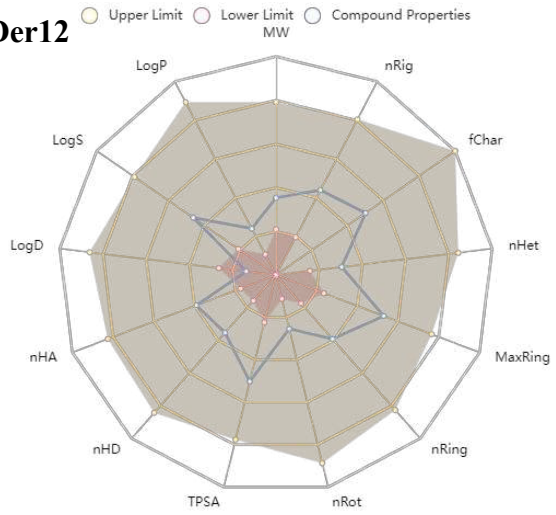
Der10



Der11

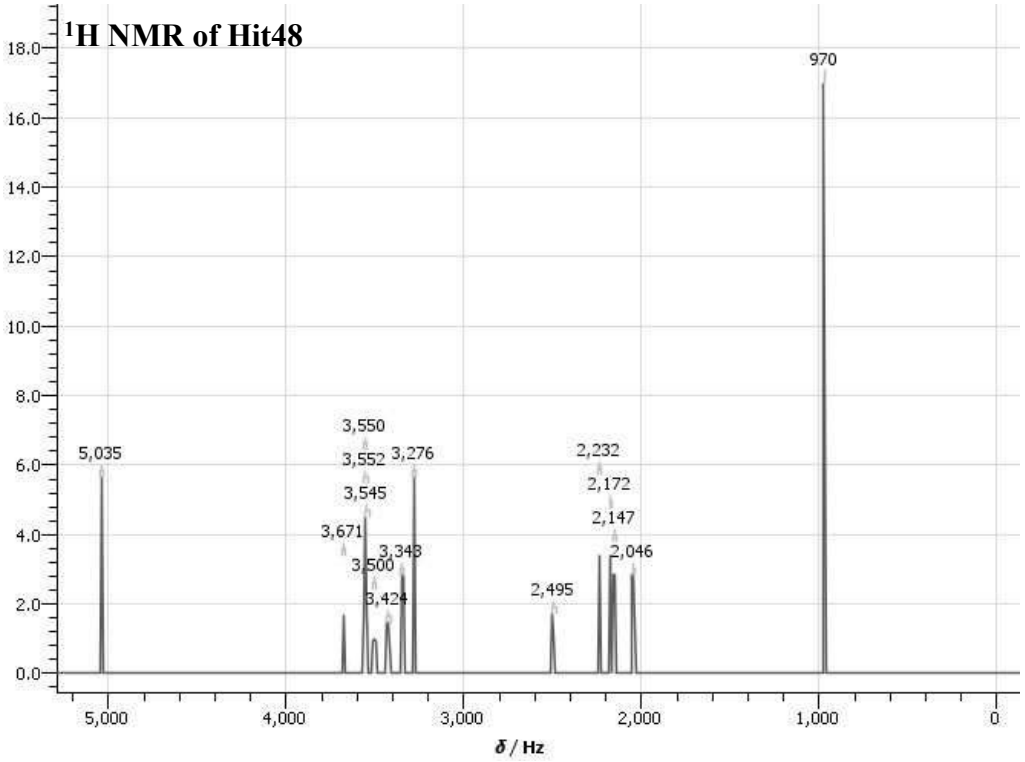

Upper Limit
Lower Limit
Compound Properties
MW
nRig
fChar
nHet
MaxRing
nRing
nRot
TPSA
nHD
nHA
LogD
LogS
LogP

Der12

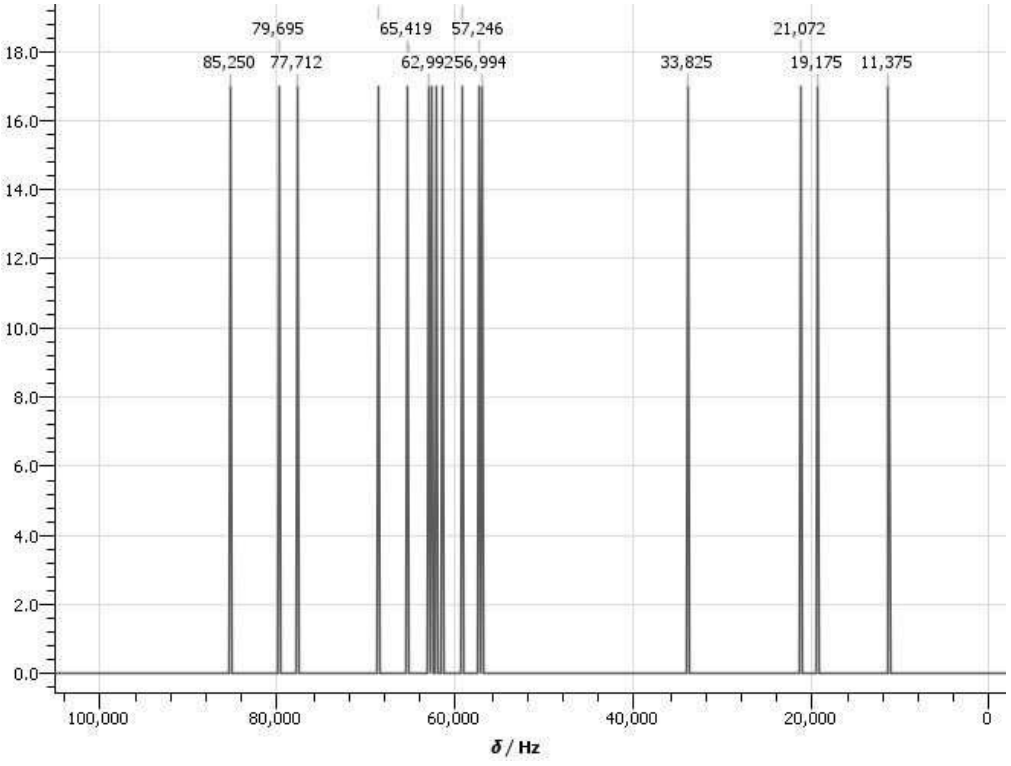

Upper Limit
Lower Limit
Compound Properties
MW
nRig
fChar
nHet
MaxRing
nRing
nRot
TPSA
nHD
nHA
LogD
LogS
LogP

APPENDIX E – PREDICTED NMR SPECTRA FOR HIT 48 AND HIT 250

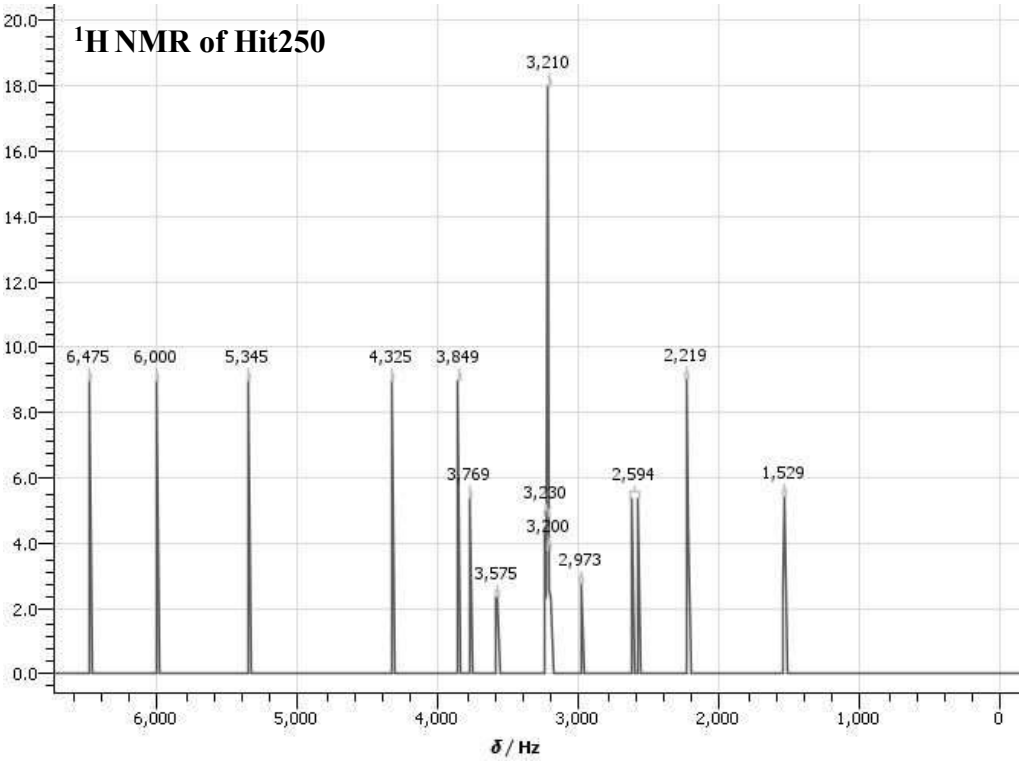
^{1}H NMR of Hit48



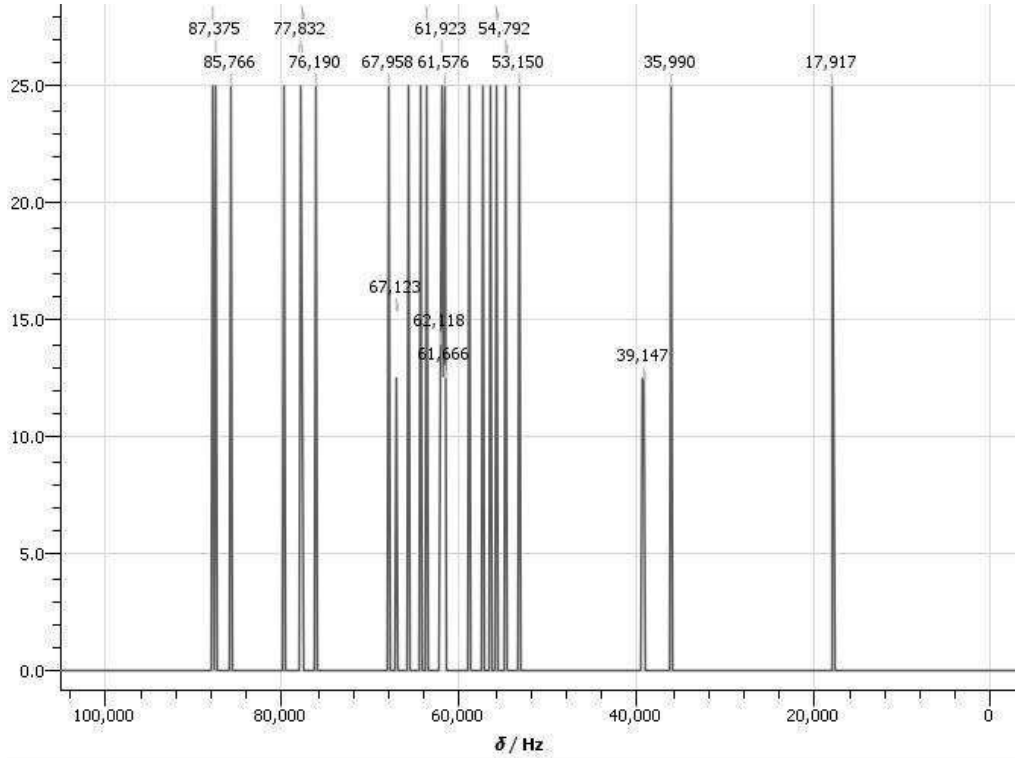
^{13}C NMR of Hit48



^{1}H NMR of Hit250



^{13}C NMR of Hit250



APPENDIX F – SCIENTIFIC PRODUCTION RELATED TO THE THESIS

Paper 1: *De Novo Design* of bio active phenol and chromone derivatives for Inhibitors of Spike glycoprotein of SARS-CoV-2 *in silico*.

Status: published.

Authors: LIMA, J. P. O.; FONSECA, A. M.; MARINHO, G. S.; ROCHA, M. N.; MARINHO, E. M.; SANTOS, H. S.; MARINHO, E. S.; NETO, P. L.; FECHINE, P. B. A.

Periodic: 3 Biotech.

Year: 2023.

3 Biotech (2023) 13:301
https://doi.org/10.1007/s13205-023-03695-9

ORIGINAL ARTICLE

De novo design of bioactive phenol and chromone derivatives for inhibitors of Spike glycoprotein of SARS-CoV-2 in silico

Joan Petrus Oliveira Lima[1] · Aluísio Marques da Fonseca[2] · Gabrielle Silva Marinho[3] · Matheus Nunes da Rocha[3] · Emanuelle Machado Marinho[1] · Helcio Silva dos Santos[4] · Rafael Melo Freire[5] · Emmanuel Silva Marinho[3] · Pedro de Lima-Neto[1] · Pierre Basílio Almeida Fechine[1]

Received: 16 March 2023 / Accepted: 29 June 2023 / Published online: 14 August 2023

Abstract
This work presents the synthesis of 12 phenol and chromone derivatives, prepared by the analogs, and the possibility of conducting an in silico study of its derivatives as a therapeutic alternative to combat the SARS-CoV-2, pathogen responsible for COVID-19 pandemic, using its S-glycoprotein as a macromolecular target. After the initial screening for the ranking of the products, it was chosen which structure presented the best energy bond with the target. As a result, derivative 4 was submitted to a molecular growth study using artificial intelligence, where 8436 initial structures were obtained that passed through the interaction filters and similarity to the active glycoprotein pocket through the MolAICal computational package. Thus, 557 Hits with active configuration were generated, which is very promising compared to the BLA reference link for inhibiting the biological target. Molecular dynamics also simulated these compounds to verify their stability within the active protein site to seek new therapeutic propositions to fight against the pandemic. The Hit 48 and 250 are the most active compounds against SARS-CoV-2. In summary, the results show that the Hit 250 would be more active than the natural compound, which could be further developed for further testing against SARS-CoV-2. The study employs the de novo approach to design new drugs, combining artificial intelligence and molecular dynamics simulations to create efficient molecular structures. This research aims to contribute to the development of effective therapeutic strategies against the pandemic.

Keywords Main protease · COVID-19 · Molecular docking · Pandemic · Deep learning

Paper 2: Phenol and chromone compounds for *in silico* inhibition of nsP2 and nsP3 of Chikungunya virus.

Status: submitted.

Authors: LIMA, J. P. O.; ROBERTO, C. H. A.; ROCHA, M. N.; OLIVEIRA, V. M.; MARINHO, E. S.; NETO, P. L.; FECHINE, P. B. A.

Periodic: DARU Journal of Pharmaceutical Sciences.

Year: 2023.

DARU Journal of Pharmaceutical Sciences

Phenol and chromone compounds for in silico inhibition of nsP2 and nsP3 of Chikungunya virus

--Manuscript Draft--

Manuscript Number:		
Full Title:	Phenol and chromone compounds for in silico inhibition of nsP2 and nsP3 of Chikungunya virus	
Article Type:	Research article	
Manuscript Classifications:	20.020: Drug analysis	
Funding Information:	Conselho Nacional de Desenvolvimento Científico e Tecnológico (308452/2022-4)	Professor Pierre Fechine
	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Finance Code 001 - PROEX 23038.000509/2020-82)	Professor Pierre Fechine
	Fondecyt (11200425)	Professor Rafael Melo Freire
Abstract:	Background The rising concern about viral neglected tropical diseases as chikungunya imposes a global challenge as the climate changes spread the vectors to new regions and increasing their urban proliferation period. Therefore, it is needed to develop a safe antiviral drug from novel potential biomolecules. Objective the in silico inhibition of CHIKV targets by novel chromone and phenol derivatives compounds from Daldinia sp. and the evaluation of their safety for human oral use. Results The docking simulations for the nsP2 showed mild binding, in the nsP3 all the derivatives presented -6 kcal mol-1 binding affinity and interacts with crucial residues in the replication cycle of CHIKV, the 5 best were chosen as the main derivatives for ADMET. The ADMET results shows high QED values, with good oral and intestinal absorption, excretion and distribution, with goo metabolism and toxicity, except for Der9 to Der12 which are hepatotoxic. Conclusion The derivatives are potential candidates do treat chikungunya, with the exception of Der9 to Der12 due to hepatotoxicity. In particular Der8 is the best candidate as also have in vitro inhibitory action against zika virus. Thus, Der8 is a potential molecule to progress to preclinical assays.	
Corresponding Author:	Pierre Fechine Federal University of Ceara: Universidade Federal do Ceara BRAZIL	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	Federal University of Ceara: Universidade Federal do Ceara	
Corresponding Author's Secondary Institution:		
First Author:	Joan Petrus Oliveira Lima	
First Author Secondary Information:		
Order of Authors:	Joan Petrus Oliveira Lima	

Paper 3: Recent advances in nanostructured materials: A look at the applications in optical chemical sensing

Status: published.

Authors: CARNEIRO, S. V.; OLIVEIRA, J. J. P.; RODRIGUES, V. S. F.; LIMA, J. P. O.; DO NASCIMENTO, J. H. O.; SANTOS-OLIVEIRA, R.; FECHINE, L. M. U. D.; FREIRE, R. M.; FECHINE, P. B. A.

Periodic: Materials Today Nano.

Year: 2023.

Obs: contributed writing in the original draft, section of "Algorithms used in chemical sensors", including AI, statistics, machine learning among other topics related to computational tools.

Materials Today Nano 22 (2023) 100345

Contents lists available at ScienceDirect

Materials Today Nano

journal homepage: https://www.evise.com/profile/#/MTNANO/login

Recent advances in nanostructured materials: A look at the applications in optical chemical sensing

S.V. Carneiro [a], J.J.P. Oliveira [a], V.S.F. Rodrigues [a], J.P.O. Lima [a], J.H.O. do Nascimento [b], R. Santos-Oliveira [c, d], L.M.U.D. Fechine [a], R.M. Freire [e], P.B.A. Fechine [a, *]

[a] Advanced Materials Chemistry Group (GQMat), Department of Analytical Chemistry and Physical Chemistry, Federal University of Ceara – UFC, Campus do Pici, CP 12100, CEP 60451-970 Fortaleza, CE, Brazil
[b] Innovation in Micro and Nanotechnologies Group (GPIMN), Department of Textile Engineering, Federal University of Rio Grande do Norte – UFRN, Center of Technology, Campus Universitário Lagoa Nova, CEP 59078-970, Natal, RN, Brazil
[c] Brazilian Nuclear Energy Commission, Nuclear Engineering Institute, Rio de Janeiro, 21941906, Brazil
[d] Universidade Unigranrio, Laboratory of Advanced Science, Duque de Caxias - RJ, 25071-202, Brazil
[e] Laboratory of Pesticide Residues and Environment, Instituto de Investigaciones Agropecuarias, INIA Centro Regional La Platina, Santiago 8820000, Chile

ARTICLE INFO

Article history:
Received 17 October 2022
Received in revised form
26 March 2023
Accepted 15 April 2023
Available online 20 April 2023

Keywords:
Optical chemical sensor
Fluorescence
Nanomaterials
Polymeric films
Sensor array
Plasmonic nanoparticles
Limit of detection
Dimensionality
Synthesis of nanoparticles

ABSTRACT

Recently, optical chemical sensors have been used to simplify traditional methods of chemical analyses. Indeed, sensing strategies based on the photoluminescence properties of nanomaterials need to be well-understood, since there are several nanomaterials with different features regarding their applicability in the optical sensor field. This review summarizes the main 0D, 1D and 2D nanomaterials used in sensing devices, with a special look in their optical properties and different methods of synthesis, as well as a briefing overview related to their colloidal stability, chemical structure and kinetic and thermodynamic parameters. In addition, the analytical performances of different optical sensing strategies are also highlighted, in which pattern recognition algorithms and the development of polymeric-based films are tools used to contribute for the application of new optical chemical sensors with exceptional sensibility and selectivity. Furthermore, this review would give some insights into advances development of optical chemical sensors, using nanomaterials of different dimensionalities.