PEPOLE'S DEMOCTRATIC REPUBLIC OF ALGERIA
وزارة التعليم العالي و البحث العلمي
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
جامعة عمار ثليجي بالأغواط
UNIVERSITY AMAR TELIDJI OF LAGHOUAT

كلية العلوم
FACULTY OF SCIENCES
قسم البيولوجيا
DEPARTEMENT OF BIOLOGIY

MASTER THESIS

Field: *Biological Sciences*

Option: *Applied Biochemistry*

TITLE

In silico study of some phenolic compounds as potential antiviral agents against the main protease of SARS-CoV-2

Submitted by: KHEMILI Aicha

Committee Members:

President :	Dr. BOUSSOUSSA Hadjer
Examinator :	Dr. NIA Samira
Supervisor :	Pr. BENAROUS Khedidja
Co-Supervisor :	Dr. BOU-SALAH Leila

Promotion: June 2021.

Dedication

This work is dedicated to:

To the spirit of my father Abass,

To my patient mother Fatima Zohra, you are the reason of my success with your care and prayers.

To my awesome sister Asma, who supported me and believed in me since the beginning, I'm lucky to have you.

To my best friend Hachani, who never let me alone, your presence in my life has given me hope and strength to overcome any obstacle in my way.

To my friends, Rihana, Macha, Didi, Emy and Roufaida, who have always been by my side. Thanks a lot for encouraging me at very moment I need you.

And to myself, who strengthened me and taught me how to face the most difficult circumstances and tough times by my own.

Acknowledgments

First and foremost, I would like to express my sincerest gratitude to **Pr. Khedidja BENAROUS** for offering me the chance to pursue my master thesis under her supervision, and to take part in the project of COVID-19 which I am really interested in, and introducing me into the field of Bioinformatics. During my studies, I have received a lot of guidance, support and suggestions from her, which were very helpful for my research. Under her supervision, I have mastered the knowledge and skills to carry out the scientific research.

My infinite appreciation and profound thanks go to my co-supervisor **Dr. Leila BOUSALEH** for sharing her knowledge and time to answer my questions, for her accurate suggestions, sincere supervision, encouragement and brilliant ideas. Thanks a lot for helping me to accomplish this thesis.

Deep thanks to **Dr. Abderahmane LINANI**, who gave me invaluable information, provided the necessary scripts and detailed explanation of the Docking technique.

I would like also to thank **Dr. Fatiha ELHOUITI** for her simple way of teaching, her advice and support during my study path, she was a source of inspiration for me. My thanks also go to the head of the Biology Department **Pr. Rachid CHAIBI** and his group, who were next to us during the academic term.

Finally, I am also thankful to the classmates for their friendliness during these five years.

Abstract

SARS-CoV-2, a rapidly spreading new strain of coronavirus, has affected more than 170 countries and received worldwide attention. Several clinical trials are underway to identify specific drugs for the treatment of this novel virus. The crystallized form of SARS-CoV-2 main protease (Mpro) was demonstrated by a Chinese researcher, which is a therapeutic drug target, its inhibition is necessary for the blockage of the viral replication. The present study aimed to evaluate the efficacy of ten medicinal plant-based phenolic compounds as SARS-CoV-2 Mpro (PDB ID: 6LU7, Resolution 2.16 Å) inhibitors, using a molecular docking with ADMET analysis. The modeling of these polyphenols were performed using Autodock Vina, Audotock tool and Discovery Studio Visualizer programs. ADMET properties were calculated using PreADMET prediction online server. The results show that Lonchocarpol A and 3-Methoxycarpachromene exhibited high binding affinity than others toward SARS-CoV-2 Mpro, and expressed good pharmacokinetic properties. Therefore, they may represent a potential treatment option for COVID-19. Further researches are urgently required to investigate their potential uses in COVID-19 treatment.

Keywords: SARS-CoV-2, COVID-19, polyphenols, Molecular docking, ADMET Analysis, Mpro.

List of abbreviations

RNA:	Ribonucleic acid
mRNA:	Messenger RNA
DNA:	Deoxyribonucleic acid
HCoV:	Humain coronaviruses
CoVs:	Coronaviruses
SARS-CoV:	Severe Acute Respiratory Syndrome-associated coronavirus
MERS-CoV:	Middle East Respiratory Syndrome-associated coronavirus
SARS-CoV-2:	Severe Acute Respiratory Syndrome Coronavirus 2
COVID-19:	Coronavirus Disease 2019
N protein:	Nucleocapsid protein
M protein:	Membrane protein
S protein:	Spike/Surface protein
E protein:	Envelope protein
HE:	Hemagglutinin-Esterase dimer
NSPS:	Non-structural proteins
ORF:	Open Reading Frames
RBD:	Receptor-Binding domain
PD:	Protease domain
ACE2:	Anti-angiotensin convertase
TMPRSS2:	Transmembrane Serine Protease
PLpro:	Papain-like protease
Mpro:	Main protease
IFN-I:	Interferon 1
STAT1:	Signal transducer and activator of transcription 1
IRF3:	Interferon 3 regulatory factor
NF-KB:	Nuclear facor – Kappa B
JACK-STAT:	Jamas Kinases- Signal transducer and activator of transcription
PHB1:	Protein prohibitin 1
PHB2:	Protein prohibitin 2
pp1a:	Polyprotein 1b
pp1b:	Polyprotein 1b
RdRp:	RNA- depandant RNA polymerase
RTC:	Replication-transcription complex
gRNA:	Genomic RNA
sgRNA:	Sub-genomic RNA
NLRP3:	NLR family pyrin domain containing 3
IL-1B:	Interleukin 1 beta
DAMPs:	Damage associated molecular patterns
PAMPs:	Pathogens associated molecular patterns
PRRs:	Patterns recognition receptor
SIRS:	Systematic inflammatory response syndrome
Ig:	Immunoglobulin
AT:	Angiotensin
ARDS:	Acute respiratory distress syndrome
RT-PCR:	Polymerase chain reaction via reverse transc
CDC:	Centers for disease control and prevention
WHO:	World Health Organization
DSV:	Discovery studio visualizer
HIV:	Human immunodeficiency virus 1
HSV-1:	Herpes simplex virus type 1
HCV:	Hepatitis C virus
ADV:	*Autodock vina*
PDB:	Protein data bank

ADT:	Autodock tools
ADMET	Absorption, Distribution, Metabolism, Excretion and Toxicity
HERG	Humain ether related gene channel
BBB	Blood-brain barrier

List of Figures

List of Tables

List of Appendix

Table of contents

GENERAL INTRODUCTION

The epidemic of coronavirus disease 2019 (COVID-19) started from Wuhan, China, December 2019, and have become one of the worst public health crisis in modern history and a major challenging problem for not only China but also countries around the world, it had also severely affected the economic level. The Countries that have been affected most severely are the USA, Brazil, India, Russia, Mexico, South American Countries, and most European Countries. **(Giovanetti et *al*., 2021).**

The symptoms associated with COVID-19 comprise flu-like symptoms including fever, cough, dyspnea, and myalgia. Other COVID-19 complications may include acute liver, cardiac, and kidney injury, as well as secondary infection and inflammatory response. Thereby, the severity of COVID-19 is variable -asymptomatic to fatal- and the lethality appears to be significantly lower than that of SARS and MERS. Generally, COVID-19 can cause respiratory, gastrointestinal, and central nervous system diseases in humans and animals, threatening human's life and causing economic loss **(Hui et *al*., 2020).**

The etiological agent of COVID-19 is the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (**Bordallo et *al*., 2020).** A new human coronavirus that appears to be more readily transmitted from human to human, but animal to human transmission is considered the origin (**Benvenuto et *al*., 2020).** It is a single, positive-strand RNA virus causing severe respiratory syndrome in humans. Depending on the genomic structure and phylogenetic analysis of SARS-CoV-2, the virus ranked among the genera Betacoronavirus, which includes SARS-CoV and MERS-CoV. However, SARS-CoV-2 exhibits differences in its genomic composition that can influence its pathogenesis **(Al-Sadeq et *al*., 2020).** The most used diagnostic method for COVID-19 is real-time PCR to detect its genomic RNA, which has been implemented through rapid sequencing and availability of genomic sequence **(Rosales-Mendoza et *al*., 2020).**

To lessen the impact of the pandemic, several efforts of the scientific community are remarkable to visualize and comprehend the complex biology driving the pandemic through structure-function studies of different SARS-CoV-2 proteins. Furthermore, condensed efforts being put in by researchers to evolve diagnostic, therapeutic, and vaccine candidates. This effort started soon after the publication of the viral genome sequence in January 2020 **(Arya et *al*., 2021),** and are still underway to reuse existing drugs and to design new therapeutic agents targeting various components of the virus **(Kneller et *al*., 2020).**

SARS-CoV-2, similar to many other single-stranded RNA viruses, employs a chymotrypsin-like cysteine protease called 3C-like protease (3CLpro) or main protease (Mpro), required to process polyproteins into mature non-structural proteins such as RNA-dependent RNA polymerase (RdRp) and helicase, which are essential for viral transcription and replication. Moreover, the substrate specificity of 3CLpro is highly conserved among different CoVs, together with the absence of a homologous human protease, making it an ideal target for the development of broad-spectrum antiviral drugs **(Kneller et *al*., 2020).**

According to the World Health Organization (WHO) report, there is an estimated 65 to 80% of the world's population who use traditional medicine for their primary health. Presently, and despite the progress achieved in medicine, the majority of population, especially in developing countries, refer to herbal traditional practices to cure themselves, this is because of their bioactive substances, especially phenolic compounds, have attracted significant attention which could be valorized to develop medications without side-effects.

The aim of the present work is to contribute by using bioinformatics techniques (molecular docking and ADMET tests), determining the potential natural antiviral agents against the main protease of SARS-CoV-2, of ten phenolic compounds from aromatic plants that are largely used in traditional Algerian medicine such as *Origanum Vulgare, Pistacia atlantica, Glycyrrhiza spp, Lonchocarpus spp, Erythrina spp, Citrus spp, Lupinus spp, Sophora tonkinensis Gagnep* and *Macaranga conifera*.

The manuscript is divided into four parts arranged as:

- The first part is Bibliographical review, which describes a brief information from literature about symptoms and pathology of COVID-19, with description of two medicinal plants.
- The second part is materials and methods, which we have defined the different used methods in this *in silico* study.
- The third part is results and discussion, herein; we have presented our results and discussed them with the most recent published papers.
- The last part is conclusion and perspectives.

BIBLIOGRAPHIC REVIEW

I. Coronaviruses

Coronaviruses (CoVs) belong to the Coronaviridae family, Nidovirales order, are actually a family of hundreds of viruses, **(Broadbent et *al*., 2020; Mousavizadeh et Ghasemi, 2020),** that cause illness ranging from the common cold to severe conditions, by affecting the respiratory tract of animals and humans. They can cause disorder not only in the respiratory tract but also in the digestive tract and systemically. They can be zoonotic, that is transmitted from animals to humans **(Khaerunnisa et *al*., 2020**). The name coronavirus is derived from a Latin word "Corona" meaning halo or crown, which refers to the distinctive spikes or crown-like projections with rounded tips that decorate their surface **(Taleghani et Taghipour, 2020).**

Based on the International Committee on Taxonomy of Viruses. The family coronaviridae consists of two subfamilies, Coronavirinae and Torovirinae, that is further classified into four main genera: Alphacoronavirus, Betacoronavirus includes human and mammal's viruses, Gammacoronavirus includes viruses of whales and birds, Deltacoronavirus includes viruses isolated from pigs and birds. Accordingly, there is no agreement yet on the exact taxonomic position of SARS-CoV-2 (**Figure 1),** but most researchers say it looks like Betacoronaviruses **(Alexandar, 2020; Harapan et *al*., 2020; Dhama et *al*., 2020; Pillaiyar et *al*., 2020).**

Coronaviruses are spherical or pleomorphic in shape, enveloped, non-segmented particles contain positive-sense single-stranded RNA (+ssRNA), they have the particularity of having the longest RNA genome among RNA viruses, consisting of 27,000 to 32,000 bases (a genome 100,000 times smaller than the human genome), having diameter of 80–120 nm **(Puttaswamy et *al*., 2020).** Until 2019, the family of coronaviruses comprised six human pathogenic species. Amongst them, four "endemic" species of human coronaviruses are responsible for a large portion of common colds and are responsible for about 10% of seasonal airway diseases not caused by influenza, the HCoV-OC43 (Betacoronaviruses) and -229E (Alphacoronaviruses), known as "classic", were identified in the 1960s. HCoV-NL63 (Alphacoronaviruses) and -HKU1 (Betacoronaviruses) are said to be "new" because they were identified more recently, at the beginning of the 2000s **(Taleghani et Taghipour, 2020).** Although these Human coronavirus (HCoV) strains can also cause more serious diseases of the lower respiratory tract, such as bronchitis, bronchiolitis, and pneumonia, especially in newborns or infants, elderly people and immunocompromised patients, their phenotypes are generally mild, and as a result, these four HCoVs received relatively little attention **(Elrashdy et *al*., 2020).** The other CoVs can cause Severe Acute Respiratory Syndrome-associated coronavirus (SARS-CoV)

(Betacoronaviruses) or Middle East Respiratory Syndrome-associated coronavirus (MERS-CoV) (Betacoronaviruses), which can be life threatening. With Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), a seventh human pathogenic species has been added in December 2019, a new coronavirus strain invaded the globe, causing a severe health disaster, and it is characterized by the most efficient and aggressive transmission **(From Medical News Today wesite, visited on 20-05-2021).**

Today, HCoVs are recognized as one of the fastest growing viruses due to their high rates of genomic nucleotide replacement and recombination. Since there is no effective therapy for these viral infections, developing anti-SARS drugs against future outbreaks remains a formidable challenge **(Jo et *al*., 2020).**

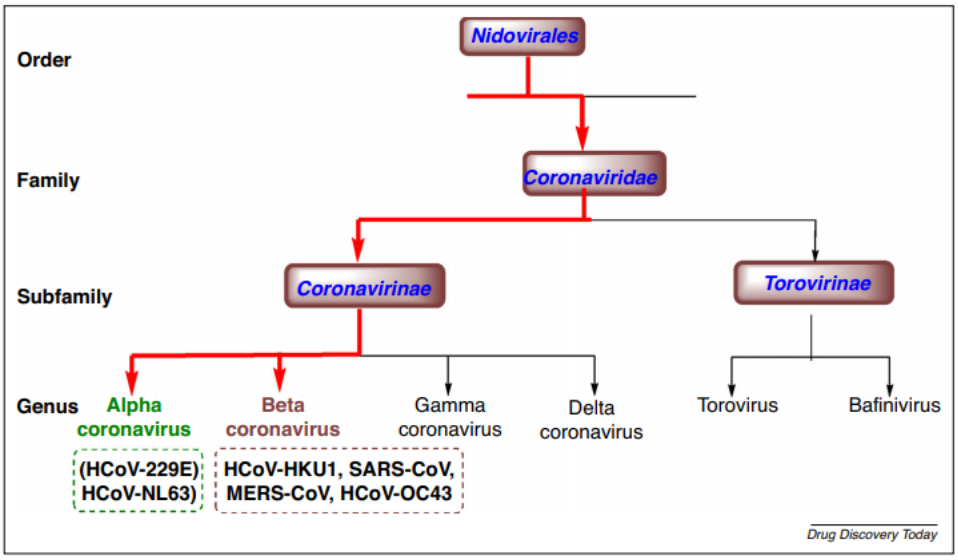


Figure 1. Schematic representation of the taxonomy of Coronaviridae (according to the International Committee on Taxonomy of Viruses) **(Pillaiyar et *al*., 2020).**

I.1. SARS-CoV

SARS is a type of coronavirus. It became a global pandemic in 2002–2003. While it is still unclear exactly where SARS-CoV came from, similar viruses were later found in bats, and some studies suggested the virus could have jumped to humans via an intermediary such as civets **(Juckel et *al*., 2020).**

The SARS-CoV epidemic occurred between November 2002 and July 2003, in Guangdong, China. The virus then spread to more than two dozen countries, during this period, more than 8,000 patients suffered from pneumonia caused by SARS-CoV, a total of 774 deaths as it spread throughout the world with an approximately 10% mortality rate **(Taleghani et Taghipour, 2020).** The infection affects both the upper and lower respiratory tracts, SARS-

CoV begins with a high fever (temperature greater than 100.4°F [>38.0°C]). Other symptoms may include headache, an overall feeling of discomfort, and body aches. After 7–10 days, the person may develop a dry cough, also, pneumonia, a severe lung infection, often develops. As SARS progresses, it can lead to failure of the lungs, liver, or heart **(Li et *al*., 2005).**

The main way that SARS seems to spread is by close person-to-person contact. The virus that causes SARS is thought to be transmitted most readily by respiratory droplets (droplet spread) produced when an infected person coughs or sneezes. In addition, it is possible that the SARS virus might spread more broadly through the air (airborne spread) or by other ways that are not now known. The virus can also be isolated from the stool but the faecal-oral route does not appear to play a significant role in the transmission of the virus **(According to Medical News Today wesite, visited on 20-05-2021).**

I.2. MERS-CoV

The Middle East Respiratory Syndrome MERS is caused by a further coronavirus named MERS-CoV. The zoonotic MERS-CoV virus is excreted by infected camelids (dromedaries) and transferred to humans in a not yet fully understood way **(Taleghani & Taghipour, 2020).** The first cases of this viral disease were detected in April 2012 in Saudi Arabia **(Platto et *al*., 2021**), and with time it spread across the Arabian Peninsula, with focal outbreaks as far away as the USA and North Korea. Most cases were spread from people who traveled to the Middle Eastern countries. So far, approximately. 2500 cases have been confirmed, of which approximately, 858 patients died. It is a virus that is found to be more dangerous than the 2002 SARS-CoV, with an estimated death rate of 34% in humans **(Taleghani & Taghipour, 2020).**

Human-to-human transmission is usually by droplet infection, with the incubation period being an average of 2 and 5 days. MERS-CoV attacks the respiratory tract, it can cause fever and cough, as well as severe pneumonia in the most severe cases. People with cardiovascular or renal insufficiency, and people who are immune-compromised or diabetic, are at greater risk of developing a pneumonia, a severe form of the disease and transition to acute respiratory distress syndrome, diarrhea often occurs as a concomitant symptom. The clinical course of an infection is often asymptomatic in healthy patients or presents mild flu-like symptoms **(Taleghani & Taghipour, 2020).**

I.3. SARS-CoV-2

SARS-CoV-2 taxonomically belongs to the coronavirus family, is considered a novel human-infecting Betacoronavirus **(Harapan et *al*., 2020).** As one of the seven members of the CoV family that infect humans **(Dhama et *al*., 2020).** This virus was first identified in the respiratory tract of patients with pneumonia in Wuhan, Hubei China, in December 2019, which was then indicated as a newly identified betacoronavirus (nCoV), and will later be named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). This virus is broadly distributed in humans and other mammals, causes CoronaVirus Disease 2019 (COVID-19). Subsequent analyzes will show that the virus has already been circulating for several weeks before December, and had appeared all over the world, caused fatal flu, it spread mainly through close contact from person to person or the spilled respirational material (cough, sneeze) of the infected people. Its diameter is approximately 65-125 nm, containing single strands of RNA and provided with crown-shaped tips on the outer surface (**Yüce et *al*., 2020).**

SARS-CoV-2 is a new Betacoronavirus after previously identified SARS-CoV and MERS-CoV which has led to lung failure and potentially fatal respiratory tract infection and caused epidemics mainly in Guangdong, China and Saudi Arabia. SARS-CoV-2, which, as its name suggests, is closely related to SARS-CoV, appears to be far more transmissible, and its mortality rate has so far proven difficult to nail down, but many estimates land at around 2 percent **(Astuti, 2020).**

The morphological structure of SARS-CoV-2 reproduced from ref **(Rosales-Mendoza et *al*., 2020) Figure 2**. The virus is formed by an envelope membrane associated with the following structural proteins: Spike glycoprotein (S), is a transmembrane protein that facilitates the binding of viral envelop to receptors expressed in host cell surfaces, Hemagglutinin-Esterase dimer (HE), which acts as a potent mediator of attachment and destruction of sialic acid receptors on the host cell surface, a Membrane glycoprotein (M), plays a role in determining the shape of the virus envelope, an Envelope glycoprotein (E), which adheres to the M protein to form the viral envelope and plays a role in the production and maturation of this virus. The viral structure also comprises a Nucleocapsid protein (N) that, along with the RNA genome, produces the nucleocapsid.

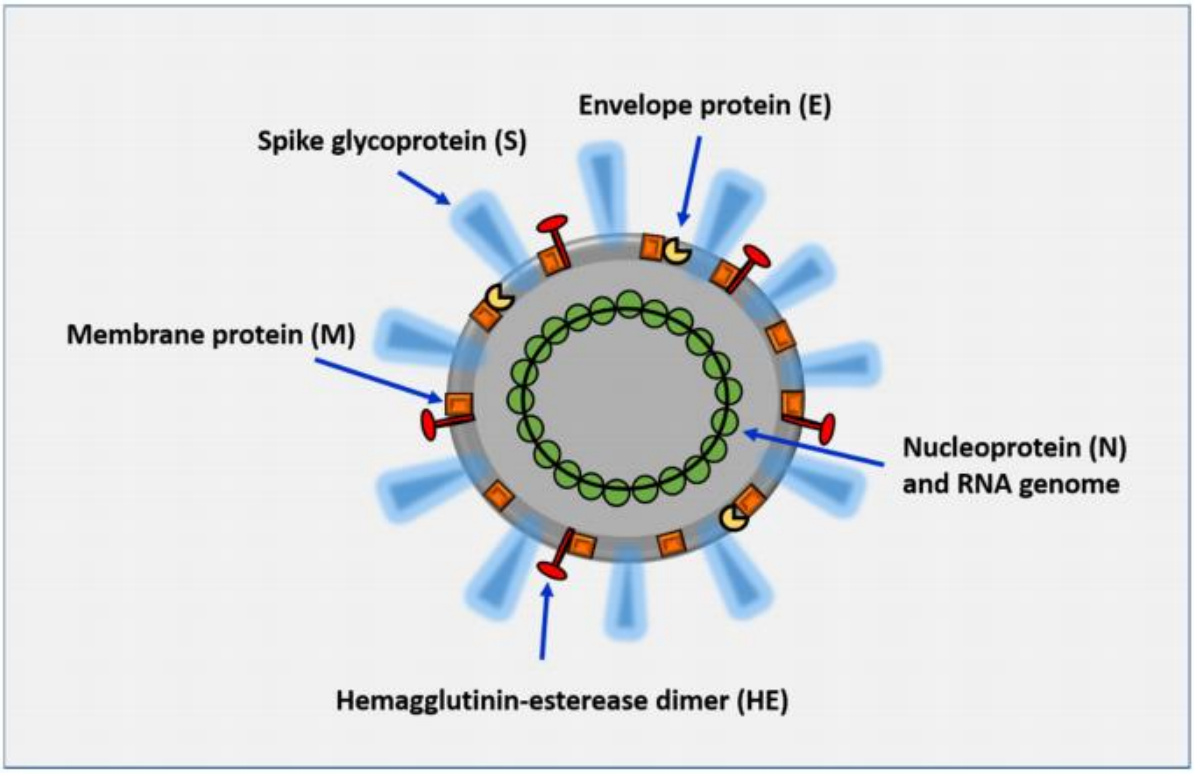


Figure 2. Schematic diagram of a 2D structure illustration of the SARS-CoV-2 virus **(Rosales-Mendoza et *al.*, 2020).**

I.3.1. Structure and morphology of the viral particle

a. Organization of the genome

SARS-CoV-2's genome structure (NCBI Reference Sequence: NC_045512.2) followed the specific gene characteristics to known CoVs, in which there are 14 Open Reading Frames (ORFs), encoding non-structural proteins (NSPs) necessary for virus replication and assembly processes, structural proteins and accessory proteins, despite accessory proteins were considered unnecessary for viral replication in vitro, some of them have been shown to play an important role in virus-host interactions in vivo (**Yüce et *al.*, 2020).**

The genome size of SARS-CoV-2 is around 30kb long, with a G + C content of 38% **(Dhama et *al.*, 2020).** That also serves as mRNA, from which ORF1a and ORF1b are translated, with a 5'-cap, 3'UTR poly (A) tail. The two-third of 5’ contains ORF1ab encoding ORF1ab polyproteins, pp1a and pp1ab protein that are cleaved by proteases it yields 16 NSPs (NSP1-16), for virus replication and assembly processes. While the one-third of 3′ consists of genes encoding eight accessory proteins, encoded by ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8a, ORF8b and ORF10 genes, their functions are still largely unexplored **(Khailany et *al.*, 2020)**. And four structural proteins, currently known as Surface/ Spike glycoprotein (S)

encoded by ORF2, an Envelope protein (E) encoded by ORF4, Membrane/Matrix protein (M) encoded by ORF5, and Nucleocapsid proteins (N) encoded by ORF9 **(Yoshimoto, 2020).** Hemagglutinin-Esterase genes (HE) are also seen intermingled within the structural genes (Dhama et al., 2020). The morphology and complete genome structure of the SARS-CoV-2 are illustrated in **Figure 3** based on the reports by **Giovanetti et *al*., 2021.**

Advanced in-depth genome analysis identified 380 amino acid substitutions between the amino acid sequences of SARS-CoV-2 and the SARS / SARSlike coronaviruses. This difference in amino acid sequence could have contributed to the difference in pathogenic divergence of SARS-CoV-2 **(Dhama et *al*., 2020).**

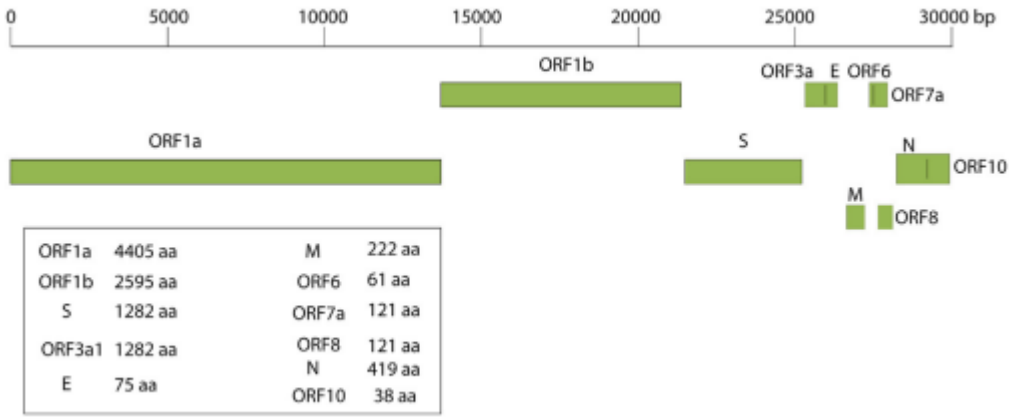


Figure 3. Schematic diagram of a genome architecture of the SARS-CoV-2 virus (**Giovanetti et al., 2021).**

b. Non-Structural proteins (NSPs)

The SARS-CoV-2 genome encodes 16 NSPs participate in transcription, translation, synthesis, processing, modification of RNA, and virus replication but also to elicit the immune response and represent targets to develop future prophylactic and therapeutic approaches against COVID-19. Some data about protein function in SARS-Cov-2 are not yet known **(Wu et *al*., 2020).**

- **NSP1** is the first protein of the polyprotein of SARS-CoV-2, it is the N-terminal cleavage product of polyprotein precursors, pp1a and pp1ab, produced through proteolysis by viral papain-like protease (PLpro), this protein is also known as the leader protein. NSP1 interferes with the host cell protein synthesis by binding to the 40S ribosomal subunit and endonucleolytic cleavage of host mRNA, inactivate translation and promotes host mRNA degradation selectively, while the viral SARS-CoV-2 mRNA remain intact **(Arya et *al*., 2021).**

- **NSP2** is the second protein of the polyprotein of SARS-CoV-2, may play a role in the modulation of host cell survival signaling pathway by interacting with two host proteins: prohibitin 1 (PHB1) and prohibitin 2 (PHB2). Indeed, these two proteins play a role in maintaining the functional integrity of the mitochondria and protecting cells from various stresses, suggest that NSP2 plays a role in disrupting the host cell environment **(Arya et *al*., 2021).**

- **The PLpro**, Papain-Like proteases, proteolytically is responsible for cleaving the viral polyprotein precursors, pp1a and pp1ab, at three sites to produce NSP1, NSP2, and NSP3, which are essential for virus replication. PLpro is a cysteine protease with its active site contains a catalytic triad of cysteine, histidine, and aspartic acid residues **(Arya et *al*., 2021).**

- **NSP4** interacts with **NSP3** and probably host cell proteins to give a role related to membrane rearrangement in SARS-CoV-2. In addition, the interaction between NSP4 and NSP3 is crucial for viral replication **(Yoshimoto, 2020).**

- **NSP5** also known as the main protease (Mpro), or 3-chymotrypsin-like cysteine protease (3CL-pro). It is a 33 kDa cysteine protease, it has the capability to cleave pp1a and pp1ab polyproteins in 11 distinct sites at P1 position, to produce a mature protein which subsequently leads replication-transcription complex (RTC) and release of 12 mature functional NSPs **(Yadav et *al*., 2020).** SARS CoV-2 Mpro is a two protomer homodimer in its active form with one active site per the homodimer chain. Each of the protomers of the Mpro consists of three domains. Domains I (amino acid residues 8–101) and II (amino acid residues 102–184) are mainly antiparallel β-barrel structures, the catalytic site comprising of cysteine (Cys145) and histidine (His41) amino acid moiety and is located at the cleft of domain I and domain II, and it is buried in an active site cavity located on the surface of the protein. Domain III (amino acid residues 201–303) is composed of five α-helices and is responsible for the enzyme dimerization **(Kneller et *al*., 2020).** The cleavage of Mpro generally take place directly after a Gln residue, and the Gln residue is typically preceded by a hydrophobic residue, most often Leu. The residue that follows the Gln is often a small amino acid such as Ser, Ala, or Asn **(Ghosh et *al*., 2020).**

Due to its responsibility in the maturation of itself and key functionalities such as the replicase and the helicase enzymes, and its importance in life cycle of coronavirus controlling the process of regulation, replication and transcription of virions. Furthermore, it is conservative in nature, it is highly conserved throughout coronavirus species, the Mpro from SARS CoV-2

is reported to share more than 96% sequence similarity with the same protease from SARS-CoV-1 and MERS-CoV, with no known human protease having the same cleavage pattern. Makes this enzyme an attractive target for potential antiviral drugs, suggesting that it should be possible to develop inhibitors would block the viral replication by targeting Mpro without off-target toxicity **(Ghosh et *al*., 2020; Abouelela et *al*., 2021).**

- **NSP6** proteins interact to form a complex to induce double-membrane vesicles **(Arya et *al*., 2021).**

- **NSP9** is a single-stranded RNA-binding protein implicated in the virulence of the virus, by acting as RNA binding protein phosphatase **(Arya et *al*., 2021).**

- **NSP10** is a small, single domain protein of 139 amino acids, plays a central role in viral transcription by acting as a scaffold protein to form the methylation complex of mRNA with NSP14 (3'-5 'exoribonuclease) and NSP16 (2 '-O-methyltransferase). NSP10 physically cooperates with NSP14 and NSP16 to improve their activities **(Arya et *al*., 2021).**

- **NSP11** is a short peptide composed of 13 amino acids and the first nine amino acids are identical to the first nine in NSP12. It is overlapping NSP10 but its function is still unknown **(Yoshimoto, 2020).**

- **NSP12, NSP7, and NSP8**. NSP12 is a 103 kDa multi-subunit RNA-dependent RNA polymerase (RdRp), responsible for viral RNA synthesis. NSP12 makes a replicase complex with an NSP7-NSP8 heterodimer and an NSP8 monomer for replication and transcription of the viral RNA genome. The **NSP7-NSP8** (a vital coronavirus replication / transcription complex enzyme "RTC") complex executes the RNA primase activity during the viral RNA synthesis. The complex forms a cylindrical hexadecameric architecture, composed of eight copies of NSP7 and eight copies of NSP8. The hexadecamer central channel has optimal dimensions with a positively charged surface to encircle the negatively charged double-stranded RNA and help RdRp in processing. The binding affinity of NSP12 to template-primer RNA augmented significantly in the presence of both NSP7 and NSP8 and the polymerase activity is improved, NSP12 displays poor processivity in RNA synthesis, that is the presence of NSP7 and NSP8 lowers the dissociation rate of NSP12 from RNA **(Arya et *al*., 2021).**

- **NSP13** is the virus helicase, allowing the duplex RNA to unwind and being accessible. In addition to its helicase activity, NSP13 of SARS-CoV-2 is also known to possess 5′-triphosphatase activity, which is responsible for introducing the 5′-terminal cap of the viral

mRNA. This 5′-terminal cap is the site of recognition for translation and plays a role in splicing, nuclear export, translation, and stability of mRNA **(Yoshimoto, 2020).**

- **NSP14** is an exoribonuclease comprised in the RTC complex. It is implied in proofreading and recombination, thereby lowering the sensitivity of the virus to RNA mutagens **(Arya et *al*., 2021).**

- **NSP15** has been biochemically characterized as an endoribonuclease that cleaves RNA at the 3′-position end of uridylates to generate a 2′-3′ cyclic phosphodiester product. The endonuclease activity of NSP15 helps the virus evade immune system by preventing its detection by the host immune sensing system. It cleaves the 5′-polyuridines from the negative strand of viral RNAs, which are pathogen-associated molecular patterns (PAMPs), major sensors of RNA viral infection. NSP15 activity reduces the accumulation of these PAMP and therefore dampens the host immune response **(Arya et *al*., 2021).**

- **NSP16** (2′-O-Ribose-Methyltransferase) catalyzes the 5'-methyl capping of viral mRNA. The 5'-methyl capping protects viral RNA from degradation by host 5'-exoribonucleases promotes mRNA translation, and prevents the viral RNA from being recognized by innate immunity mechanisms **(Arya et *al*., 2021).**

c. Structural proteins

- **S protein:** The Spike/Surface protein is transmembrane protein with a molecular weight of about 150 kDa found in the outer portion of the virus **(Astuti, 2020).** Forms trimers at the protrusions of the virus, decorate the virion surface like the spikes in a crown and inspired the name for this group of viruses, distinct spikes of 9 to 12 nm **(Dhama et *al*., 2020).** The most vital protein which controls the biological processes such as viral particle attachment, fusion and lastly entry in the host cell **(Li et *al*., 2003),** comprises two functional subunits: in the cascade of the viral entry, S1 domain is responsible for receptor binding as it recognizes and binds to the host receptor through the Receptor-Binding Domain (RBD), and S2 domain is responsible for cell membrane fusion. The RBD of the S1 subunit contains five antiparallel □-strands, whereas □-helical and loop motifs form the connecting entities between the □-sheets. Between two □-sheets, an extended insertion forms the receptor binding motif, which binds to ACE2 at its N-terminal helix **(Peter et Schug., 2021).** The RBD interacts with ACE2 receptor in Protease Domain (PD) of the host human cell, causing viral infection **(Basu et *al*., 2020).**

The SARS-CoV-2 spike protein has a strong binding affinity with human ACE2; the Angiotensin Convertase 2, based on biochemical interaction studies and crystal structure

analysis. The SARS-CoV-2 and SARS-CoV spike proteins share 76.5% identity in amino acid sequences and the more importantly, the SARS-CoV-2 and SARS-CoV spike proteins have a high degree of homology **(Zhang et *al*., 2020).**

- **N protein:** The nucleocapsid known as the N protein is the structural component of CoV, located in the region of the endoplasmic Reticulum-Golgi which is structurally linked to the nucleic acid material of the virus. Because the protein is bound to RNA, it is involved in processes related to the RNA replication, virion formation, and immune evasion. The N protein is also strongly phosphorylated and appears to lead to structural changes improving the affinity for viral RNA **(Astuti, 2020).**

- **M protein:** The Membrane protein is one of the most abundant and well-conserved proteins in the virion structure and plays a role in determining the shape of the virus envelope. This protein promotes the viral assembly and budding of viral particles through interaction with N protein and accessory proteins, by stabilizing N protein-RNA complex, inside the internal virion **(de Oliveira et de Oliveira, 2021).**

- **E protein:** The last component is the envelope protein which is the smallest protein in the SARS-CoV-2 structure that plays a role in the production, maturation and release of virions **(de Oliveira et de Oliveira, 2021).**

- **Hemagglutinin-Esterase dimer (HE):** The HE enzyme is localized in the viral envelope of several viruses, is a marker for the evolution of influenza and the coronavirus. HE acts as a mediator for reversible attachment with O-acetylsialic acids by acting as lectins and a receptor-disrupting enzyme, allowing the fusion of the viral envelope with the plasma membrane of the cell host and destruction of receptors upon infection of the cell **(Yadav et *al*., 2020).**

The S protein, E protein, M protein, N protein, 3CL protease, PL protease, RNA polymerase and helicase protein have been suggested to be viable antiviral drug targets **(Yoshimoto, 2020).**

I.3.2. Replication cycle

Coronaviruses comprise a family of enveloped RNA viruses with positive strand genomes, which means that the expression of most genes, except for late structural genes, is transcribed directly from the genome. The first obstacle for any type of virus is entry into the host cell, using receptors. These receptors not only play a role in adhesion to the host surface

but also facilitate membrane fusion. The attachment of the virus's surface proteins with those of host leads to a chain reaction which activates the mechanisms of entry into the host, at some point during or after viral entry into the host cell, the virus can release its genome to begin to hijack the cellular machinery, allowing its genome to replicate with the eventual production of viral products **(Wong et Saier, 2021).**

SARS-CoV-2 enters the body through the airways, from the nose and mouth. Part of its surface protein (the RBD region of S protein) attaches to the ACE2 cellular receptor, expressed on the surface of cells, thus facilitating viral entry into target cells. This binding convinces the cell that the virion is not a threat allowing the virus entry **(Vijayakumar et *al*., 2020)**. The virus will only infect ACE2 expressing cells, notably type II pneumocytes **(Bordallo et *al*., 2020),** these cells represent 83% of the ACE2-expressing cells in humans, but cells from other tissues and organs, such as human epithelial tissues including the oral and nasal mucosa, nasopharynx, lungs, stomach, small intestine, colon, lymph nodes, thymus, bone marrow, spleen, liver, kidney and even the brain can also express this receptor, which sometimes leads to systemic infection in late pathogenesis of Covid diseases **(Wong et Saier, 2021).** ACE2 is a membrane-bound peptidase, whose main function is to regulate heart rate and lower blood pressure by catalyzing the hydrolysis of angiotensin II to angiotensin (1-7), which acts as a vasodilator. It thus reduces the amount of angiotensin-II and increases angiotensin (1-7) **(Veenma, 2020).** Dipeptidyl peptidase-4 DPP4 is utilized by MERS-CoV, and both of SARS-CoV and SARS-CoV-2 use ACE2 **(Wong et Saier, 2021).**

The process of entry of SARS-CoV-2 into the host cell begins with the attachment of the S protein to the ACE2 receptor. This binding occurs in the protein S RBD, which is divided into two subunits, S1-RBD which bind to ACE2 and S2-RBD which fuses with the cell membrane **(Astuti, 2020).** The sites of the S1 and S2 subunits are cleaved by a Transmembrane Serine Protease called TMPRSS2, and the latter will eliminate ACE2. SARS-CoV-2 can also use the endosomal proteases cathepsin B and L for the priming of the S protein in cells not expressing TMPRSS2 **(Bordallo et *al*., 2020).** After the S-ACE2 glycoprotein interaction and membrane fusion, the virus enters in the cell using the endosomic compartment with the ACE2 receptor, this early endosome becomes a late endosome and fuses with a lysosome to form an endolysosome. At this time, the virus leaves the endolysosome and reaches the cytoplasm, where its viral genome will be translated as seen in **Figure 4 (Bordallo et *al*., 2020).**

The genomic material released by this virus is mRNA that is ready to be translated into protein **(Astuti, 2020).** It hijacks the replication of the cell to form new virus particles that can infect other cells. This happens in four steps; the translation of viral replication machinery, replication of the genome, translation of viral structure proteins and finally, the virion assembly. When the viral RNA is released into the host cell, the viral genome is first unveiled in the cytoplasm, ORF1a and ORF1ab are translated using host cell protein translation machinery to produce pp1a and pp1ab viral replicase polyproteins by contributing a ribosomal frame shifting event. These polyproteins are cleaved by protease enzymes PLpro and a serine type Mpro, which are encoded by ORF1a-b. This results in 16 NSP; NSP1-11 are encoded in ORF1a and NSP12-16 are encoded in ORF1b. These replicase-transcriptase proteins, together with other viral proteins and possibly cellular proteins, assemble into the RNA Replicase /Transcription complex (RdRp). This complex namely replicase-transcriptase complex (RTC) drives the production of negative sense RNA through both replication and transcription. Negative sense RNA intermediates are generated to serve as the templates for the synthesis of positive sense genomic RNA (gRNA) and sub-genomic RNA (sgRNA). The RdRp uses the (+) strand gRNA as a template, which will become the genome of the new virus particle. sgRNAs produced through the transcription are translated into some accessories proteins and viral structural proteins; S proteins, E proteins, M proteins and N proteins, together they form the new viral particles. S, E and M proteins enter the endoplasmic reticulum and the nucleocapsid protein is combined with the (+) strand genomic RNA to become a nucleoprotein complex. In the endoplasmic Golgi apparatus complex, the proteins merge into a complete virus particle, and are excreted from primary cells to extracellular regions through the Golgi apparatus via exocytosis **(Astuti, 2020; Veenma, 2020).** Meanwhile, the stress of viral production on the endoplasmic reticulum eventually leads to apoptosis, or cell death. The mature virions can infect new target cells which results in the production of more virus particles. Thus, the viral replication cycle and progression begin **(Basu et *al*., 2020).**

The interaction between viral S protein and ACE2 on the surface of the host cell is of significant interest since it initiates the infection process. Thus, the binding affinity between protein S and the ACE2 receptor determines the level of viral replication and the severity of the disease **(Liu et *al*., 2020).**

Inside host cells, survival of SARS-CoV-2 is maintained by several strategies to evade the host immune mechanism, NSP1 can interfere with IFN-I responses by several mechanisms such as silencing host translational system, induction of host mRNAs degradation, and

repression of transcription factor STAT1 (Signal Transducer and activator of transcription 1) phosphorylation. NSP3 antagonizes the production of interferon and cytokines by blocking phosphorylation of Interferon 3 Regulatory Factor IRF3 and disrupting the NF-κB (nuclear factor-kappa B) signaling pathway. NSPs 14 and 16 cooperate to form a 5 'viral cap similar to that of the host. Thus, the viral RNA genome is not recognized by cells of the immune system. PLpro also behaves as a deubiquitinase which can deubiquinate certain host cell proteins, including IRF3 and NF-κB, resulting in immune suppression. The accessory proteins ORF3b and ORF6 could disrupt the IFN-I signaling pathway by inhibiting IRF3 and NF-κB-dependent IFNβ expression and blocking the JAK-STAT signaling pathway, respectively **(Liu et *al.*, 2020).**

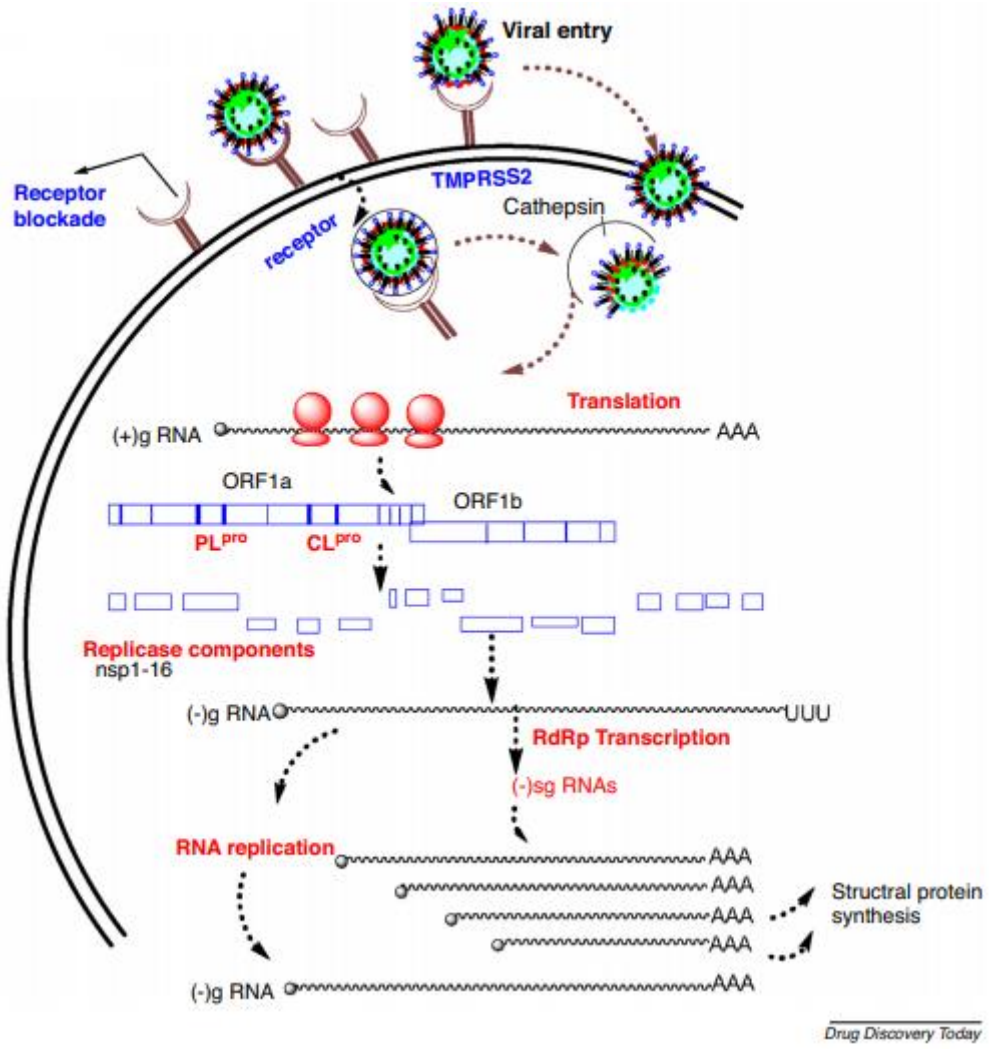


Figure 4. Schematic diagram of a 2D illustration of the SARS-CoV-2's life cycle **(Pillaiyar et *al.*, 2020).**

II. COVID-19

An epidemic of pneumonia, described at the time as viral in appearance of unknown cause, emerged in the city of Wuhan (Hubei province, China) in December 2019. The first confirmed case of COVID-19 outside of China was diagnosed on January 13, 2020 in Bangkok, Thailand **(Di Gennaro et *al*., 2020).** This pneumonia is an infectious disease and globally affected a larger population through human-to-human transmission caused by a virus belonging to the coronavirus family that has not been seen before in humans, currently identified as SARS-CoV-2, it marked the history with third large scale coronavirus epidemic into the human population after SARS-CoV in 2002 and MERS-CoV in 2012. On March 11 2020, World Health Organization (WHO) named this pandemic as a CoronaVirus Disease-19 (COVID-19) officially **(Ghosh et *al*., 2020),** this virus is different from the SARS-CoV responsible for the SARS epidemic in 2002, it is also different from the MERS-CoV responsible for an epidemic evolving since 2012 in the Middle-East. Taken together, both previous outbreaks of other members of the coronavirus family (SARS-CoV and MERS-CoV) did not produce even 1% of the global harm already inflicted by COVID-19.

COVID-19 is a respiratory illness with flu-like symptoms, manifested by a dry cough, fever, severe headache, and fatigue **(Ghosh et *al*., 2020),** People infected with SARS-CoV-2 have a wide range of symptoms ranging from mild respiratory illness to severe respiratory illness with critically ill cases resulting in damage to organ function, such as heart damage, kidney damage acute, hepatic dysfunction and acute respiratory distress syndrome, which can lead to long-term decrease in lung function and arrhythmia; eventually, some critical cases can lead to death. Serious conditions and death have generally been specific to older patients or to patients with weakened immune systems, such as specific heart conditions **(Taleghani et Taghipour, 2020).**

Due to the high reproduction rate of novel SARS-CoV-2 and unavailability of effective drugs and vaccines, the number of reported cases and death are increasing day by day. So far, 127K cases have been confirmed in Algeria. The number of cases worldwide is over 160M and still rising **(last updated 23-05-2021 from worldometer siteweb).**

Owing to the seriousness of the situation, some countries have applied stricter control methods by asking people to stay at home, avoid going out without a reason, take precautions to protect themselves against the disease. These precautions may include also getting vaccinated, avoiding crowds, using disinfectants, canceling travel plans and wearing face masks

which are suitable for that particular time **(WHO, 2020),** to combat the spread of COVID-19 and saving their citizens from the pandemic. At the same time research institutes, drug corporations, biotechnology institutes, research groups in different universities all over the world are racing and still ongoing to develop effective drugs or potential vaccines for COVI-19 **(Rosales-Mendoza et *al*., 2020; Taleghani et Taghipour, 2020).**

Among all the treatment strategies, the uses of already known viral medications have great advantages as the pharmacokinetics, pharmacodynamics, and safety profiles of these drugs have already been established and they can have the potential to save lives and to stabilize the situation. This drug repurposing strategy has gained significance as they are expected to be faster with less investment **(Das et *al*., 2020).** Although few antiviral strategies are being used to treat patients, lack of specific antiviral drugs or vaccines against SARS-CoV-2 is further aggravating the situation. Thus, there is an urgent need to identify and develop effective antivirals against SARS-CoV-2 to fight this deadly virus **(Adem et *al*., 2020).** To date no specific medications have been developed, and therefore considering the risk factors associated with this disease, there is an urgent need for a treatment method to treat it in order to limit the transmission **(Mpiana et *al*., 2020).**

II.1. Origin of the COVID-19 pandemic

Pandemics are not new and have occurred at different stages in human history. While there have been many outbreaks and human catastrophes, there has been a notable rise in the frequency of pandemics from the year 2000. This is particularly due to increased emergence of viral disease amongst animals. The COVID-19 epidemic was predictable, because of several reasons; the first summer months of 2019, flu cases in China had topped the total in the previous four years, with around 270 deaths. And an unusually high number of abnormal interstitial pneumonia was observed in many Italian and French hospitals before the end of 2019. Retrospective analyzes of some of the patients admitted to the hospitals of Colmar and Mulhouse (France) had revealed cases with radiological features typical of the COVID-19 disease **(Platto et *al*., 2021).**

Links between the clue cases and the city's South China Seafood Market were noted. The first cases of COVID-19 were indeed all identified in people who attended this market in the days preceding the appearance of signs of their infection. At this market a large range of alive and freshly slaughtered animals were sold, such as rats, poultry, marmots, bats and snakes. The restaurants of this market are famous for serving different types of wild animals for human

consumption. This might be the point where zoonotic (animal-to-human) transmission might have occurred. Previously the animals present in the live-animal market were identified to be the intermediate hosts of the SARS outbreak in China. Several wildlife species were found to harbor potentially evolving CoV strains that can overcome the species barrier **(Dhama et *al*., 2020).** However, even today, the origin of the SARS-COV-2 virus has not been clearly identified, many studies have demonstrated that bats are suspected as the key reservoir of the viruses by finding as much as 96.2% identical genome sequencing of SARS-CoV-2 with bat CoV RaTG13 **(Astuti, 2020).** Although, others say that bats were hibernating when the virus emerged and attribute it to pangolins **(Veenma, 2020).** In this type of market, live animals are crowded together in cages in non-existent hygienic conditions with dispersal of dangerous excrement. Viruses (SARS-CoV-2 in this case) could in this environment modify their genome, for example generating variants better suited to humans, thus increasing their power of propagation, first from animals to animals, animals to human, then from human to human **(Platto et *al*., 2021).**

SARS-CoV-2 is genetically closer to viruses that infect bats than to MERS-CoV or SARS-CoV with 96.2% similarity (SARSr-CoV; RaTG13). Thus, a Chinese horseshoe bats (Rhinolophus sinicus) being the most probable origin. In contrast, the SARS-CoV-2 genome is less similar to the genomes of SARS-CoV (about 79%) or MERS-CoV (about 50%) **(Al-Sadeq, et Nasrallah, 2020).** Nevertheless, since no epidemic linked to direct transmission from bats to humans has been demonstrated to date, researchers believe it is likely that transmission to humans must rather take place via an intermediate host species in which the viruses can evolve and then be selected into forms capable of infecting human cells. In order to identify this intermediate species, the phylogenetic relationships between the new virus and those of animal species living near the area of emergence are generally examined, it is this method which established that the civet was probably the intermediate host of SARS-CoV in the early 2000s, and the dromedary that of MERS-CoV ten years later. The question of the intermediate host in the transmission of coronaviruses from bats to humans is not resolved and this intermediate animal is not identified with certainty, but the pangolin—better known as scaly anteaters— is suspected, it was initially identified as a carrier of a CoV close to SARS-CoV-2, however several elements cast doubt on this possibility, in particular because the genetic sequences of the virus responsible for the current epidemic and those of the coronavirus which infects the pangolin retain significant differences **(According to Infection Control today, consulted on 01-06-2021).** Initially, the Pangolin origin of COVID-19 in humans had received significant

favor, particularly based on the finding that the virus of pangolins could use ACE2 as the receptor in the host cells. However, more recent results of various experimental approaches, show that the genome of CoVs infecting pangolins of a short genetic sequence encoding the recognition domain of the ACE-2 receptor, related to that which allows SARS-CoV-2 to enter human cells, has at one time suggested that we had a possible intermediate host, but the rest of its genome is too distant from SARS-CoV-2 to be a direct ancestor. The question of the intermediate host in the transmission of SARS-CoV-2 from bats to human is thus not settled, and it appears possible that the transmission occurs directly **(Platto et *al*., 2021; Giovanetti et *al*., 2021).**

However, a later phylogenetic analysis by **Damas et *al*., 2020** cast doubts on the hypothesis of pangolins as SARS-CoV-2 intermediate hosts. Therefore, it seems very likely that the previous detection of CoV in pangolins might have been due to their exposure to infected humans, or other wildlife in the wildlife trade network. Thus, also in the case of SARS-CoV-2 no intermediate host has thus so far been identified, and the possibility of a direct spillover of the virus from bats to humans seems a reasonable option **(Platto et *al*., 2021).**

It is therefore crucial to understand how this virus crossed the species barrier and became highly transmissible from human to human. The study of the evolutionary mechanisms and molecular processes involved in the source and the emergence of this pandemic virus is essential in order to better protect ourselves from the potential emergence of these viruses, to develop therapeutic and vaccine strategies, and to mitigate the risk of future outbreaks **(Harapan et *al*., 2020).**

II.2. Transmission mode

After January 1, less than 10% of patients had market exposure and more than 70% had no market exposure **(Harapan et *al*., 2020).** Therefore, human-to-human transmission is believed. Thus, SARS-CoV-2 spreads primarily through human-to-human transmission, but there is evidence of transmission between humans and animals. Several animals like mink, dogs, domestic cats, lions, tigers and raccoon dogs have tested positive for SARS-CoV-2 after contact with infected humans **(Jo et *al*., 2020).**

The first transmission to humans was in Wuhan, China. Since then, the virus has mostly spread through person-to-person contact. SARS-CoV-2 is transmitted from an infected person to an uninfected person by two main routes:

- Direct contact usually within 1 meter, with the infected person or by indirect contact through infected surfaces in the environment or objects used by an infected person **(Harapan et *al*., 2020).**

- Aerial (or airborne) transmission of the virus via droplets or an aerosol emitted by the infected person, Droplets in different sizes are emitted from the mouth and nose, humans can produce respiratory droplets in several ways, when talk, shout, sing, cough or sneeze. Aerosols are suspensions of smaller particles (a few nanometers to 100 µm), like the vapor produced by our breathing in cold weather. Naturally produced droplets by humans consist mostly of water, but also include many cell types, like epithelial and immune cells, physiological electrolytes, like Na+ and K+, and various infectious agents. Concretely, in the absence of a mask, an infected person emits droplets loaded with viruses, the largest of which are deposited by gravity on surfaces in the immediate vicinity. A healthy person can then become infected by touching the contaminated area with their hands and then bringing them to their mouth, nose or eyes **(Harapan et *al*., 2020).**

- The virus can persist for several hours on a contaminated inert surface. The duration of its persistence varies depending on the nature of the surface, the surrounding temperature, humidity and light conditions. But that's not all, the smaller the diameter of the droplets emitted by the infected person, the more these droplets can be carried away by the ambient air, and remain in suspension there. The virus can thus accumulate in the indoor air of a poorly ventilated room and lead to its airborne transmission **(Harapan et *al*., 2020).**

New insights indicate that SARS-CoV-2 not only spreads through respiratory droplets (direct contact) or environmental contact (indirect contact), but also through the fecal-oral route. Several case studies have reported gastrointestinal symptoms and evidence of viral RNA or live infectious virus present in faces **(Meng et *al*., 2020).**

The disease is most contagious when a person’s symptoms are at their peak. However, it is possible for someone without symptoms to spread the virus. A new study suggests that 10% of infections are from people exhibiting no symptoms. In addition, it remains to be proven whether convalescent patients are a potential source of transmission. The most common way that this illness spreads is through close contact with someone who has the infection **(Meng et *al*., 2020).**

The COVID-19 virus is spreading like wildfire worldwide, affecting almost every country in the world. However, evidence to date suggests that two groups of people are at a higher risk of getting severe COVID-19 disease. These are older people, and those with underlying medical conditions. WHO emphasizes that all must protect themselves from COVID-19 in order to protect others. Current observations suggest that people of all ages are generally susceptible to this new infectious disease, another interesting study, recorded that SARS-CoV-2 was shown to favorably infect older adult males with rare cases reported in children **(Harapan et *al*., 2020).** However, those who are in close contact with patients with symptomatic and asymptomatic COVID-19, including health care workers and other patients in the hospital, are at higher risk of SARS-CoV-2 infection **(Veenma, 2020).**

Thus, the large capacity of the virus to multiply (For example, we find many in nasal secretions: a thimble of secretion would contain hundreds of billions of viruses!), and its mode of contamination explain the rapid spread of the disease in the population **(Veenma, 2020).**

II.3. The SARS-CoV-2 infection

The virus enters the body by inhaling contaminating particles, mainly droplets and aerosols from infected hosts, first lodging in the upper respiratory tract and then reaching the lungs **(Bordallo et *al*., 2020).** Regardless of the type of coronavirus, immune cells such as mast cells that are present in the submucosa of the respiratory tract and nasal cavity are considered the main barrier against this virus **(Dhama et *al*., 2020).** The virus penetrates the cell and hijacks the host's replication system, the virus is replicated and released to infect other cells. In the progress the cell is damaged and undergoes pyroptosis, Pyroptosis is a pro inflammatory cell death process induced by the assembly of a multiprotein complex called inflammasome. The inflammasome activation is a well-known mechanism of tissue damage related to viral infection. It can be triggered by endoplasmic reticulum stress response or by ion influx through viroporins. Viroporin are viral proteins that undergo oligomerization forming ions channel and have been related to many functions in multiple stages of viral life cycle and enhancement of pathogenic effects. CoVs encodes two viroporins: E and 3a. SARS-CoV-2 viroporin E acts as a Ca2+ selective ion channel that also activates the nucleotide binding domain, leucine-rich repeat and pyrin domain-containing protein 3 (NLRP3) inflammasome. Viroporin 3a forms a homotetramer complex that work as an ion channel to promote virus release. The viroporin 3a ion channel leads to K+ efflux and mitochondrial reactive oxygen species production that activate NLRP3 inflammasome. Thus, both viroporins are responsible for inflammasome

activation, with subsequent release of IL-1ß and pyroptosis. Inflammasome promotes the proteolytic cleavage of pro-IL-1β into the active form IL-1β, a pro-inflammatory cytokine, as well as cleavage of Gasdermin-D into Gasdermin N, forming pores and inducing the inflammatory cell death. The death of the infected cell releases cellular and viral fragments known as damage associated molecular patterns DAMPs and pathogens associates molecular patterns PAMPs, they are released and recognized by neighboring innate immunity cells such as macrophages, monocytes and dendritic cells (endothelial, epithelial-cells and alveolar macrophages) through their pattern recognition receptor PRRs, that excrete IL-6, MCP1, IP10, MIP1β and MIP1α, which attract T cells, monocytes and macrophages, to the site of the infection, which will also produce these pro-inflammatory cytokines and chemokines and cause a positive feedback loop **(Bordallo et *al*., 2020; Veenma, 2020)**

From there it can go two ways. In the dysfunctional immune response, the immune cells will start to accumulate in the lungs causing an overproduction of pro-inflammatory cytokines and damage the lung infrastructure, furthermore the inflammation can spread through the body causing an even more severe inflammation. The infection will not be cleared by macrophages and T cells causing a cytokine storm, this cytokine storm can travel through the blood to other organs, typically in the cardiac, hepatic and renal systems. This will lead to Systemic Inflammatory Response Syndrome (SIRS) which is an exaggerated defense response of the body in response to the virus. In the healthy immune response, the virus will be cleared through T cells and macrophages which are recruited by dendritic cells to the infection, before the virus spreads. B cells differentiate into plasma cells and produce antibodies that are able to neutralize and contain viral spreading **(Bordallo et *al*., 2020).** Alveolar macrophages recognize neutralized viruses and apoptotic cells and clear them by phagocytosis. These processes lead to clearance of the virus and minimal lung damage, ultimately resulting in recovery **(Veenma, 2020)**. Most infected people produce antibodies Ig directed against a protein present on the surface of the virus, protein S. The production of M-type (IgM) and G-type (IgG) immunoglobulin begins after the first week and peaks between the 2nd and 3rd week after infection. Then, the level of these antibodies seems to decrease over time, but it is not yet possible to say how long their presence persists **(Anka et *al*., 2020).**

ACE2 involvement with coronavirus infection is of further interest since ACE2 is a potent negative regulator restraining over activation of the renin-angiotensin system (RAS) that may be involved in elicitation of inflammatory lung disease in addition to its well-known role in regulation of blood pressure and balance of body fluid and electrolytes. It catalyzes degradation

of angiotensin II to angiotensin "AT" (1−7). The balance between angiotensin II and angiotensin (1−7) is critical since angiotensin II binds to angiotensin AT1 receptor to cause vasoconstriction, whereas angiotensin (1−7) elicits vasodilation mediated by AT2. Although the notion that ACE2 mediates coronavirus invasion is largely accepted, it remains unclear how the levels or activities of ACE2, AT1 receptors, and AT2 receptors are altered in coronavirus induced diseases due to the limited number of studies **(Anka et *al*., 2020; Liu et *al*., 2020).**

II.4. Incubation

The incubation period of a virus is the time between the entry of the virus into the body and the development of the disease, that is the onset of the first symptoms. For COVID-19, the WHO estimates that this incubation period varies between 1 and 12 days, with an average time of 3 to 5 days. A relatively long period therefore, which raises another problem and it is infection during the incubation period. During this period, the subject can be contagious: he can carry the virus before symptoms appear or when weak signals appear. This is why barrier measures must always be applied (cough or sneeze into your elbow or into a handkerchief, greet without shaking hands or hugs, etc.) even in the absence of symptoms, in order to avoid contaminating your patient entourage **(Yesudhas et *al*., 2020).**

II.5. Symptoms

The majority of patients with COVID-19 represent relatively mild cases. According to recent studies and data from the National Health Commission of China, the proportion of severe cases among all patients with COVID-19 in China was around 15% to 25%.

The symptoms of COV infection vary between animals, birds, and humans, but have similarities with SARS-CoV **(Harapan et *al*., 2020).** Major symptoms of COVID-19 include fever or feeling of fever, fatigue, cough-like respiratory signs, sore throat, headache, soreness, shortness of breath, diarrhea (at least 3 soft stools during the day), sudden loss of smell (without nasal obstruction) and disappearance of taste. People with COVID-19 have had a wide range of symptoms reported, ranging from mild symptoms to severe illness. In the most severe forms, when the virus spreads from the upper respiratory tract into the lungs to cause viral pneumonia and lung damage leading to Acute Respiratory Distress Syndrome (ARDS), this impairs the body's ability to maintain critical levels of oxygen in the blood stream, that is an inability of the lungs to perform gas exchange, which can cause multiple body systems to fail and can be fatal: this constitutes an emergency that requires immediate hospital management (**Anka et *al*., 2020).**

It should be noted that some patients with Covid-19 do not have symptoms of the disease (they are called “asymptomatic”). This shows the importance of carrying out barrier gestures, since a person can be infected with SARS-CoV-2 without knowing it and transmit the disease.

A study based on the follow-up of patients hospitalized following infection with SARS-CoV-2 shows that a quarter of them still have three or more symptoms six months after infection and that 60% are still affected by at least one symptom. But this phenomenon also concerns people who have developed mild forms of COVID-19. The mechanisms behind this persistence of symptoms remain to be clarified (**Anka et *al*., 2020).**

II.6. Diagnosis

Currently, the main test used to diagnose an infection by SARS-CoV-2 is a technique called "RT-PCR", for Polymerase Chain Reaction via Reverse Transcriptase. "RT-PCR" is a virological test also called a molecular test, is used here to reveal the presence of viral genetic material in samples collected by healthcare professionals from the nose or throat of patients (techniques using "saliva" samples are also possible), by transcribing the viral RNA into DNA and then targeting the DNA for amplification and replication until it can be detected by the detectors. **Figure 5 (A).** This test is used to determine if a person is a carrier of the virus at the time of the test, if the patient currently infected with the virus that causes COVID-19 and it give a very accurate results **(Taleghani et *al*., 2020; Yüce et *al*., 2020).**

Second type of test more and more used, “the antigen” test. Here, it is a matter of detecting the presence of certain specific viral proteins in nasal swabs taken from the patient. A positive antigen test result is considered accurate when instructions are carefully followed, but there's an increased chance of false-negative results; meaning it's possible to be infected with the virus but have a negative result. It is advisable, to carry out an “RT-PCR” test in order to confirm the result, the antigenic tests being considered as less reliable than the latter **(Taleghani et *al*., 2020; Yüce et *al*., 2020).**

Finally, we can note the existence of a "serological" test also known as "antibody" testing, conducted from blood samples. Its principle is based on the detection of antibodies circulating in the blood of patients, works based on the principle of the interaction of the antibody immobilized on a support and the antibody present in the sample (blood), either on a lateral flow platform or in ELISA. **Figure 5 (B).** These antibodies are produced by immune cells when infected with the virus in order to destroy it, such as IgM and IgG, the antibody amounts produced on the first a few days of the infection may be insufficient for the detection. Therefore,

they can persist in the blood for a long time, so they are mostly indicative of a past infection. This test should not be used to diagnose a current infection with the virus that causes COVID-19, this latter information may be important in identifying people potentially immune to SARS-CoV-2. But there's a lack of evidence on whether having antibodies means the patient protected against reinfection with COVID-19 **(Taleghani et *al*., 2020; Yüce et *al*., 2020).**

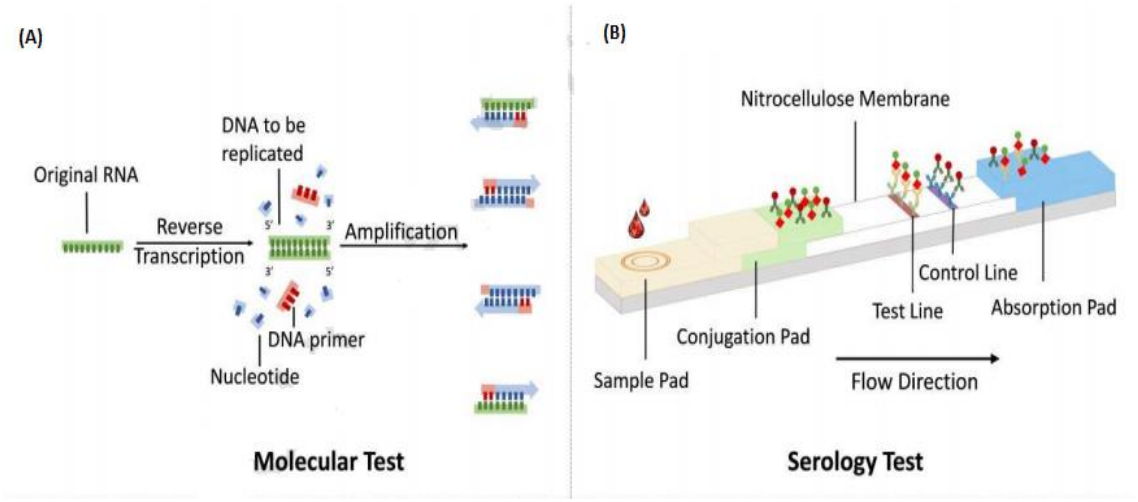


Figure 5. Schematic diagram of the main methods for diagnosing the presence of the SARS-COV-2 virus from COVID-19 **(Taleghani et al., 2020).**

II.7. Treatments and prevention

Several scientists have rushed to find possible treatments to save lives and produce vaccines for future prevention, although the vaccine is now widely available around the world, the development of antivirals, targeting specific proteins involved in the pathogenesis of the disease COVID-19, may be feasible longer term **(Wong et Saier, 2021).**

Vaccination against COVID-19 will be an important tool to help stop the pandemic. Getting vaccinated and continuing to follow prevention recommendations will provide the best protection against COVID-19 for everyone. The SARS-CoV-2 vaccines used went through all the required phases of testing, review and evaluation faster than a typical drug or vaccine, as researchers came together like never before to focus, cooperate and share resources to create a vaccine. This type of large-scale effort helped to make the various test phases more efficient. They are already familiar with other coronaviruses like SARS-CoV-2, and these COVID-19 vaccines were developed based on decades of existing research on other coronaviruses. But there is no way to know how COVID-19 will affect you and transmission prevention measures

and standard recommendations given by various trusted organizations to prevent the spread of infection include must be adopted, as COVID-19 can be easily spread from person to person. The CDC (Centers for Disease Control and Prevention) recommends that people wash their hands regularly and thoroughly because hands touch many surfaces and can catch viruses, wear masks in public places where it is difficult to maintain physical distance. Cover the nose and mouth with a tissue or flexed elbow when sneezing and coughing, this will help slow the spread of the virus in people who are unaware they have contracted it, including those who are asymptomatic. Avoid consuming undercooked meat or eggs, get a COVID-19 vaccine when it is available to you. Air travel should be limited unless serious medical attention is required, the implementation of a temperature control or a scan is mandatory at the airport and at the border to identify suspected cases **(Harapan et *al*., 2020).**

The rapid development of effective drugs for COVID-19 therapy is a difficult task because the process of developing conventional drugs usually takes a long time and costs billions, which is why researchers have used drugs for other viruses like SARS-CoV-1, MERS-CoV and HIV as well as antimalarial drugs were inspected for activity against SARS-CoV-2 **(Abouelela et al., 2021)**. Including Remdesivir (designed for Ebola virus), Lopinavir / Ritonavir (designed for HIV) has inhibitory activity on SARS-CoV-2 Mpro **(Yadav et *al*., 2020),** chloroquine and hydroxychloroquine (designed for malaria) and Tocilizumab (designed for rheumatoid arthritis), in addition, other potential drugs from existing antiviral agents have also been proposed **(Gimeno et *al*., 2020; Harapan et *al*., 2020).** Nevertheless, it is important to point out that despite the high level of similarity of SARS-CoV-2 to other members of the coronavirus family, their binding sites have differences in shape and size, which means that reuse SARS medication may not be successful, and improved sampling should be considered **(Gimeno et *al*., 2020).** At present, there is insufficient evidence to suggest that these antivirals are safe and effective treatments for COVID 19, and most of the evidence suggests that at safe doses they are not effective.

III. Studied Plants

Plants have always been the primary choice for preventing and treating various diseases faced by human beings, they contain specific or broad-spectrum active compounds for almost any type of disease. People living in Algeria have also benefited from plants in the prevention and treatment of various diseases for centuries. People living in rural areas still have an especially rich medicinal plant repertoire.

In addition, there is still no effective drug therapy for SARS-CoV-2. Natural compounds have been consumed since ancient times because they exhibit less toxicity, low cost availability, minimal side effects and are rich in therapeutic resources **(Singh et *al*., 2020).** Natural products or phytochemicals can serve as prophylactic agents, halt virus progression, inhibit inflammatory cytokines secretion, and reduce infection, complications and mortality of SARS-CoV-2. They are of diversified ranges, including essential oils, polyphenols, flavonoids, have been a valuable source of therapeutic agents, molecules with therapeutic potentials, and an important source of more efficient drugs that are based on the chemical structure of natural products, and that can be used in the prevention and treatment of different diseases starting from cancers to different viral infections **(Attia et *al*., 2021; Umar et *al*., 2021).** For example, flavonoids have shown significant antiviral activities, specifically apigenin, quercetin, amentoflavone, daidzein, puerarin, epigallocatechin, epigallocatechin gallate, gallocatechin gallate and kaempferol were reported to inhibit the proteolytic activity of SARS-CoV 3CLpro **(Jo et *al*., 2020).** Curcumin and luteolin also show therapeutic potential against HIV targeting viral protease and HIV-1 trans-activator of transcription. Kaempferol also exhibits anti HSV-1 **(Abouelela et *al*., 2021),** quercetagetin showed strong HCV replication inhibitory activity in vitro, green tea polyphenols can inhibit the endonuclease activity of influenza A virus RNA polymerase. Therefore, the existing scientific evidence strongly suggests that natural flavonoids/polyphenol can act against SARS-CoV2 **(Singh et *al*., 2020).**

Therefore, in this study, we focused on plants whose successful have several types of activities in some diseases, and there is no docking results of these ligands with this chosen target, **Table 1.**

Table 1. The natural sources of some studied phenolic compounds

Molecule	Source	References
4-[[(2’ ,5’ -Dihydroxybenzoyl)oxy]methyl]phenyl O-b-D-glucopyranoside	*Origanum vulgare*	Zhang et *al*., 2014
4-[[(3’ ,4’ -Dihydroxybenzoyl)oxy]methyl] phenyl O-b-D-[6-O- (3’’,5’’-dimethoxyl-4’’-hydroxybenzoyl)] glucopyranoside		
Acacetin 7-O-[4’’’-O-acetyl-b-D-apiofuransyl-(1=>3)]-b-D-xylopy ranoside		
Acacetin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside		
Apigenin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside		
Acacetin-7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>2)]-bD-glucopyranoside		
Lonchocarpol A	-Bark samples of Lonchocarpus and Erythrina	Salvatore et *al*., 1998
	-Bark and root of Citrus, Lupinus and Sophora	
	-Roots and rhizomes of *Sophora tonkinensis Gagnep*	Ahn et *al*., 2019
	-Leaves of *Macaranga conifera*	Jang et *al*., 2002
licochalcone A	*Glycyrrhiza spp*	de Freitas et *al*., 2020
licochalcone C	*Glycyrrhiza spp*	de Freitas et *al*., 2020
3-Methoxycarpachromene	*Pistacia atlantica*	Maamri et *al*., 2021

III.1. *Origanum vulgare*

III.1.1. Botanical description

Origanum vulgare (Oregano, wild marjoram) is a woody perennial plant, with a self-supporting growth habit, it is a photoautotroph. Growing 20–80 cm tall, with opposite leaves 1–4 cm long, produced in erect spikes. with very fragrant flowers when crumpled and they are

purple in color, 3–4 mm long, produced in erect spikes in summer. It is thus recognizable by its odor and its spicy and hot flavor **(Jan et *al*., 2020).**

Oregano is a plant with erect stems, usually hairy, sometimes hairless. They bear leaves with entire or toothed edges (up to 30 pairs per stem), usually oval and blunt-tipped, they are hairy or glabrous and carry invisible sessile secretory glands (up to 800 per cm2). The flowers are grouped in inflorescences or spikes. Each flower is located in the axil of an oval, slightly membranous, glabrous or sometimes pubescent, red-purple or sometimes glaucous bract. The bract is longer than the calyx of the flower. Inside the 3-4 mm long calyx is the corolla (4-10mm long) which is pink or purple in color **(Krim et *al*., 2021).**

Figure 6. Pictures of *Origanum vulgare* leaves (A) and Floral part (B) **(Telabotanica, visited on 23-06-2021).**

III.1.2. Taxonomy

According to Tela Botanica Platform (https://www.tela-botanica.org/) (2016), the systematics of *Origanum vulgare* is as follows:

Domain :	Eukaryota
Kindgom :	Plantae
Phylum :	Spermaphyta
Sub-phylum :	Angiospermae
Class :	Dicotyledonae
Order :	Lamiales
Family :	Lamiaceae
Genus :	Origanum
Species :	*Origanum vulgare*

III.1.3. Geographical location

O. vulgare is the most widespread among all the species of the genus distributed in the Mediterranean regions, in particular Morocco, Tunisia, Algeria, Egypt, Spain and Portugal **(Ligaj et *al*., 2021).** It is also found in culture in Cuba and Reunion Island, but the Mediterranean is its largest range **(Bendifallah et *al*., 2015).** This is the only species of genus Origanum which is found in India. It grows profusely on mountainous and hilly areas with extensive ranges of altitudes **(Jan et *al*., 2020).**

● Distribution limit

Figure 7. Distribution of the genus *Origanum* in the world **(Bouras & Hachemi, 2019).**

III.1.4. Traditional use

O. vulgare have an aromatic thyme-like flavor, their volatile constituents notably contribute to the aroma and flavor of the herb **(Zhang et *al*., 2014).** The leaves before blooming have been used to flavor foods, as a spice for flavoring the dishes in Mediterranean cooking, it has been employed also to flavor alcohol and beer **(Krim et *al*., 2021; Coccimiglio et *al*., 2016).**

The herb having a long history of use in gastronomy and traditional medicine, it is among the most important medicinal plants in Algeria which grow to the spontaneous state **(Hanganu et *al*., 2020)**, used as ethnopharmacological drug to treat many diseases including cold, respiratory problems and digestive disorders through decoctions and infusions **(Krim et *al*., 2021).** The plant is known for its carminative, stomachic, emmenagogue, and expectorant effects to treat cramps, flatulence, coughs, or menstrual problems **(Oniga et *al*., 2018).**

The Origanum oil is applied in chronic rheumatism, tooth ache, and ear ache, it is also being used in veterinary liniments and also stimulates the growth of hair. In homoeopathy, it has been used for hysteric condition. It is also being used in healing lotions for wounds along with other herbs as an external application **(Rao et *al*., 2011).**

III.1.5. Chemical and biological studies of *Origanum vulgare*

Previously, various groups from different regions have worked on this plant and have reported a variety of compounds, as phenolic acids, including rosmarinic acid and caffeic acid, flavonoids, tocopherols, and essential oils such as carvacrol, thymol, and phytosterols, rom different parts of the plant. There are 2951 articles published in Science Direct platform (last update 22-06-2021), which prove its several activities; the antimicrobial activity of *O. vulgare* is due to its high content of volatile oils, the phenolic compounds including flavonoids and phenolic acids are responsible for its antioxidant activity. Moreover, these phenolic antioxidants possess diverse biological activities, for instance, anti-ulcer, anti-inflammatory, antidiabetic, antiviral, cytotoxic and antitumor **(Zhang et *al*., 2014).** Numerous data on the exophthalmic, disinfecting, antidiarrheal, diuretic, antispasmodic, carminative, and detoxifying properties of oregano are available too in the literature **(Ligaj et *al*., 2021).**

III.2. *Pistacia atlantica*

III.2.1. Botanical description

Pistacia atlantica Desf. (Atlas pistachio) is a deciduous tree, its vernacular name is "ElBetoum", up to 20 m in height and up to 1 m in diameter, can reach 300 years. It is a drought-tolerant species, which is adapted to semi-arid and arid areas. It can be found solitary or forming scattered stands, even in harsh areas in which few species can grow (1–3). On the leaves and petioles of this plant are found peculiar galls, like pale greenish-brown and rough excrescences, which are produced by insects **(Benamar et *al*., 2018; Gourine et *al*., 2010).** Pictures of some botanical characters of this species are reported in **figure 8**.

Figure 8. Pictures of *P. atlantica* leaves (A) and its fruits (B) **(Labdelli et *al*., 2020).**

III.2.2. Taxonomy

According to Tela Botanica Platform (https://www.tela-botanica.org/)(2011), the systematics of *Pistacia atlantica* is as follows:

Domain :	Eukaryota
Kindgom :	Plantae
Phylum :	Tracheobionta
Sub-phylum :	Angiospermae
Class :	Dicotyledonae
Order :	Sapindales
Family :	Anacardiaceae
Genus :	Pistacia
Species :	*Pistacia atlantica*

III.2.3. Geographical location

P. atlantica has a wide ecological plasticity, it grows in the Middle Eastern and Mediterranean regions, is widely distributed in Algeria, especially in semi-arid and arid regions where it grows wild in canyons (dayas) (**Benamar et *al*., 2018).**

In Algeria, it occurs in the wild from sub-humid environments to extreme Sahara sites. As a thermophilous xerophyte, *P. atlantica* grows in dry stony or rocky hill sides, edges of field, roadsides, near the base of dry stone walls and other similar habitats. The species grows well on clay or silty soils, although it can thrive also on calcareous rocks where roots develop inside cracks. Hence, *P. atlantica* has a wide ecological plasticity **(Said et *al*., 2011).**

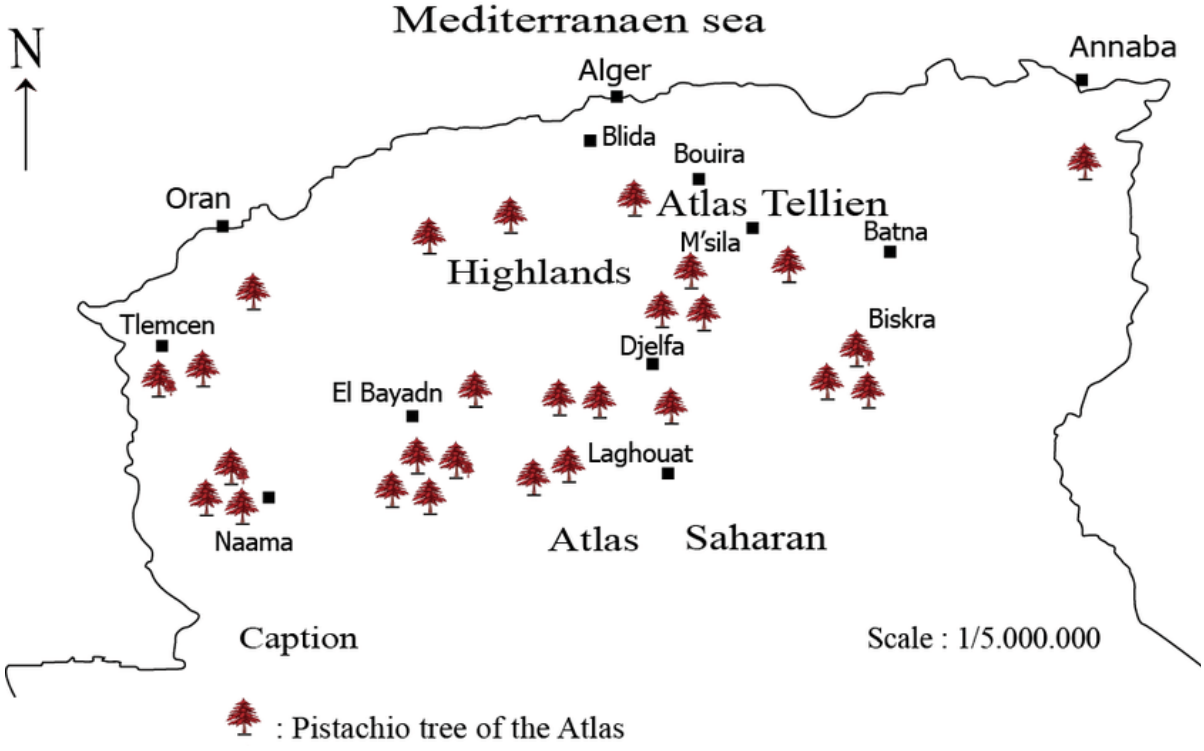


Figure 9. The distribution of *P. atlantica* in Algeria **(Chebouti-Meziou et *al*., 2014).**

III.2.4. Traditional use

P. atlantica is a food and medicinal plant used for treatment of pain due to its analgesic properties, dyspepsia, peptic ulcer, diabetes, and eye infection, used for stress, as a tonic, and used in decoction against for its antidiarrheal activity **(Toul et *al*., 2017).**

The fruits, called "el khodiri" by local population, find various applications in cooking. In southern Algeria, dried fruits are ground and mixed with dates or figs and used as a food

tonic (**Benamar et al., 2018).** Moreover, the galls of *P. atlantica* which are edible and sold in markets are used as an embalming gradient by rural habitants **(Gourine et *al*., 2011).**

III.2.5. Chemical and biological studies of *Pistacia atlantica*

The main phenolic compounds of the fruits and leaves of *P. atlantica* are benzoic acid derivates, hydroxycinnamic acid derivative, and flavonoids which are potent antioxidants and free radical scavengers **(Toul et *al*., 2017).** There are many pharmacological activities, such as antibacterial, antifungal, antihyperglycemic, scolicidal, and anticancer activities, have been reported and confirmed by current scientific documents (almost 600 articles published in Science Direct). The leaves and twigs of *P. atlantica* with its active substance 3-methoxycarpachromene showed antiprotozoal activity against *Plasmodium falciparum* **(Mahjoub et *al*., 2018).**

MATERIALS AND METHODS

1. Structure-Activity Relationship (SAR)

Through this work, we have screened the published papers on COVID-19 with phenolic compounds as inhibitors of the Mpro of SARS-CoV-2, we have found more than 2000 papers, we have regrouped them and study their inhibition. Following the same research strategy, we have selected some ligands that have not been published before. The ten chosen molecules are: 4-[[(2’ ,5’ Dihydroxybenzoyl)oxy]methyl]phenyl O-b-D-glucopyranoside, 4-[[(3’ ,4’ -Dihydroxybenzoyl)oxy]methyl] phenyl O-b-D-[6-O- (3’’,5’’-dimethoxyl-4’’-hydroxybenzoyl)] glucopyranoside, Acacetin 7-O-[4’’’-O-acetyl-b-D-apiofuransyl-(1=>3)]-b-D-xylopy ranoside, Acacetin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside, Apigenin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside, Acacetin-7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>2)]-bD-glucopyranoside, lonchocarpol A, licochalcone A, licochalcone C and 3-Methoxycarpachromene.

Furthermore, to identify the inhibition mechanism and the involved interactions that may exist between this enzyme and ligands implicated in this study, such as polar interactions (hydrogen bonds, ionic bonds) and hydrophobic interactions (contact between hydrophobic groups). Some are more important than others, thus, hydrogen bonds constitute one of the components playing an essential role in intermolecular interactions. This is mainly due to the ubiquity of polar groups in biological macromolecules. We have used the *AutoDock Vina* program (ADV) for molecular docking over 30 docking’s software that are currently available. This software has proven its efficiency and precision in docking of one or more ligands.

1.1. Protein preparation

First, the crystallographic structure of Mpro SARS-CoV-2 used in this work, is downloaded from the protein data bank (PDB ID: **6LU7**) (https://www.rcsb.org/structure/6LU7) presenting a high resolution of 2.16 Å. The 6LU7 protein contains two chains, A and B, which form a homodimer. Chain A was used for macromolecule preparation (**Figure 10)**, containing 306 amino acids, and approximately 31.4 KDa of molecular mass. Chain B, composed of 6 residues represent the native co-crystal ligand present in the crystal structure named n-[(5-methylisoxazol-3- yl)carbonyl]alanyl-l-valyl-n~1~-((1r,2z)-4-(benzyloxy)-4-oxo-1-{[(3r)-2-oxopyrrolidin-3- yl]methyl}but-2-enyl)-l-leucinamide (inhibitor N3), with total residue count of 312 amino acids and a molecular weight of 34 KDa. This crystal structure was originally resolved by Liu et al. (2020).

In molecular docking (MD), the protein must be prepared following three main steps, the first, by removing all water molecules, heteroatoms, any ligands and co-crystallized solvent. Then, Polar hydrogens and partial charges were added to the structure using Autodock tools (ADT) version 1.5.4 and co-crystallized inhibitors have been removed. Finally, the third step, consists of setting the grid box center and size of the binding pocket by the grid coordinates x, y and z using ADT.

The molecular docking results were loaded into the Discovery studio visualizer v20.0 (DSV) program **(Serseg et *al*., 2020).** The best poses were chosen according to their minimum binding energy values in kcal/mol and the ratio of the most repeated poses (RR) in percent (%) besides polar hydrogens, and hydrophobic interactions inside the binding pocket of Mpro **(Linani et *al*., 2020).**

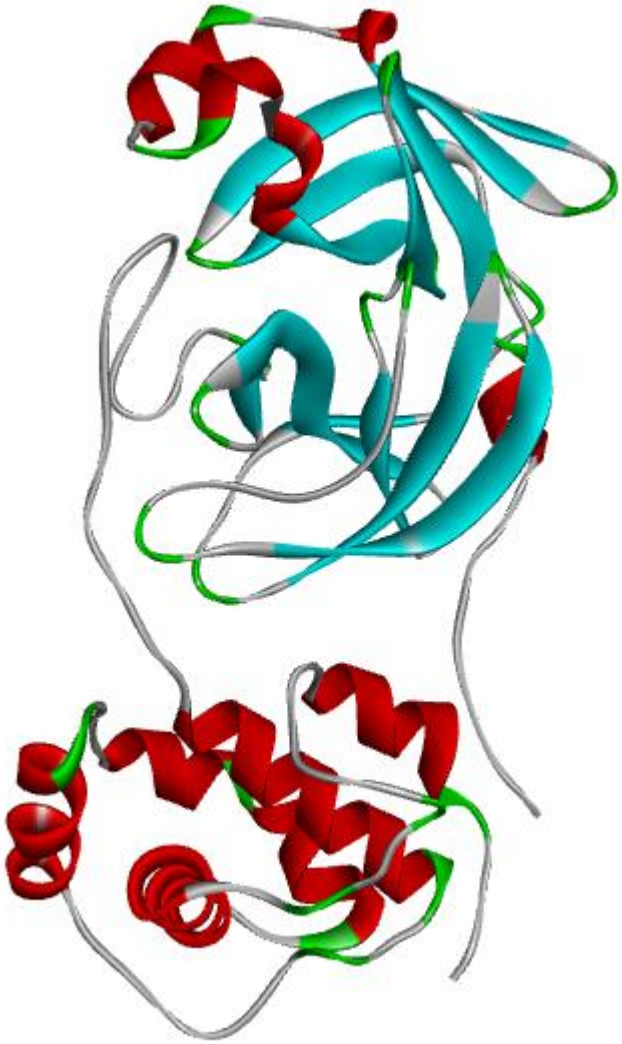


Figure 10: Structure of SARS-COV-2 Mpro without N3 inhibitor ((PDB ID: **6LU7**) obtained by Discovery Studio Visualizer v. 4.0.

1.2. Ligand preparation

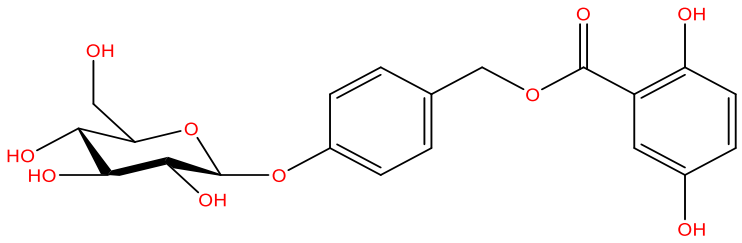
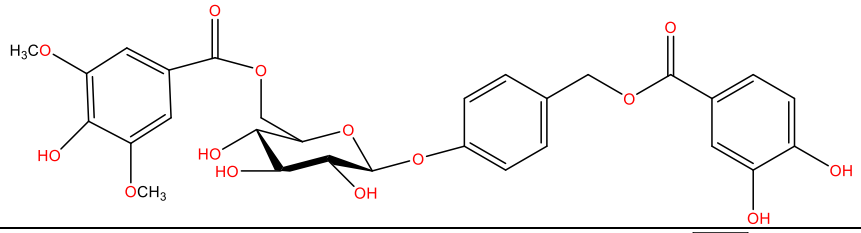
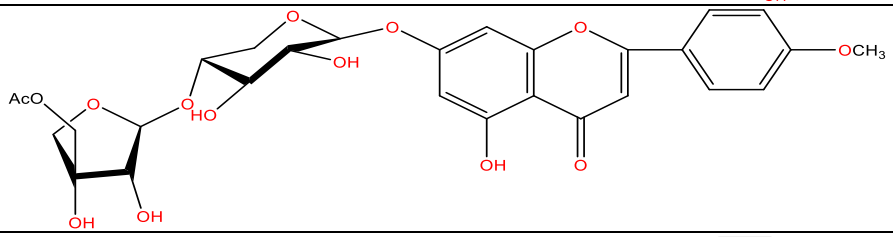
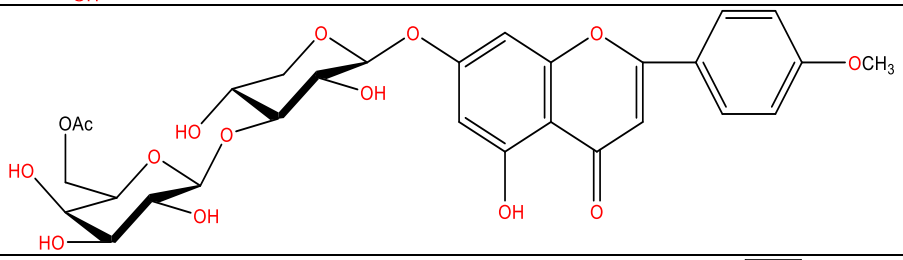
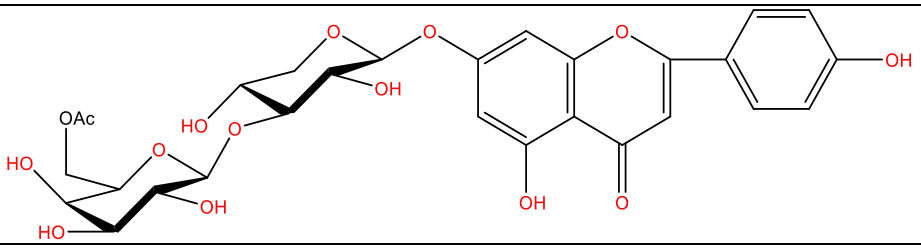
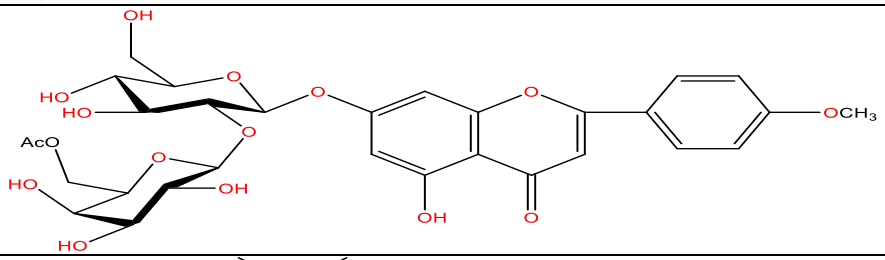
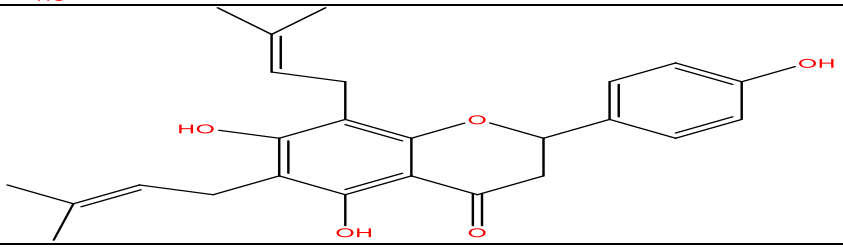
Polyphenols have been reported to exhibit antiviral bioactivities, we investigated the effect of these ten compounds as potential inhibitors of COVID-19 Mpro (PDB ID: **6lu7**) using molecular docking (MD).

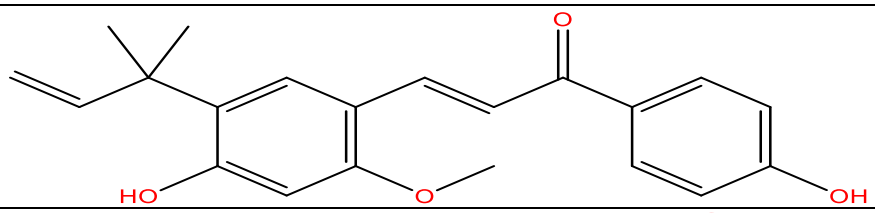
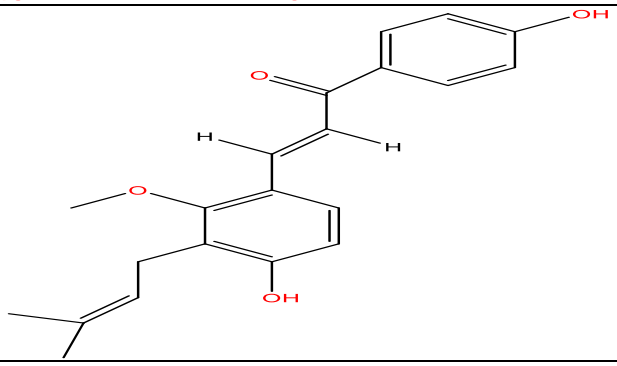
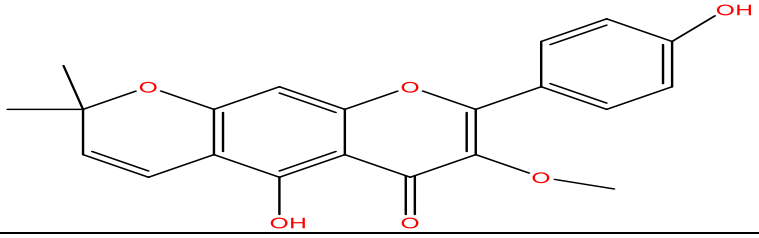
The ten compounds are coded to facilitate their discussion and citation in the following sections of this work, the codes are presented as below:

- **(M1):** 4-[[(2’ ,5’ -Dihydroxybenzoyl)oxy]methyl]phenyl O-b-D-glucopyranoside,
- **(M2):** 4-[[(3’ ,4’ -Dihydroxybenzoyl)oxy]methyl] phenyl O-b-D-[6-O- (3’’,5’’-dimethoxyl-4’’-hydroxybenzoyl)] glucopyranoside **,**
- **(M3):** Acacetin 7-O-[4’’’-O-acetyl-b-D-apiofuransyl-(1=>3)]-b-D-xylopyranoside**,**
- **(M4):** Acacetin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside**,**
- **(M5):** Apigenin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside**,**
- **(M6):** Cacetin-7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>2)]-bD-glucopyranoside**,**
- **(M7):** Lonchocarpol A**,**
- **(M8):** Licochalcone A**,**
- **(M9):** Licochalcone C
- **(M10):** 3-Methoxycarpachromene**,**

The 2D structures of three chemical molecules **(M7, M8, M9)** were drawn by ChemDraw. Then, 3D structures downloaded and converted to *pdbqt* format by ADT software. While the 3D structures of the seven chemical compounds **(M1, M2, M3, M4, M5, M6** and **M10)** were retrieved in the SDF format from PubChem database (https://pubchem.ncbi.nlm.nih.gov). And Discovery studio Visualizer 2021 (DSV) was used to convert them into *pdbqt* type. The file conversion into pdbqt format was done to allow the file to be loaded in Autodock Vina for MD simulation. The 2D structures of the studied compounds and their PubChem CID are shown in **Table 2.**

Table 2. The chemical formulas, 2D structures and PubChem CID of the ten studied molecules.

Inhibitors	Chemical structure	PubChem CID
M1 **BF:** $C_{20}H_{22}O_{10}$		/
M2 **BF:** $C_{29}H_{30}O_{14}$		/
M3 **BF :** $C_{28}H_{30}O_{14}$		/
M4 **BF :** $C_{29}H_{32}O_{15}$		/
M5 **BF :** $C_{28}H_{30}O_{15}$		/
M6 **BF :** $C_{30}H_{34}O_{16}$		/
M7 **BF:** $C_{25}H_{28}O_5$		124035

M8 **BF:** $C_{21}H_{22}O_4$		5318998
M9 **BF:** $C_{21}H_{22}O_4$		9840805
M10 **BF :** $C_{21}H_{18}O_6$		/

BF: Brut formula **CID:** compound ID

1.3. Molecular docking settings

Molecular docking (MD) is an efficient approach for screening appropriate therapies against a specific drug target of lethal pathogens. This powerful tool is used to model the interaction between small ligands (inhibitors) and macromolecules (enzyme), to form a stable complex. Thus paving the way for drug discovery, reducing costs and time (especially for emerging diseases such as COVID-19) and accelerating analyzes of target interactions with drug candidates **(Azim et *al*., 2020).** Therefore, different computational studies have been published in order to better understand the mechanism of M-pro and attempt to inhibit its function **(Gimeno et *al*., 2020).**

In the current study, molecular docking was achieved to predict the binding mode of 10 compounds from different sources with a target from COVID-19; the main protease (PDB code: **6LU7**), using *AutoDock Vina* program on LENOVO laptop, Windows 10, 64 bits, with specifications of processor I5 7th generation, operation system 8GB of RAM. The number of docking runs was set at 10 runs. The software uses rectangular boxes for the binding site, the center of the box has been set and displayed using ADT. The grid boxes for 6LU7 of center (x= -12.441, y= 12.198 117 and z= 67.681) with size (x=18, y=22, z=24), with one Å separated grid points positioned in the middle of the active site. The default settings were used, except that the

number of output conformations was set to one. The number of solutions obtained is equal to 10 conformations for each ligand and enzyme. All these solutions are very well handled. The "random seed" is random. The preferred conformations were those of lower binding energy within the active site **(Serseg et *al*., 2020).** Finally, the generated docking results were directly loaded into Discovery Studio visualizer 2021, to study possible docking modes between compounds and the main protease of COVID-19.

2. ADMET settings

Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) are pharmacokinetic and toxicological properties of a ligand that deals primarily with its absorption, distribution, metabolism, excretion and toxicity in the human body. Evaluation of ADMET properties of potential drug molecules has become key step and appropriate methodology employing *in silico* drug development and discovery processes, to ensure the pharmacokinetics (PK) properties and avoid failure in the last phases **(Mahrosh et Mustafa, 2021; Mpiana et *al*., 2020).** Thereby, the process of optimizing drug development is multidimensional. In the end, a balance must be reached in order to obtain the best both in terms of the activity and the properties of the compound **(Linani et *al*., 2020).**

Drug absorption is the movement of a drug into the bloodstream. This process is indeed very complex and depends on many parameters, including a drug's physicochemical properties, formulation, and route of administration **(Lagorce et *al*., 2017).**

Drug distribution describes the reversible movement of a drug from one location to another within the body. Usually, it refers to a drug transfer from the blood to various tissues. Some factors affecting drug distribution include regional blood flow rates, molecular size, polarity and binding to serum proteins, forming a complex **(Lagorce et *al*., 2017).**

Metabolism is the biotransformation of drugs to facilitate their excretion. Metabolic processes are mainly catalyzed by redox enzymes, termed cytochrome P450 enzymes (CYP3A4, CYP2D6, CYP2C9, CYP2C19, and CYP1A2), for the most part, produced in the liver, to test the metabolic stability of a molecule. As metabolism occurs, the initial compound is converted to new compounds called metabolites **(Lagorce et *al*., 2017).**

Excretion. Compounds and their metabolites need to be removed from the body via excretion, usually through the kidneys (urine) or in the feces **(Lagorce et *al*., 2017).**

Toxicity is another very important factor, which often overshadows the ADME behavior. Failure of drugs at clinical trial stage due to adverse effects generated because of their toxicity proves very expensive and detrimental in the drug development process. Toxicity is the degree to which a substance can damage an organism or substructures of the organism, such as cells and organs, and remains one of the most significant reasons for late-stage drug development failure. Early identification of toxicity would thus be very valuable (**Lagorce et *al*., 2017; Nisha et *al*., 2016).**

The ten compounds were taken for ADMET test profile prediction with the help of web based server Pre-ADMET v2.0 (https://preadmet.bmdrc.kr/) **(Lee et *al*., 2003**). We have used as the **absorption** parameter; the human intestinal absorption (HIA), buffer solubility (mg/l), for the **distribution**; blood-brain barrier penetration (C.brain/C.blood) and skin permeability, for **metabolism**; Cytochrome P450 3A4 inhibition and substrate. In terms of **toxicity,** we have demonstrate the Ames test, and the hERG inhibition.

RESULTS

1. Structure-activity relationship (SAR)

Table 3 shows the binding energy between the ten compounds and Mpro with various interactions, including hydrogen bonds and the interaction distance between amino acids and the active site of enzyme after docking analysis. Binding energy is the main parameter that is generated as a result of molecular docking. It gives us the idea of the strength and affinity of the interaction between the ligand and the receptor. The greater the binding energy, is the weaker the interaction and vice versa. Thus, we intend to search for the ligand, which has the least binding energy, therefore, the best affinity among the molecules to be tested (**Nisha et *al*.; 2016).**

Molecular docking results ranked by binding energy value of tested ligands were; M3, M7, M10>M2>M4>M5>M6>M9>M1, M8. The binding energies were, -8.2, -8.2, -8.2, -8.1, -7.9, -7.8, -7.7, -7.5, -7.3 and -7.3 kcal/mol, respectively. It was shown that all ligands have good binding affinity towards the receptor in different ways. The interaction analysis was performed to get highly selective compounds that preferentially bind to the SARS-CoV-2 Mpro substrate-binding pocket. Results of ligands docking ranked by number of hydrogen bonds produced were; M1, M8>M2, M9, M10>M3, M4>M5, M6, M7. It is well known that the high binding affinity of the drug compounds depends on the type and amount of bonds that occurs with the target protein. According to minimum energy values, the RR percentage and the number of hydrogen bonds M10, M2 and M7 were the best inhibitors among the other tested compounds. M3, M4, M5, M6 and M9 were reported as satisfied inhibitors. As well as M1 and M8 were the last ranked inhibitor models for enzyme with the minimum saved criteria comparing to other inhibitors. Otherwise, all the obtained results save RR percentage of 60-100%, 80% of the saved interactions were a hydrophilic type. All results are summarized in **Table 3.**

Table 3. Molecular docking analysis of several compounds against SARS-COV-2 Mpro (PDB: 6LU7)

Molecule	Energy (Kcal/mol)	RR %	Closest residues	Hydrophobic interactions	Number	Hydrogen bonds	Length (Å)
M1	-7.3	100	**Arg188, Met165 Glu166, His41 Cys145, Asn142 Gly143, Thr26**	Pi-alkyl Pi-Pi T-shaped	02	Cys145 Thr26 Clu166 Met165 Arg188	2.36 2.13 2.08 2.81 2.59
M2	-8.1	100	Pro168, **Glu166 Leu141, Gln189 Arg188, Met165 Met49, His41**	Pi-Alkyl Alkyl Pi-Pi T-shaped	05	His163 Leu141 Glu166 Thr90 Met165	0.56 2.93 2.48 2.03 2.83
M3	-8.2	60	**Met165, Cys145 Asn119, His41**	Pi sulfur, Pi alkyl	01	Asn119 His41	2.51 2.40
M4	-7.9	100	**Thr26, Leu27 Gly143, Cys145 His41, Met165 Glu166,** Pro168	Pi-Pi T-shaped Alkyl Pi-Alkyl	04	Glu166 Gly143	2.58 2.21
M5	-7.8	70	**Gly143, Cys145 Met165,** Pro168 **His41, Glu166**	Pi-Alkyl Pi-Pi T-shaped	02	Glu166 Gly143	2.44 2.22
M6	-7.7	80	**His41, Gly143 Cys145, Met165 Glu166,** Pro168 **Gln189**	Pi-Pi T-shaped Alky Pi-Alkyl	04	Glu166 Gln189	2.54 2.71
M7	-8.2	100	**Met165, His172 His163, Cys145 Thr26**	Alkyl Pi-Alkyl	05	Thr26	2.71
M8	-7.3	100	**Leu141, Thr190 Gln192, Ser144 Cys145,** Pro168 **Met165**	Pi-alkyl Alkyl	04	Cys145 Leu141 Ser144 Gln192 Thr190	3.46 2.19 2.97 2.66 2.37
M9	-7.5	100	**Thr190, Gln189 Arg188, Met165 His41, Ser144 Cys145, Leu27**	Alkyl Pi-Alky	05	Ser144 Thr190	2.25 2.62 2.22
M10	-8.2	100	**Glu166, Asn142 Gln192, Thr190 Met165**	Pi-alkyl	01	Asn142 Glu166 Gln192 Thr190	2.31 2.34 2.57

RR: Repeating ratio %

2. ADMET analysis

The ADMET parameters were determined and verified for qualification with their standard ranges **(Appendix 1).** The ADMET properties prediction using Pre-ADMET server of tested polyphenols M1-M10 showed in **Table 4.**

Table 4. Predicted ADMET properties of the ten studied compounds

Pharmacokinetics	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Absorption										
Human intestinal absorption (HIA %)	38.14	44.76	46.56	27.51	15.33	42.95	92.40	94.08	94.08	94.14
Buffer solubility (mg/l)	386.45	228.25	38.90	37.70	48.62	61.76	20.94	0.86	2.60	13.69
Distribution										
Blood-brain barrier penetration (C.brain/C.blood)	0.03	0.03	0.03	0.03	0.03	0.03	5.62	3.12	4.29	0.12
Skin permeability (logKp, cm/hour)	4.17	-3.41	-4.17	-4.35	-4.35	-4.23	-2.14	-1.18	-1.84	-2.84
Metabolism										
Cytochrome P450 3A4 inhibition	Inhibitor	Inhibitor	Inhibitor	Inhibitor	Inhibitor	Inhibitor	Inhibitor	Inhibitor	Inhibitor	Inhibitor
Cytochrome P450 3A4 substrate	Weakly	Weakly	Weakly	Weakly	Weakly	Weakly	Substrate	Weakly	Weakly	Non
Toxicity										
Ames test	Mutagen	Mutagen	Non-mutagen	Non-mutagen	Non-mutagen	Non-mutagen	Non-mutagen	Mutagen	Non-mutagen	Mutagen
HERG_inhibition	High-risk	High-risk	High-risk	High-risk	Ambiguous	High-risk	Medium-risk	High-risk	Medium-risk	Medium-risk

HERG: Human ether related gene channel

DISCUSSION

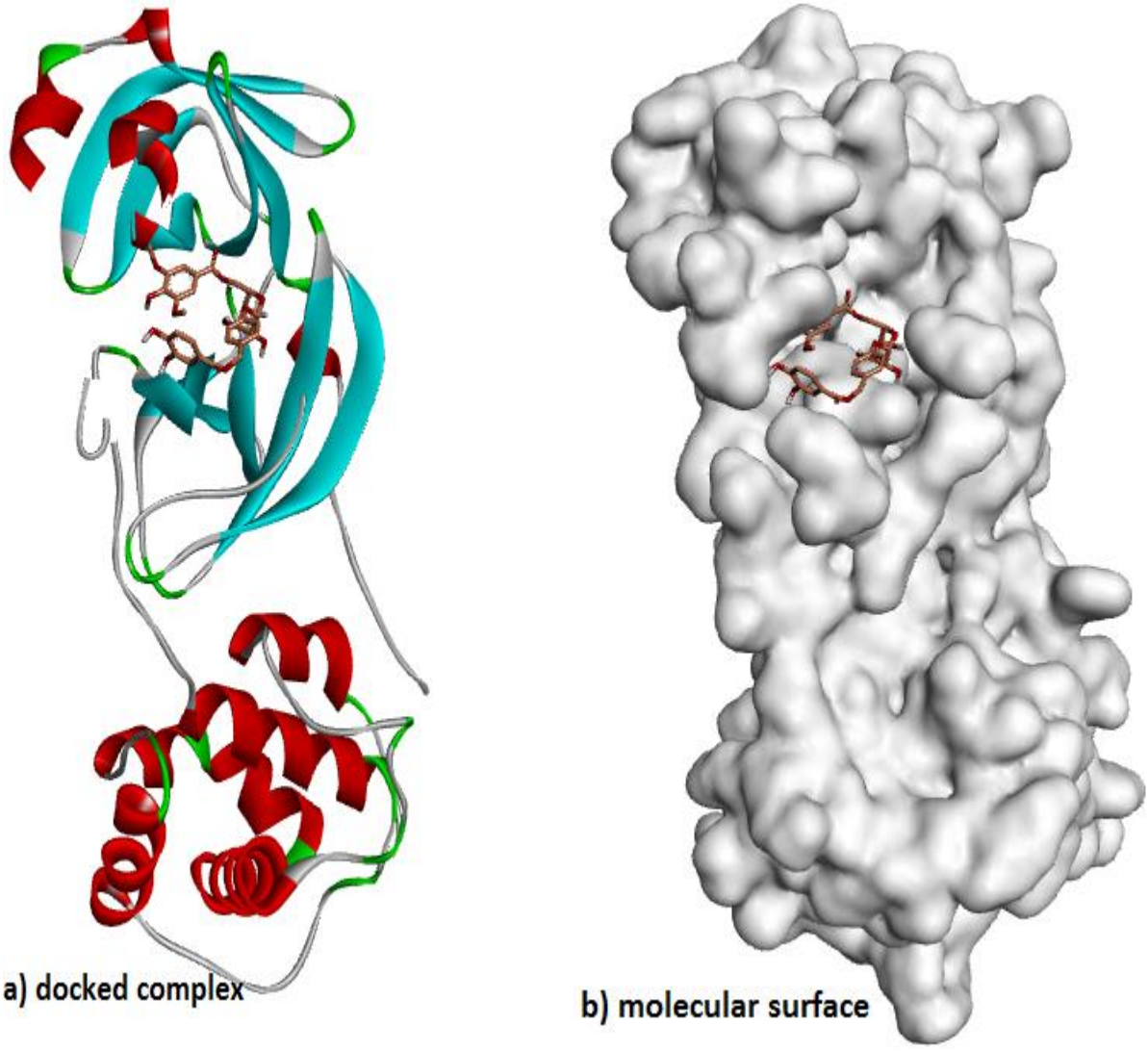
1. SAR study and Molecular docking

The objective of our present *in silico* study is to look over new drugs by elucidating the interaction mode of the above selected ligands with the catalytic site of SARS-COV-2 Mpro. The MD analysis of Mpro with 10 selected phenolic compounds from different medicinal plants showed that the majority of the compounds bind to Mpro with favorable binding energy ranging from -7.3 kcal/mol (for M1, M8) to -8.2 kcal/mol (for M10, M7, M3).

M1, 4-[[(2' ,5' -Dihydroxybenzoyl)oxy]methyl]phenyl O-b-D-glucopyranoside, is a bicyclic organic compound that belongs to the glycosylated phenolic compounds, from *Origanum vulgare*, of chemical formula $C_{20}H_{22}O_{10}$. It has two (02) benzene rings (methyl 2,5-dihydroxybenzoate, and benzene) and one sugar (D-glucose), linked together with glycosidic bond, its structure is shown in **Table 2**. According to the MD results, M1 was docked into the active site cavity by internal fixation by a U-shape, due to its low molecular weight (MW = 422.39g / mol) relative to the molecular weight of Mpro (MW: 31400g / mol), this explains its affinity to this enzyme represented by an energy value of -7.3(Kcal/mol), and 100% as RR value. Whereas, its energy was ranked 9th among other inhibitors **(Table 1).** To understand its mode of action, we localize the two reactive functions in its structure, methyl 2,5-dihydroxybenzoate and D-glucose, which have cooperated to achieve the best orientation and pointed towards the inside of the active site cavity forming the U-shape. They were essential in the inhibition mechanism and enzyme stabilization due to ten (10) formed favorable interactions. Methyl 2,5-dihydroxybenzoate formed many interaction types, hydrogen bonds with three amino acids, Glu166 with heteroatom O of the carboxyl group, Arg188 and Met165 with hydroxyl function of C2, we also recorded two hydrophobic bonds of the Pi-alkyl and Pi-Pi T-Shaped type with His41 and Met165, it formed as well, one Pi-donor hydrogen type, His41 with hydroxyl function of C5. So that M1 was also formed hydrogen interactions between hydroxyl groups of C2, C4 of D-glucose and Cys145, Thr26, respectively, which increases the polarity, and two carbon hydrogen bond type, Asn142 with OH of C2 and Gly143 with OH of C4 **(Table 3, Appendix 2).**

M2, 4-[[(3' ,4' -Dihydroxybenzoyl)oxy]methyl] phenyl O-b-D-[6-O- (3'',5''-dimethoxyl-4''-hydroxybenzoyl)] glucopyranoside, is a polycyclic compound that belongs to glycosylated phenolic compounds, from *Origanum vulgare*, of chemical formula $C_{29}H_{25}O_{14}$, with a molecular weight 602.55 g/mol. It has three (03) benzene rings (3,4 dihydrobenzoic acid, toluene, 4-hydroxy-3.5 dimethoxy benzoic acid) and one sugar (D-glucose) **(Table 2).** It differs from M1 by having a third benzene ring; 4-hydroxy-3.5 dimethoxy benzoic acid, its presence

decreases the binding energy, thus, increases the stability of this ligand towards the enzyme, representing one of the best energy scores -8.1 (Kcal/mol) with a 100 % RR value, ranked the 4th among other inhibitors **(Table 3).** One (01) hydrogen bond interaction [Met165 (2.83 Å)], four (04) hydrophobic interactions of Pi-Pi T-shaped, alkyl and Pi-alkyl type (His41, His41, Met49, Met165), a single C-H bond (Gln189) were observed in this ring. On the other hand, Mpro-M2 complex was further stabilized by three H-bond interactions with the D-glucose [Glu166 (2.3 Å), His163 (2.3 Å) and Leu141 (2.3 Å)], a single electrostatic interaction [Glu166 (3.38 Å)] with the toluene. Besides these, a couple of interactions was observed in the last ring; 3,4 dihydrobenzoic acid, a H-bond interaction [Thr190 (2.03 Å)] and a hydrophobic interaction [Pro168 (4.35 Å)] **(Table 3, Figure 11 c)).** All these interactions types and numbers have been acting jointly to achieve the best orientation inside the active site cavity forming the U-shape. **(Figure 11 a), b)).**



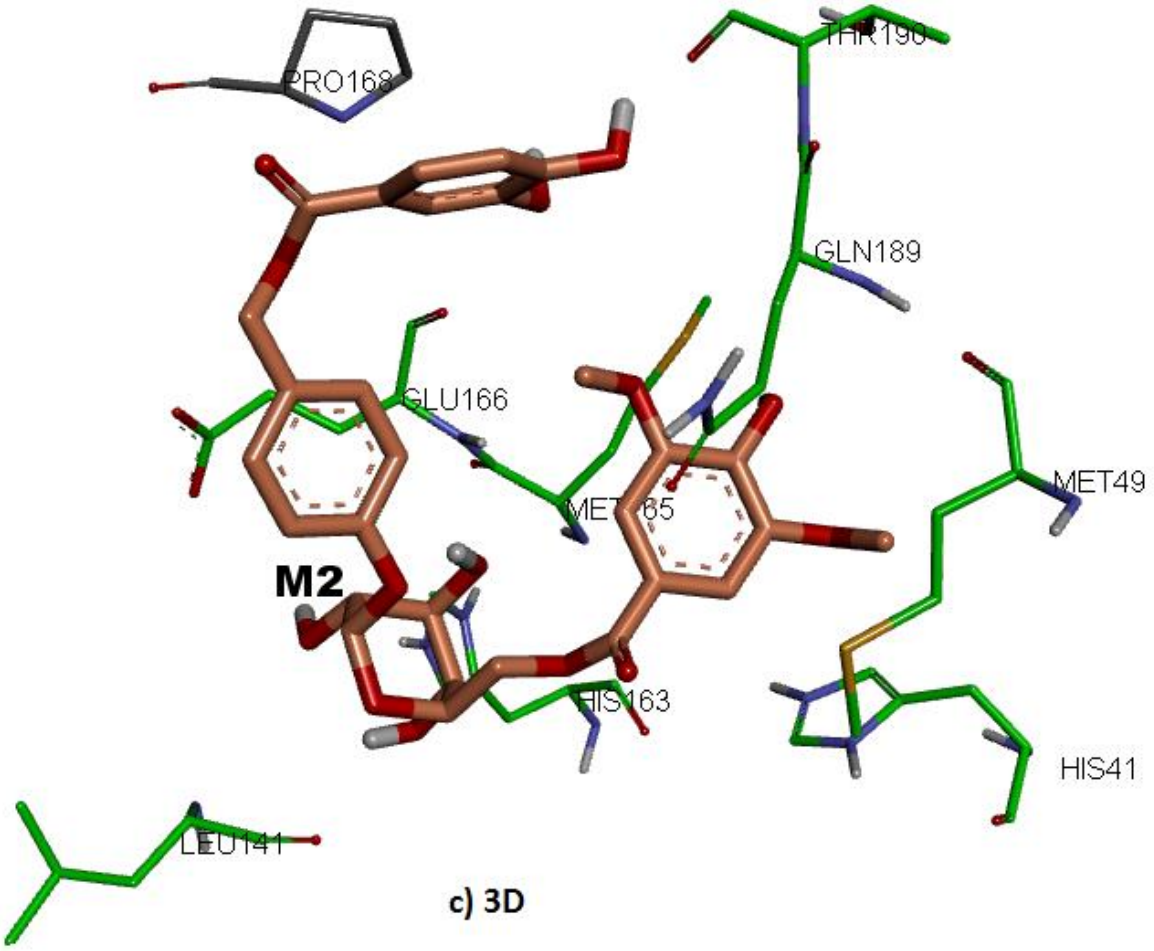


Figure 11. Molecular docking interaction of 4-[[(3’ ,4’ -Dihydroxybenzoyl)oxy]methyl] phenyl O-b-D-[6-O- (3’’,5’’-dimethoxyl-4’’-hydroxybenzoyl)] glucopyranoside with Mpro (6LU7) (Catalytic amino acids in green, ligand in pink).

M3, Acacetin 7-O-[4’’’-O-acetyl-b-D-apiofuransyl-(1=>3)]-b-D-xylopyranoside, belongs to glycosylated phenolic compounds, from *Origanum vulgare*, of chemical formula $C_{28}H_{30}O_{14}$, and with a molecular weight 590.53 g/mol, it is a flavone glycoside, composed of an aglycone; acacetin and two sugars; D-xylose and D-apiose **(Table 2).** The binding energy of M3 to the active site is even smaller than that of the other inhibitors, indicating that M3 has a higher binding activity, it was -8.2 Kcal/mol, it is the least stable inhibitor with a RR value equal to 60%, its energy is ranked among first inhibitors **(Table 3).** M3 was found to form two (02) H-bond interactions with residues Asn119 and His41, Asn119 with acetoxy group of D-apiose (2.51 Å), and His41 with Hydroxy function of C3 of D-xylose (2.40 Å). In addition, M3 built Pi-sulfur contact with Met165 (5.65 Å) and a hydrophobic interaction (Pi-Alkyl) was formed with Cyc145 (4.96 Å), as seen in **Figure 12 c)**. The position and nature of ligand was responsible for the preferred horizontal orientation inside the active site cavity **(Figure 12 a), b)),** and was also essential in the inhibition mechanism and stabilizing the enzyme.

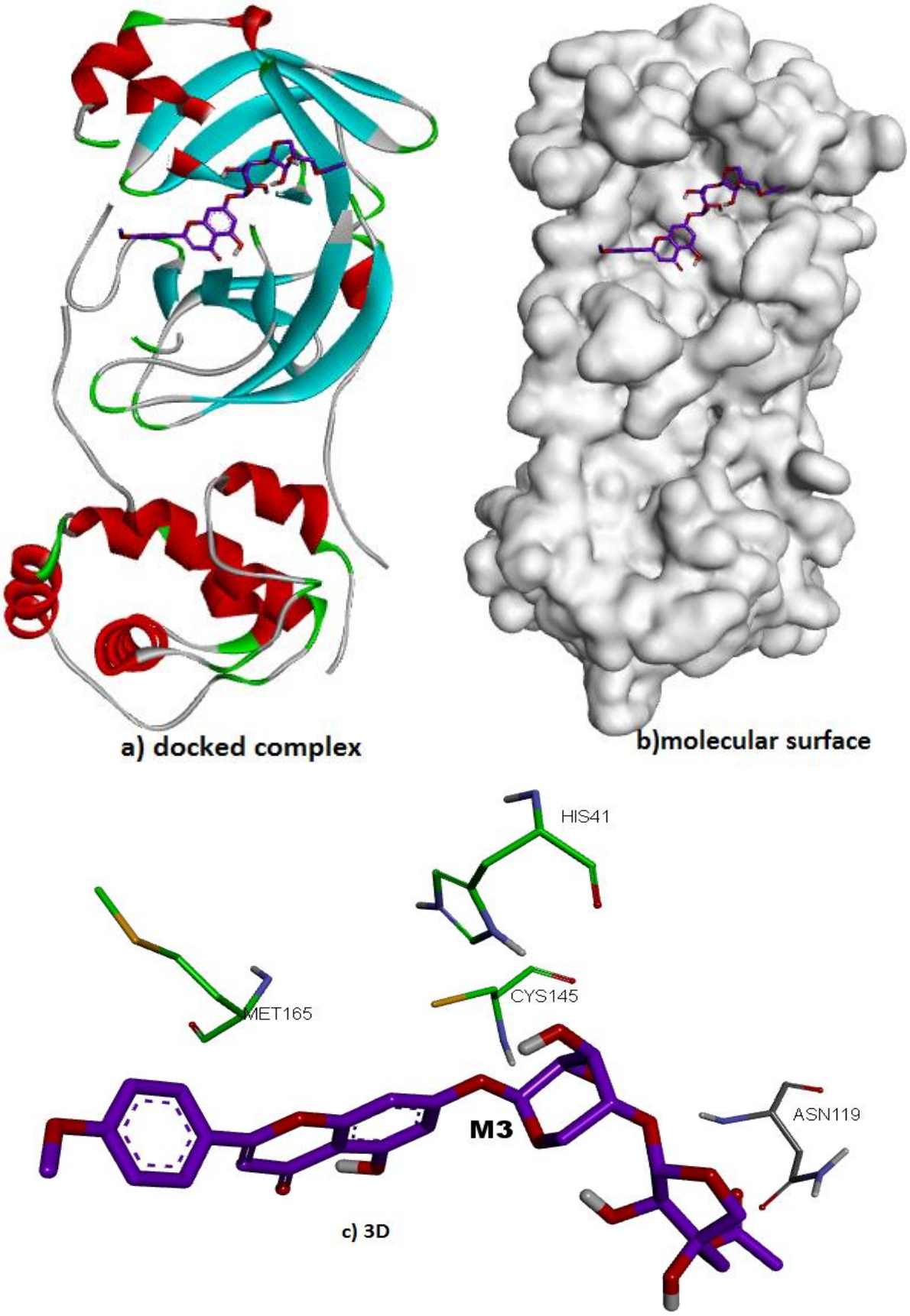


Figure 12. molecular docking interaction of Acacetin 7-O-[4’’’-O-acetyl-b-D-apiofuransyl-(1=>3)]-b-D-xylopyranoside with Mpro (6LU7) (Catalytic amino acids in green, ligand in dark purple).

M4, Acacetin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside, is a polycyclic compound, belongs to glycosylated phenolic compounds, from *Origanum vulgare*, of chemical formula $C_{29}H_{32}O_{15,}$ molecular weight 620.56 g/mol, its structure is similar to M3, except the two sugars who are replaced by D-galactose and D-xylose, as seen in the **Table 2.** From a binding energy point of view, M4 was docked into the active site of Mpro with -7.9 kcal/mol and 100% of RR value, its energy was ranked 5th among the other inhibitors **(Table 3).** The Mpro-M4 complex was stabilized by multiple interactions that participate in the inhibition mechanism and internal stabilizing the Mpro, forming the shape of the letter L, remarkably that the last part (D-galactose) is not accompanied inwards; it is kind of out of site cavity. C-H bond interactions between Glu166 (3.05 Å, 3.14 Å, 3.55 Å), Pro168 (3.52 Å, 3.15 Å), and sugars, and between Thr26 (3.37Å) and acacetin, hydrophobic interactions with Leu27 (5.01 Å), His41 (4.92 Å) and Met165 (4.39 Å, 5.41 Å), we also evidenced that Cys145 of Mpro interacted with M4 via Pi-donor hydrogen bond with distance 4.03 Å, a couple of H-bond interactions [Glu166 (2.3 Å) with C2 of D-xylose and Gly143 (2.3 Å) with hydroxyl group of acacetin] were observed too **(Table 3, Appendix 3).**

M5, Apigenin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-D-xylopyranoside, is a polycyclic compound, that belongs to the glycosylated phenolic compounds, from *Origanum vulgare*, of chemical formula $C_{28}H_{30}O_{15}$, and with a molecular weight 606.53 g/mol. It was found that M5 very similar to M4, except for the methoxyl at C-4’ position which is absent in M5 **(Table 2),** that’s why the binding energy values are close, the binding energy of M5 to the Mpro was -7.8 kcal/mol moderately stable with a RR value equal to 70%, in 6th class, while that of M4 was -7.9 kcal/mol in 5th class, forming also the shape of the letter L, but this time the ligand by its functions are all inside the active site cavity, engaging in the inhibition and stabilizing of the enzyme **(Appendix 4).** M5 is composed of an aglycone; apigenin and two sugars; D-galactose and D-xylose, the biglycosidic side interacted with Mpro via H-bond, between OH of D-galactose C2 and Glu166 (2.44 Å), four (04) C-H bonds with both carbohydrates. The apigenin side also formed a single H-bond [Gly143 and hydroxyl (2.22 Å)], Pi-sulfur bond [Cys145 (4.23 Å)], alkyl bond [Met165(4.46 Å)] and Pi-Pi T-Shaped bond [His41(4.98 Å)] **(Table 3, Appendix 4).**

M6, Acacetin-7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>2)]-bD-glucopyranoside, is a polycyclic compound, that belongs to glycosylated phenolic compounds, from *Origanum vulgare*, of chemical formula $C_{30}H_{34}O_{16}$ and a molecular weight 650.59g/mol. Its structure is similar to M3 and M4, except the two sugars are replaced by D-galactose and D-Glucose **(Table**

2). According to MD results, the ligand was docked with -7.7 Kcal/mol as binding energy classify the 7th among others, with a RR value equal to 80% **(Table 3).** M6 was oriented within the active site like a pocket, the D-galactose was more pointed to the catalytic site **(Appendix 5)** and forms two (02) H-bonds [Glu166 with acetoxy group (2.54 Å), Gln189 with D-galactose C3 (2.71 Å)], four (04) hydrophobic interactions between Pro168 (5.49 Å, 4.06 Å), Met165 (4.51 Å) and His41 (4.90 Å) and the aglycone acacetin were formed **(Table 3).**

M7, Lonchocarpol A, 6,8 dipernyl naringenin, is a trihydroxy flavanone belongs to phenolic compounds, of chemical formula $C_{25}H_{28}O_5$, molecular weight 408.46 g/mol **(Table 1).** Among all the test candidates in this study, M7 displayed the lowest binding energy of -8.2 kcal/mol, thus, M7 displayed much better binding than others with good stability represented by a RR of 100% **(Table 3),** with good internal fixation in the active site cavity due to its low molecular weight **(Figure 4 a), b)).** When M7 was docked into the active site cavity of Mpro, one (1) H-bond interaction [Thr 26 with OH of C-4' (2.22 Å)], five (05) hydrophobic interactions, one (01) Alkyl type [Met165 (5.34 Å)] and four (04) Pi-Alkyl type; [Met165 (5.12 Å), Cys145 (4.39 Å), His 163 (5.18 Å, 4.62 Å)] were formed **(Table 3, Figure 13 c)).** All these interaction types and numbers have been acting jointly to achieve the best orientation inside the active site cavity.

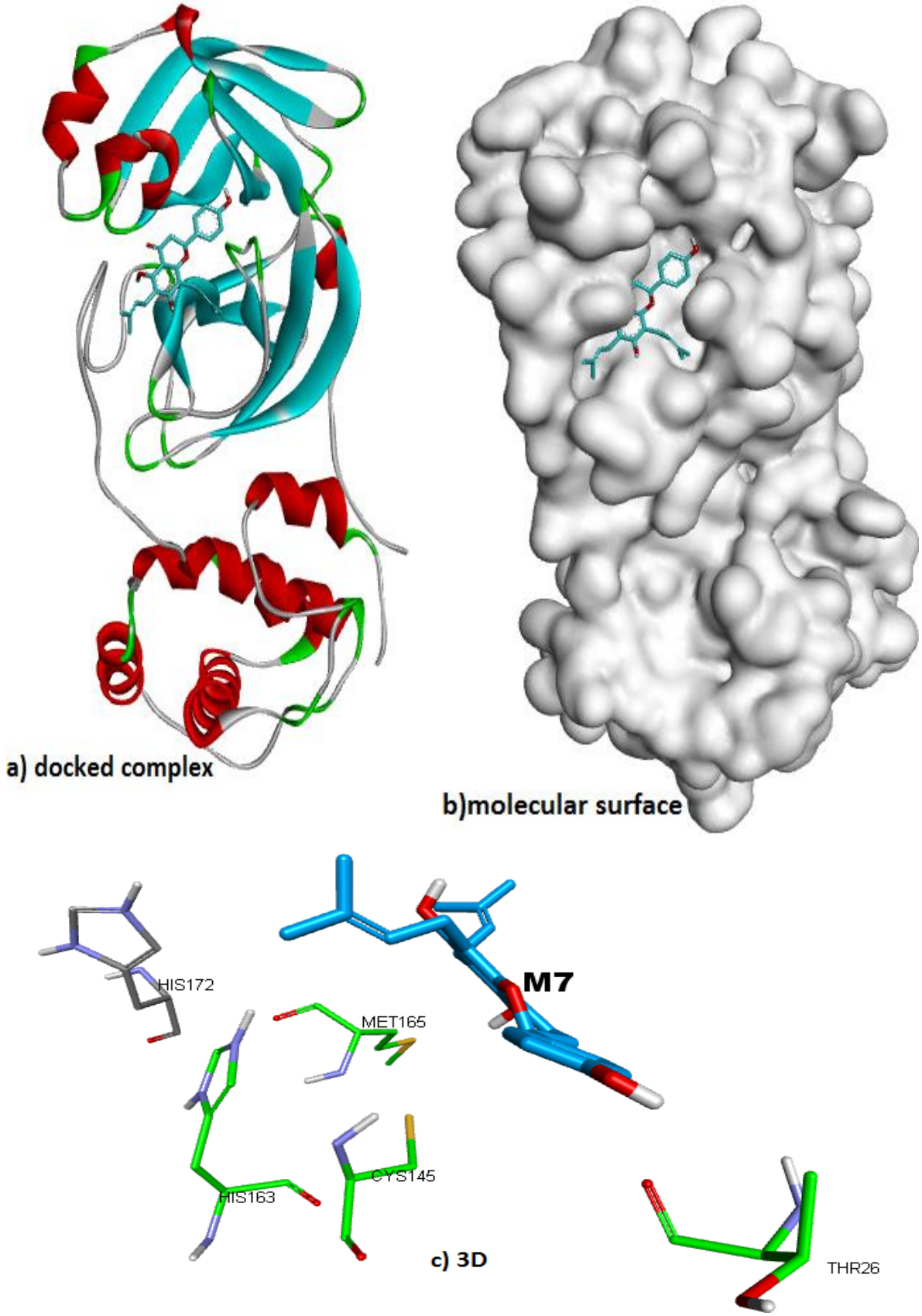


Figure 13. Molecular docking interaction of Lonchocarpol A with Mpro (6LU7) (Catalytic amino acids in green, ligand in bleu).

M8, Licochalcone A, 3-[4-hydroxy-2-methoxy-5-(2-methylbut-3-en-2-yl)phenyl]-1-(4-hydroxyphenyl)-2-propen-1-one, is a member of chalcones, of chemical formula $C_{21}H_{22}O_4$, molecular weight 338.40 g/mol **(Table 2).** The MD results shows that 100% of poses are in the active site, with the slightest energy of -7.3 kcal/mol, ranked 10^{th} among other inhibitors **(Table 3).** It shows also that it could extremely fit inside the substrate-binding pocket as seen in **(Appendix 6),** considering the best pose of ligand fixation in Mpro active site cavity, we explain that by the structure flexibility, shape, and various chemical function properties**.** It forms a strong hydrogen bond network with five amino acids, two (02) with hydroxyl group of a ring; Thr190 (2.37 Å), Gln192 (2.66 Å), three (03) with hydroxyl group of the other ring; Cys145, Leu141, Ser144, with distances, 3.46 Å, 2.19 Å, 2.97 Å, respectively. Likewise, there is four (04) hydrophobic interactions between M8 and Mpro at the binding site, a couple of Pi-Alky type with Met165 (4.82 Å) and Cys145 (5.19 Å), another couple of Alky type with Pro168 (5.39 Å), and Met165 (4.62 Å) **(Table 3, Appendix 6).**

M9, Licochalcone C**,** (*E*)-3-[4-hydroxy-2-methoxy-3-(3-methylbut-2-enyl)phenyl]-1-(4-hydroxyphenyl)prop-2-en-1-one, is a retrochalcone, of chemical formula $C_{21}H_{22}O_4$, molecular weight 338.40 g/mol **(Table 2).** As stated in MD results, the predicate binding affinity to the Mpro is -7.5 Kcal/mol, and the blind docking shows that 100% of poses are in the active site. It was docked into the active site cavity by internal fixation, forming an I-shape and located in a deep tunnel within the enzyme, this fixation was essential in the inhibition mechanism **(Appendix 7).** The Mpro-M9 complex was stabilized by eight (08) multiple bond interactions, one (01) H-bonds with residues Ser144 (2.25 Å) and two (02) H-bonds with Thr190 (2.62 Å, 2.22 Å) via hydroxyl groups. In addition, M9 built C-H bond with Gln189 and five (05) hydrophobic interactions (Pi-Alkyl and Alkyl) was formed with Cys145 (4.52 Å, 5.07 Å), Met165 (5.05 Å), His41 (4.70 Å) and Leu27 (4.31 Å) (**Table 3, Appendix 7)**.

M10, 3-Methoxycarpachromene, is a tetracyclic flavone **(Table 2),** from *Pistacia atlantica. Desf*, of chemical formula $C_{21}H_8O_6$, and molecular weight 336.37 g/mol. It has the lightest molecular weight among other inhibitors, for this reason, it is well anchored in the active site cavity as seen in **Figure 14 a) b),** thus, all rings contribute in inhibition mechanism and stabilizing the complex Mpro-M10. Exemplify the best inhibitor, consuming the least amount energy -8.2 Kcal/mol, the MD shows that 100% of poses are in the active site. Four (04) H-bond interactions [Asn142 (2.31 Å), Glu166 (2.34 Å), Gln192 (2.34 Å) and Thr190 (2.34 Å)] were observed when M10 was docked to the active site of Mpro. Mpro-M10 complex

was further stabilized by a single hydrophobic interaction, (Met165 (4.54 Å)] **(Table 3, Figure 14 c)).**

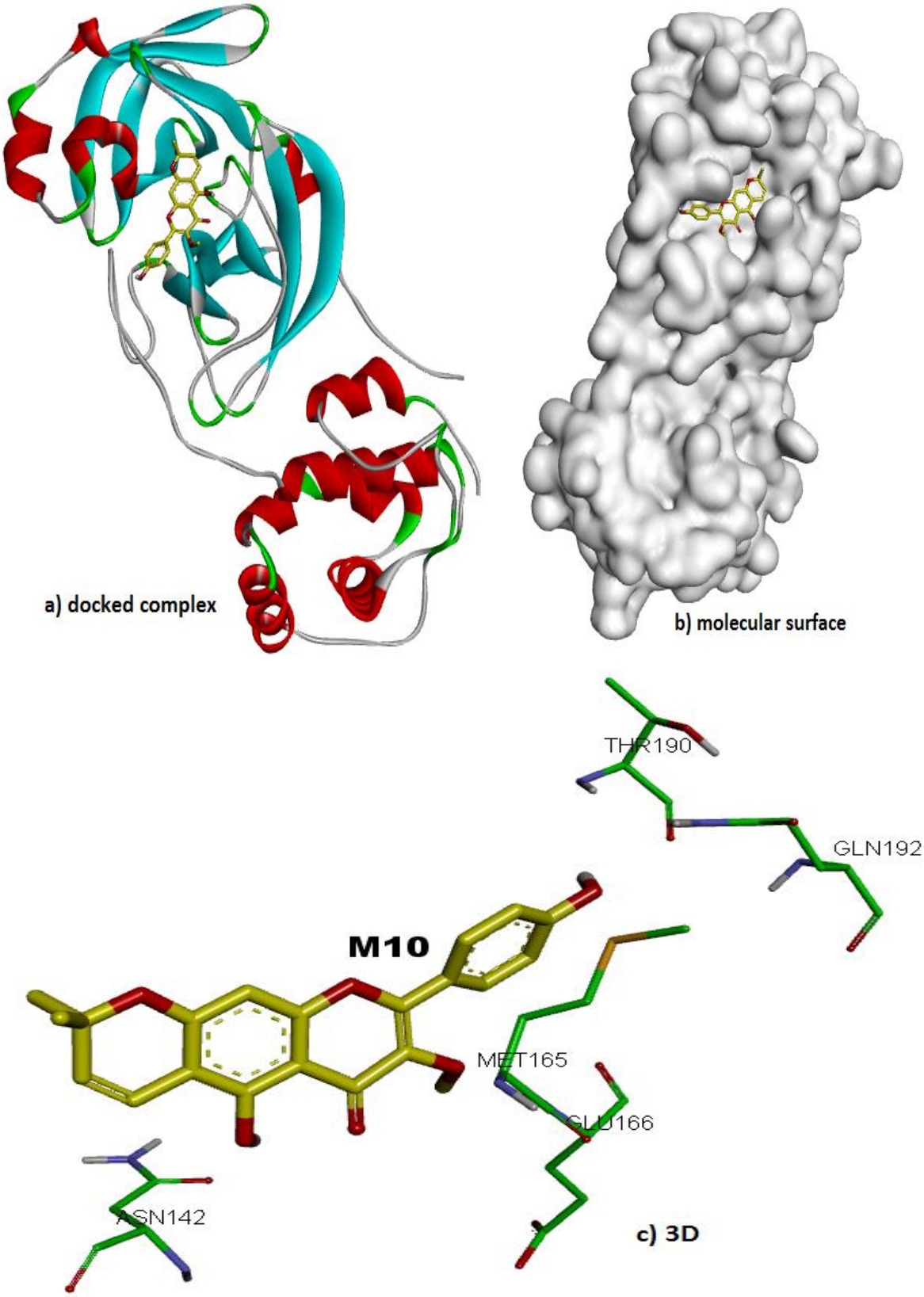


Figure 14. Molecular docking interaction of 3-Methoxycarpachromene with Mpro (6LU7) (Catalytic amino acids in green, ligand in yellow).

Based on the analysis of the previous studies, most of the binding pocket amino acids are hydrophilic, we propose that all the tested inhibitors in this study bind 70% to the hydrophilic residue in the Mpro active site (Glu, Arg, Asn, Gln, His, Thr, Ser). The amino acids: His41, Cys145 and Glu166 are important residues in the substrate-binding for the proteolytic activity, we believe that the availability and/or accessibility of Cys145, glu166 as well as His41 of Mpro may be ceased down due to the formation of multiple bonds (hydrogen or hydrophobic) between compounds and these three residues of Mpro can inhibit its catalytic activity **(Serseg et *al*., 2020)**. Thus, it can be concluded that M2, M4, M10 and M7, may possibly inhibit the proteolytic activity of Mpro, but to confirm its usefulness to treat patients with COVID-19 we will analyze ADMET results.

2. ADMET analysis

In terms of absorption, results revealed that the Human intestinal absorption (HIA) used to determine drugs intestinal absorption to track drugs transport targets **(Lee et *al*., 2003)**, as well as M7, M8, M9 and M10 were all more than 50% indicating that these compounds had good intestinal absorption. Greater HIA denotes that the compound could be better absorbed from the intestinal tract upon oral administration. M1, M2, M3, M4, M5 and M6 were all lesser than 50%, suggesting that these compounds had strong polarity due to glycosidic cycles and were not easily absorbed by the body. With regards to HIA, absorbance of less than 30% is considered to be poorly absorbed. Drugs that absorb poorly when taken orally must be administered in some less desirable way, like intravenously or by inhalation (e.g. zanamivir). Routes of administration are an important consideration. Absorption critically determines the compound's bioavailability **(Han et *al*., 2019).** The results also reveal that M1 and M2 had better Buffer solubility score than other compounds, estimated by 386.45 and 228.25 mg/l, respectively. M8 and M9 displays the lowest Buffer solubility score, estimated by 0.86 and 2.60 mg/l respectively. For blood–brain barrier membrane permeability that is formed by endothelial cells and prevents non-selective molecule to enters the central nervous system **(Lee et *al*., 2003; Lefauconnier et Hauw., 1984),** logBBB > 0.3, the compounds were thought to cross the blood–brain barrier easily. A logBBB < -1 suggested that the compounds did not easily cross the blood–brain barrier. All tested ligands have ability to pass through the blood brain barrier **(Han et *al*., 2019).** With regards to skin permeability, that investigates the effects of drug penetration across the skin barrier **(Khan et *al*., 2015),** the log Kp> -2.5, the compound is considered to be relatively low skin permeability **(Han et *al*., 2019).** In terms of metabolism, the present results expressed that all compounds have inhibitory effect for cytochrome P450 3A4 enzymes and

could be metabolized by P450 CYP 3A4, the enzyme responsible for biotransformation for about 90% of prescribed drugs **(Lynch et Price, 2007)**. M7 could behave either as substrate or inhibitor for P450 CYP3A4 enzyme, and may be metabolized in the liver. AMES toxicity test is employed to know whether a compound is mutagenic or not **(Ames et *al*., 1973; Nisha et *al*., 2016; Zeiger, 2019).** The results also suggest that M1, M2, M8 and M10 may be mutagen in AMES test but M3, M4, M5, M6, M7 and M9 may be non-mutagen. M1, M2, M3, M4, M6 and M8 may inhibit the hERG channel unlike the others which inhibit with a medium risk, human ether-a-go-go (hERG) channel which is a protein reported responsible for modulating the heart function by reducing the K+ from cardiac cells **(Lynch et Price, 2007).**

From docking analysis and prediction of drug likeness and pharmacokinetic properties of ten tested polyphenols, Lonchocarpol A (M7) and 3-Methoxycarpachromene (M10) exhibited high binding affinity toward COVID–19 Mpro and expressed good drug likeness and pharmacokinetic properties. Therefore, it may represent a potential treatment option for COVID-19.

GENERAL CONCLUSION AND PERSPECTIVES

COVID-19 is an evolving disease triggered by SARS-CoV-2 and has turned into a global pandemic in the blink of eyes. Until date, there is no specific antiviral drug. However, there are many ongoing clinical trials evaluating potential treatments. SARS-CoV-2 Mpro is shown to be highly potent and vital target for the inhibition of COVID-19 contamination.

The aim of this study was to examine some bioactive phenolic compounds that may be used to inhibit the main protease and stop the SARS-CoV-2 life cycle and its replication.

The presented study screened *in silico* the biological activity of ten polyphenols as antiviral components against the SARS-Cov-2 main protease. Our results revealed that flavonoids such as lonchocarpol A and 3-Methoxycarpachromene have a potential inhibition with a strong binding affinity to SARS-CoV-2 Mpro, and better ADMET profile, predicted to restrain the action of SARS-CoV-2 Mpro, thereby obstructing further translation of viral protein that assists in damaging the vital organs of the host, which may help in developing optimized potent SARS-CoV-2 inhibitors.

In Algeria, we have a lot of effective medicinal plants that could help in treating COVID-19, we need to do further studies (*in vitro* and *in vivo*) of our best compounds as inhibitors, on the other hand, screening for new potent compounds with more effectiveness and affinities, showing less secondary effects and a good ADMET profile.

BIBLIOGRAPHIC REFERENCES

Abouelela, M. E., Assaf, H. K., Abdelhamid, R. A., Elkhyat, E. S., Sayed, A. M., Oszako, T., ... et Abdelkader, M. S. A. (2021). Identification of Potential SARS-CoV-2 Main Protease and Spike Protein Inhibitors from the Genus Aloe: An In Silico Study for Drug Development. *Molecules*, *26*(6), 1767.

Adem, S., Eyupoglu, V., Sarfraz, I., Rasul, A., et Ali, M. (2020). Identification of potent COVID-19 main protease (Mpro) inhibitors from natural polyphenols: an in silico strategy unveils a hope against CORONA.

Ahn, J., Kim, Y. M., Chae, H. S., Choi, Y. H., Ahn, H. C., Yoo, H., ... et Chin, Y. W. (2019). Prenylated flavonoids from the roots and rhizomes of Sophora tonkinensis and their effects on the expression of inflammatory mediators and proprotein convertase subtilisin/kexin type 9. *Journal of natural products*, *82*(2), 309-317.

Al-Sadeq, D. W., et Nasrallah, G. K. (2020). The incidence of the novel coronavirus SARS-CoV-2 among asymptomatic patients: a systematic review. *International Journal of Infectious Diseases*, *98*, 372-380

Alexander E. Gorbalenya, Susan C. Baker, Ralph S. Baric, Raoul J. de Groot, Christian Drosten, Anastasia A. Gulyaeva, Bart L. Haagmans, Chris Lauber, Andrey M. Leontovich, Benjamin W. Neuman, Dmitry Penzar, Stanley Perlman, Leo L.M. Poon, Dmitry V. Samborskiy, Igor A. Sidorov, Isabel Sola, John Ziebuhr, (2020). The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature microbiology*, *5*(4), 536.

Ames, B. N.; Durston, W. E.; Yamasaki, E.; Lee, F. D., Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. *Proceedings of the National Academy of Sciences of the United States of America* 1973, *70* (8), 2281-5.

Anka, A. U., Tahir, M. I., Abubakar, S. D., Alsabbagh, M., Zian, Z., Hamedifar, H., ... et Azizi, G. (2020). Coronavirus disease 2019 (COVID-19): An overview of the immunopathology, serological diagnosis and management. *Scandinavian Journal of Immunology*, e12998.

Arya, R., Kumari, S., Pandey, B., Mistry, H., Bihani, S. C., Das, A., ... et Kumar, M. (2021). Structural insights into SARS-CoV-2 proteins. *Journal of molecular biology*, *433*(2), 166725.

Astuti, I. (2020). Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. *Diabetes et Metabolic Syndrome: Clinical Research et Reviews*, *14*(4), 407-412.

Attia, G. H., Moemen, Y. S., Youns, M., Ibrahim, A. M., Abdou, R., et El Raey, M. A. (2021). Antiviral zinc oxide nanoparticles mediated by hesperidin and in silico comparison study between antiviral phenolics as anti-SARS-CoV-2. *Colloids and Surfaces B: Biointerfaces*, *203*, 111724.

Azim, K. F., Ahmed, S. R., Banik, A., Khan, M. M. R., Deb, A., et Somana, S. R. (2020). Screening and druggability analysis of some plant metabolites against SARS-CoV-2: An integrative computational approach. *Informatics in Medicine Unlocked*, *20*, 100367.

Basu, A., Sarkar, A., et Maulik, U. (2020). Molecular docking study of potential phytochemicals and their effects on the complex of SARS-CoV2 spike protein and human ACE2. *Scientific reports*, *10*(1), 1-15.

Benamar, H., Marouf, A., et Bennaceur, M. (2018). Phytochemical composition, antioxidant and acetylcholinesterase inhibitory activities of aqueous extract and fractions of Pistacia atlantica subsp. atlantica from Algeria. *Journal of Herbs, Spices et Medicinal Plants*, *24*(3), 229-244.

Bendifallah, L., Tchoulak, Y., Djouabi, M., Oukili, M., et Ghezraoui, R. (2015). Phytochemical Study and Antimicrobial Activity of Origanum vulgare L.(Lamiaceae) in Boumerdes Mountainous Region (Algeria). *Journal of Medical and Bioengineering Vol*, *4*(6).

Benvenuto, D., Giovanetti, M., Ciccozzi, A., Spoto, S., Angeletti, S., et Ciccozzi, M. (2020). The 2019-new coronavirus epidemic: evidence for virus evolution. *Journal of medical virology*, *92*(4), 455-459.

Bordallo, B., Bellas, M., Cortez, A. F., Vieira, M., et Pinheiro, M. (2020). Severe COVID-19: what have we learned with the immunopathogenesis? *Advances in Rheumatology*, *60*(1), 1-13.

BOURAS, N., HACHEMI, A. (2019). Etude préliminaire des activités biologiques (insecticide et antifongique) des huiles essentielles de deux plantes aromatiques Thymus sp. Et Origanum sp.

Mémoire de fin d'étude. Université Abdelhamid Ibn Badis-Mostaganem Faculté des Sciences de la Nature et de la Vie.

Broadbent, A., Combrink, H., et Smart, B. (2020). COVID-19 in South Africa. *Global Epidemiology*.

Chebouti-Meziou, N., Merabet, A., Chebouti, Y., Bissaad, F. Z., Behidj-Benyounes, N., et Doumandji, S. (2014). Effect of cold and scarification on seeds germination of Pistacia atlantica L. for rapid multiplication.

Coccimiglio, J., Alipour, M., Jiang, Z. H., Gottardo, C., et Suntres, Z. (2016). Antioxidant, antibacterial, and cytotoxic activities of the ethanolic Origanum vulgare extract and its major constituents. *Oxidative medicine and cellular longevity*, *2016*.

Damas, J., Hughes, G. M., Keough, K. C., Painter, C. A., Persky, N. S., Corbo, M., ... et Lewin, H. A. (2020). Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proceedings of the National Academy of Sciences*, *117*(36), 22311-22322.

Das, S., Sarmah, S., Lyndem, S., et Singha Roy, A. (2020). An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *Journal of Biomolecular Structure and Dynamics*, 1-11.

de Freitas, K. S., Squarisi, I. S., Acésio, N. O., Nicolella, H. D., Ozelin, S. D., Reis Santos de Melo, M., ... et Tavares, D. C. (2020). Licochalcone A, a licorice flavonoid: antioxidant, cytotoxic, genotoxic, and chemopreventive potential. *Journal of Toxicology and Environmental Health, Part A*, *83*(21-22), 673-686.

de Oliveira, M. D. L., et de Oliveira, K. M. T. (2021). Comparative computational study of SARS-CoV-2 receptors antagonists from already approved drugs.

Dhama, K., Khan, S., Tiwari, R., Sircar, S., Bhat, S., Malik, Y. S., ... et Rodriguez-Morales, A. J. (2020). Coronavirus disease 2019–COVID-19. *Clinical microbiology reviews*, *33*(4), e00028-20.

Di Gennaro, F., Pizzol, D., Marotta, C., Antunes, M., Racalbuto, V., Veronese, N., et Smith, L. (2020). Coronavirus diseases (COVID-19) current status and future perspectives: a narrative review. *International journal of environmental research and public health*, *17*(8), 2690.

Elrashdy, F., Redwan, E. M., et Uversky, V. N. (2020). Why COVID-19 transmission is more efficient and aggressive than viral transmission in previous coronavirus epidemics? *Biomolecules*, *10*(9), 1312.

Ghosh, R., Chakraborty, A., Biswas, A., et Chowdhuri, S. (2020). Computer aided identification of potential SARS CoV-2 main protease inhibitors from diterpenoids and biflavonoids of Torreya nucifera leaves. *Journal of Biomolecular Structure and Dynamics*, 1-16.

Gimeno, A., Mestres-Truyol, J., Ojeda-Montes, M. J., Macip, G., Saldivar-Espinoza, B., Cereto-Massagué, A., ... et Garcia-Vallvé, S. (2020). Prediction of novel inhibitors of the main protease (M-pro) of SARS-CoV-2 through consensus docking and drug reposition. *International journal of molecular sciences*, *21*(11), 3793.

Giovanetti, M., Benedetti, F., Campisi, G., Ciccozzi, A., Fabris, S., Ceccarelli, G., ... et Ciccozzi, M. (2021). Evolution patterns of SARS-CoV-2: Snapshot on its genome variants. *Biochemical and Biophysical Research Communications*, *538*, 88-91.

Gourine, N., Yousfi, M., Bombarda, I., Nadjemi, B., Stocker, P., et Gaydou, E. M. (2010). Antioxidant activities and chemical composition of essential oil of Pistacia atlantica from Algeria. *Industrial Crops and Products*, *31*(2), 203-208.

Han, Y., Zhang, J., Hu, C. Q., Zhang, X., Ma, B., et Zhang, P. (2019). In silico ADME and toxicity prediction of ceftazidime and its impurities. *Frontiers in pharmacology*, *10*, 434.

Hanganu, D. A. N. I. E. L. A., Benedec, D. A. N. I. E. L. A., Olah, N. K., Ranga, F., Mirel, S. I. M. O. N. A., Tiperciuc, B. R. Î. N. D. U. Ș. A., et Oniga, I. (2020). Research on enzyme inhibition potential and phenolic compounds from Origanum vulgare ssp. vulgare. *Farmacia*, *68*, 1075-1080.

Harapan, H., Itoh, N., Yufika, A., Winardi, W., Keam, S., Te, H., ... et Mudatsir, M. (2020). Coronavirus disease 2019 (COVID-19): A literature review. *Journal of infection and public health*.

Hedley, P. L.; Jorgensen, P.; Schlamowitz, S.; Wangari, R.; Moolman-Smook, J.; Brink, P. A.; Kanters, J. K.; Corfield, V. A.; Christiansen, M., The genetic basis of long QT and short QT syndromes: a mutation update. *Human mutation* 2009, *30* (11), 1486-511.

Hui, K. P., Cheung, M. C., Perera, R. A., Ng, K. C., Bui, C. H., Ho, J. C., ... et Chan, M. C. (2020). Tropism, replication competence, and innate immune responses of the coronavirus SARS-CoV-2 in human respiratory tract and conjunctiva: an analysis in ex-vivo and in-vitro cultures. *The Lancet Respiratory Medicine*, *8*(7), 687-695.

Jan, S., Rashid, M., Abd_Allah, E. F., et Ahmad, P. (2020). Biological Efficacy of Essential Oils and Plant Extracts of Cultivated and Wild Ecotypes of Origanum vulgare L. *BioMed research international*, *2020*.

Jang, D. S., Cuendet, M., Hawthorne, M. E., Kardono, L. B., Kawanishi, K., Fong, H. H., ... et Kinghorn, A. D. (2002). Prenylated flavonoids of the leaves of Macaranga conifera with inhibitory activity against cyclooxygenase-2. *Phytochemistry*, *61*(7), 867-872.

Jo, S., Kim, S., Shin, D. H., et Kim, M. S. (2020). Inhibition of SARS-CoV 3CL protease by flavonoids. *Journal of enzyme inhibition and medicinal chemistry*, *35*(1), 145-151.

Juckel, D., Dubuisson, J., et Belouzard, S. (2020). Les coronavirus, ennemis incertains. *Médecine/sciences*, *36*(6-7), 633-641.

Khaerunnisa, S., Kurniawan, H., Awaluddin, R., Suhartati, S., et Soetjipto, S. (2020). Potential inhibitor of COVID-19 main protease (Mpro) from several medicinal plant compounds by molecular docking study.

Khailany, R. A., Safdar, M., et Ozaslan, M. (2020). Genomic characterization of a novel SARS-CoV-2. *Gene reports*, *19*, 100682.

Khan, N. R.; Harun, M. S.; Nawaz, A.; Harjoh, N.; Wong, T. W., Nanocarriers and their Actions to Improve Skin Permeability and Transdermal Drug Delivery. *Current pharmaceutical design* 2015, *21* (20), 2848-66.

Kneller, D. W., Phillips, G., O'Neill, H. M., Jedrzejczak, R., Stols, L., Langan, P., ... et Kovalevsky, A. (2020). Structural plasticity of SARS-CoV-2 3CL M proactive site cavity revealed by room temperature X-ray crystallography. *Nature communications*, *11*(1), 1-6.

Krim, S., Rihani, R., Marchal, L., Foucault, A., Bentahar, F., et Legrand, J. (2021). Two-phase solvent extraction of phenolics from Origanum vulgare subsp. glandulosum. *Journal of Applied Research on Medicinal and Aromatic Plants*, *20*, 100273.

Labdelli, A., Rebiai, A., Tahirine, M., Adda, A., et Merah, O. (2020). Nutritional content and antioxidant capacity of the seed and the epicarp in different ecotypes of Pistacia atlantica desf. subsp. atlantica. *Plants*, *9*(9), 1065.

Lagorce, D., Douguet, D., Miteva, M. A., et Villoutreix, B. O. (2017). Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. *Scientific reports*, *7*(1), 1-15.

Lee, S.; Lee, I.; Kim, H.; Chang, G.; Chung, J.; No, K., The PreADME Approach: Web-based program for rapid prediction of physico-chemical, drug absorption and drug-like properties. *EuroQSAR designing drugs and crop protectants: processes, problems and solutions* 2003, 418-20.

Lefauconnier, J. M.; Hauw, J. J., [The blood-brain barrier. II. Physiological data (conclusion)]. *Revue neurologique* 1984, *140* (2), 89-109.

Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., ... et Farzan, M. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*, *426*(6965), 450-454.

Li, W., Zhang, C., Sui, J., Kuhn, J. H., Moore, M. J., Luo, S., ... et Farzan, M. (2005). Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *The EMBO journal*, *24*(8), 1634-1643.

Ligaj, M., Kobus-Cisowska, J., Szczepaniak, O., Szulc, P., Kikut-Ligaj, D., Mikołajczak-Ratajczak, A., ... et Jarzębski, M. (2021). Electrochemical screening of genoprotective and antioxidative effectiveness of Origanum vulgare L. and its functionality in the prevention of neurodegenerative disorders. *Talanta*, *223*, 121749.

LINANI, A., BENAROUS, K., et Yousfi, M. (2020). Novel Structural Mechanism of Glutathione as a potential peptide inhibitor to the main protease (Mpro): CoviD-19 treatment, Molecular docking and SAR study.

Liu, C., Zhou, Q., Li, Y., Garner, L. V., Watkins, S. P., Carter, L. J., ... et Albaiu, D. (2020). Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases.

Lynch, T.; Price, A., The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *American family physician* 2007, *76* (3), 391-6.

Maamri, S., Benarous, K., et Yousfi, M. (2021). Identification of 3-Methoxycarpachromene and Masticadienonic Acid as New Target Inhibitors against Trypanothione Reductase from Leishmania Infantum Using Molecular Docking and ADMET Prediction. *Molecules*, *26*(11), 3335.

Mahjoub, F., Rezayat, K. A., Yousefi, M., Mohebbi, M., et Salari, R. (2018). Pistacia atlantica Desf. A review of its traditional uses, phytochemicals and pharmacology. *Journal of medicine and life*, *11*(3), 180.

Mahrosh, H. S., et Mustafa, G. (2021). An in silico approach to target RNA-dependent RNA polymerase of COVID-19 with naturally occurring phytochemicals. *Environment, Development and Sustainability*, 1-14.

Meng, L., Hua, F., et Bian, Z. (2020). Coronavirus disease 2019 (COVID-19): emerging and future challenges for dental and oral medicine. *Journal of dental research*, *99*(5), 481-487.

Mousavizadeh, L., et Ghasemi, S. (2020). Genotype and phenotype of COVID-19: Their roles in pathogenesis. *Journal of Microbiology, Immunology and Infection*.

Mpiana, P. T., Tshibangu, D. S., Kilembe, J. T., Gbolo, B. Z., Mwanangombo, D. T., Inkoto, C. L., ... et Tshilanda, D. D. (2020). Identification of potential inhibitors of SARS-CoV-2 main protease from Aloe vera compounds: a molecular docking study. *Chemical Physics Letters*, *754*, 137751.

Nisha, C. M., Kumar, A., Nair, P., Gupta, N., Silakari, C., Tripathi, T., et Kumar, A. (2016). Molecular docking and in silico ADMET study reveals acylguanidine 7a as a potential inhibitor of β-secretase. *Advances in bioinformatics*, *2016*.

Oniga, I., Pușcaș, C., Silaghi-Dumitrescu, R., Olah, N. K., Sevastre, B., Marica, R., ... et Hanganu, D. (2018). Origanum vulgare ssp. vulgare: Chemical composition and biological studies. *Molecules*, *23*(8), 2077.

Peter, E. K., et Schug, A. (2021). The inhibitory effect of a coronavirus spike protein fragment with ACE2. *Biophysical Journal*, *120*(6), 1001-1010.

Pillaiyar, T., Meenakshisundaram, S., et Manickam, M. (2020). Recent discovery and development of inhibitors targeting coronaviruses. *Drug discovery today*, *25*(4), 668-688.

Platto, S., Wang, Y., Zhou, J., et Carafoli, E. (2021). History of the COVID-19 pandemic: origin, explosion, worldwide spreading. *Biochemical and biophysical research communications*, *538*, 14-23.

Puttaswamy, H., Gowtham, H. G., Ojha, M. D., Yadav, A., Choudhir, G., Raguraman, V., ... et Chauhan, L. (2020). In silico studies evidenced the role of structurally diverse plant secondary metabolites in reducing SARS-CoV-2 pathogenesis. *Scientific reports*, *10*(1), 1-24.

Rao, G. V., Mukhopadhyay, T., Annamalai, T., Radhakrishnan, N., et Sahoo, M. R. (2011). Chemical constituents and biological studies of Origanum vulgare Linn. *Pharmacognosy research*, *3*(2), 143.

Rosales-Mendoza, S., Márquez-Escobar, V. A., González-Ortega, O., Nieto-Gómez, R., et Arévalo-Villalobos, J. I. (2020). What does plant-based vaccine technology offer to the fight against COVID-19?. *Vaccines*, *8*(2), 183.

Said, S. A., Fernandez, C., Greff, S., Derridj, A., Gauquelin, T., et Mevy, J. P. (2011). Inter-population variability of leaf morpho-anatomical and terpenoid patterns of Pistacia atlantica Desf. ssp. atlantica growing along an aridity gradient in Algeria. *Flora-Morphology, Distribution, Functional Ecology of Plants*, *206*(4), 397-405.

Salvatore, M. J., King, A. B., Graham, A. C., Onishi, H. R., Bartizal, K. F., Abruzzo, G. K., ... et Witherup, K. M. (1998). Antibacterial activity of Lonchocarpol A. *Journal of natural products*, *61*(5), 640-642.

Serseg, T., Benarous, K., et Yousfi, M. (2020). Hispidin and Lepidine E: two Natural Compounds and Folic acid as Potential Inhibitors of 2019-novel coronavirus Main Protease (2019-nCoVMpro), molecular docking and SAR study. *arXiv preprint arXiv:2004.08920*.

Singh, S., Sk, M. F., Sonawane, A., Kar, P., et Sadhukhan, S. (2020). Plant-derived natural polyphenols as potential antiviral drugs against SARS-CoV-2 via RNA-dependent RNA polymerase (RdRp) inhibition: An in-silico analysis. *Journal of Biomolecular Structure and Dynamics*, 1-16.

Taleghani, N., et Taghipour, F. (2020). Diagnosis of COVID-19 for controlling the pandemic: A review of the state-of-the-art. *Biosensors and Bioelectronics*, 112830.

Toul, F., Belyagoubi-Benhammou, N., Zitouni, A., et Atik-Bekkara, F. (2017). Antioxidant activity and phenolic profile of different organs of Pistacia atlantica Desf. subsp. atlantica from Algeria. *Natural product research*, *31*(6), 718-723.

Umar, H. I., Josiah, S. S., Saliu, T. P., Jimoh, T. O., Ajayi, A., et Danjuma, J. B. (2021). In-silico analysis of the inhibition of the SARS-CoV-2 main protease by some active compounds from selected African plants. *Journal of Taibah University Medical Sciences*, *16*(2), 162-176.

Vijayakumar, B. G., Ramesh, D., Joji, A., et Kannan, T. (2020). In silico pharmacokinetic and molecular docking studies of natural flavonoids and synthetic indole chalcones against essential proteins of SARS-CoV-2. *European journal of pharmacology*, *886*, 173448.

Vivianne Veenma. (2020). The Pathogenesis of SARS-CoV-2 The life cycle of SARS-CoV-2 and the consequence for the human body Master Thesis. University of Groningen, faculty of science and engineering.

Wong, N. A., et Saier, M. H. (2021). The SARS-coronavirus infection cycle: a survey of viral membrane proteins, their functional interactions and pathogenesis. *International journal of molecular sciences*, *22*(3), 1308.

Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., ... et Li, H. (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, *10*(5), 766-788.

X. Liu, B. Zhang, Z. Jin, H. Yang, Z. Rao, The crystal structure of covid-19 main protease in complex with an inhibitor n3 (2020).

Y**adav, M., Dhagat, S., et Eswari, J. S. (2020).** Emerging strategies on in silico drug development against COVID-19: challenges and opportunities. *European Journal of Pharmaceutical Sciences*, 105522.

Yesudhas, D., Srivastava, A., et Gromiha, M. M. (2020). COVID-19 outbreak: history, mechanism, transmission, structural studies and therapeutics. *Infection*, 1-15.

Yoshimoto, F. K. (2020). The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *The protein journal*, *39*, 198-216.

Yüce, M., Filiztekin, E., et Özkaya, K. G. (2020). COVID-19 diagnosis —A review of current methods. Biosensors and Bioelectronics, 112752.

Zeiger, E., The test that changed the world: The Ames test and the regulation of chemicals. *Mutation research* 2019, *841*, 43-48.

Zhang, H., Penninger, J. M., Li, Y., Zhong, N., et Slutsky, A. S. (2020). Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive care medicine*, *46*(4), 586-590.

Zhang, X. L., Guo, Y. S., Wang, C. H., Li, G. Q., Xu, J. J., Chung, H. Y., ... et Wang, G. C. (2014). Phenolic compounds from Origanum vulgare and their antioxidant and antiviral activities. *Food chemistry*, *152*, 300-306.

Webography

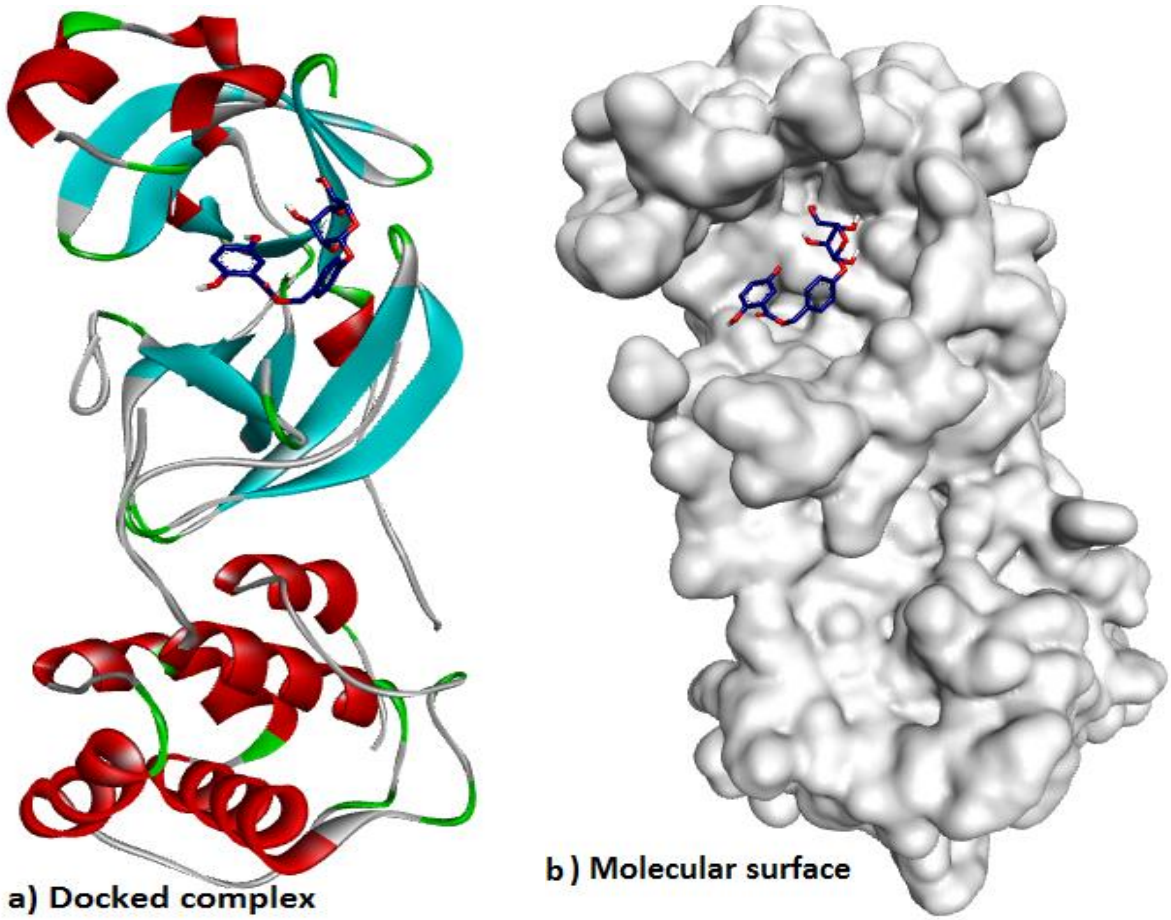
https://www.worldometers.info/coronavirus/ : consulted on 23-05-2021.
https://preadmet.bmdrc.kr/ : work with on 12-06-2021.
https://www.infectioncontroltoday.com/view/lab-leak-called-viable-possibility-for-covid-19-pandemic: consulted on 01-06-2021.
https://www.medicalnewstoday.com/articles/256521 : visited on 20-05-2021.
https://www.tela-botanica.org/ : visited on 23-06-2021.

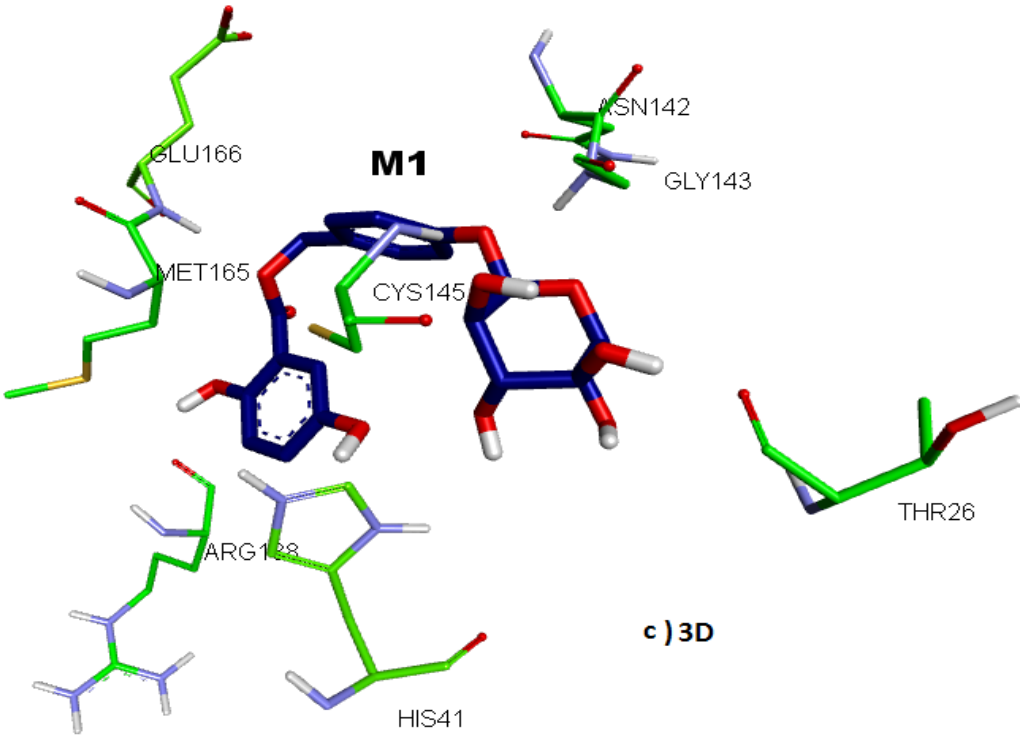
APPENDIX

Appendix 1. ADMET stand ranges.

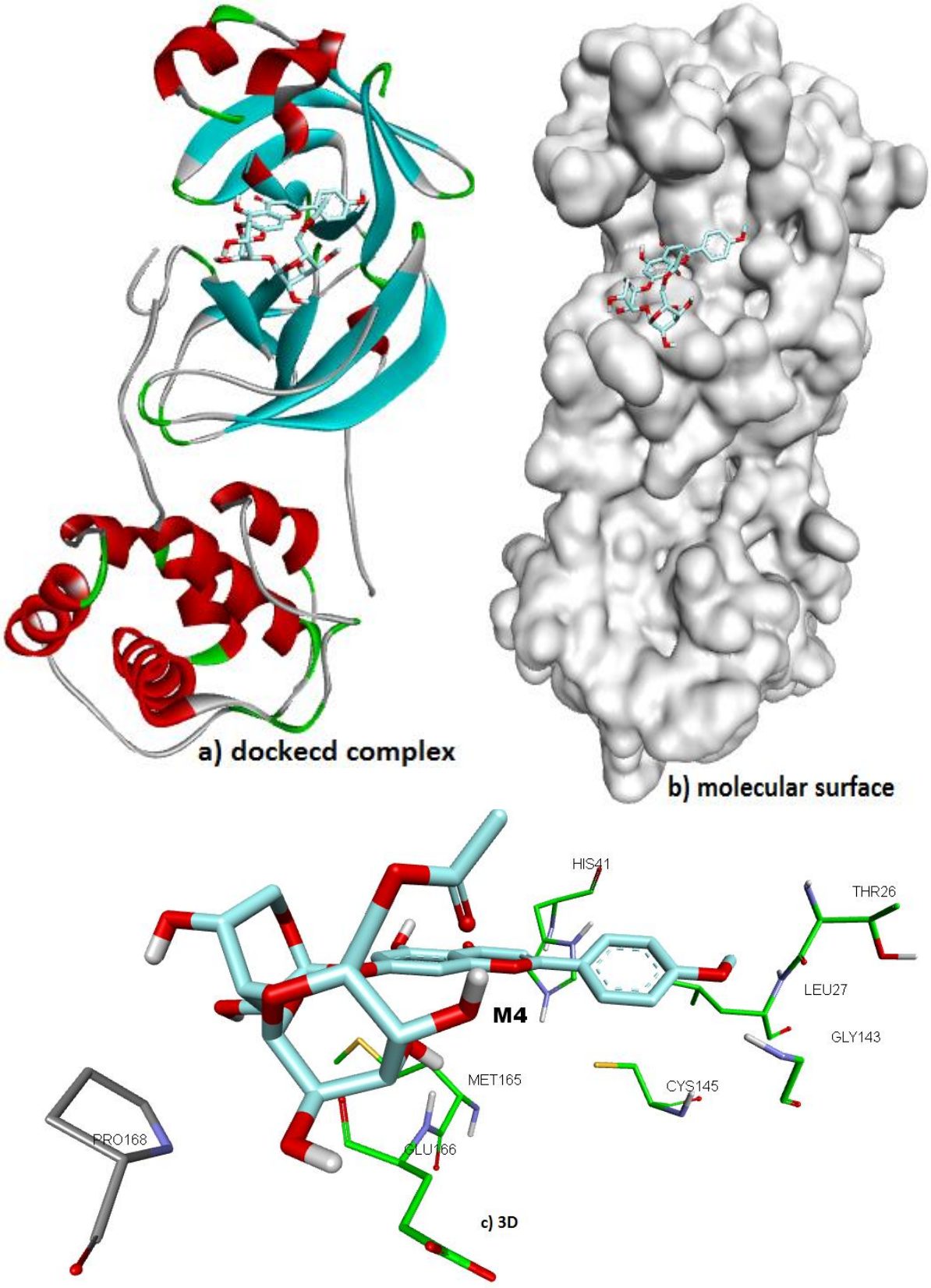
Absorption		
Human intestinal absorption (HIA %)	Norms ≈ 100%	Lee et *al*., 2003
Buffer solubility (mg/l)	Bigger better	/
Distribution		
Blood-brain barrier penetration (C.brain/C.blood)	logBBB > 0.3	Lee et *al*., 2003; Lefauconnier et *al*., 1984
Skin permeability (log Kp, cm/hour)	Log kp >-2.5	Khan et al., 2015
Metabolism		
Cytochrome P450 3A4 inhibition	Inhibitor (bad) / or not	Lynch et *al*., 2007
Cytochrome P450 3A4 substrate	Substrate (good) / or not	
Toxicity		
HERG_inhibition	Risky (bad) or not	Hedley et *al*., 2009
Ames test	Mutagen (bad) or not	Ames et *al*., 1973; Zeiger, 2019.

Appendix 2. Molecular docking interaction of 4-[[(2' ,5' Dihydroxybenzoyl)oxy]methyl]phenyl O-b-D-glucopyranoside with Mpro (6LU7) (Catalytic amino acids are in green, ligand is in dark blue).

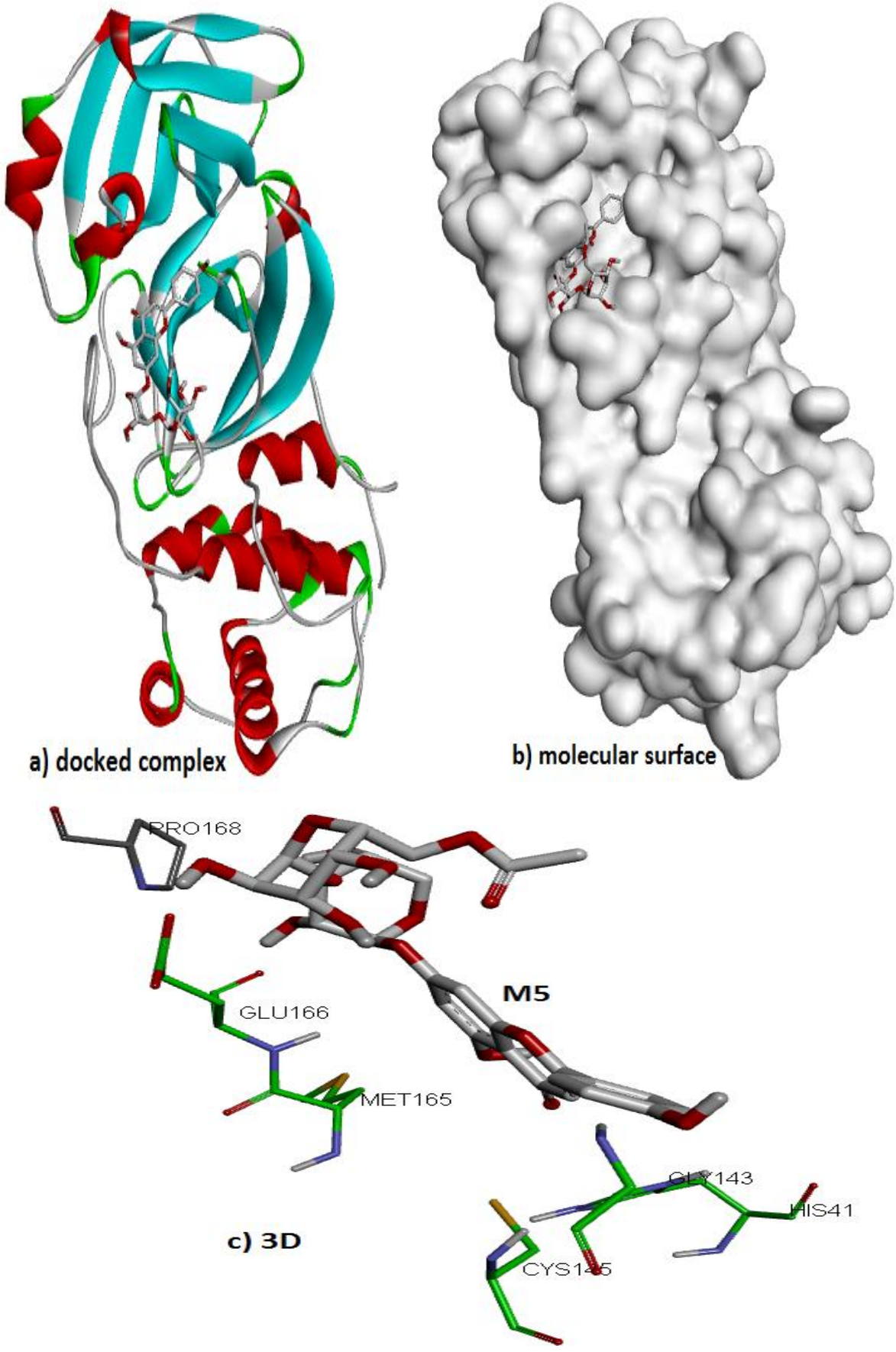



GLU166
M1
ASN142
GLY143
MET165
CYS145
THR26
ARG188
HIS41
c) 3D

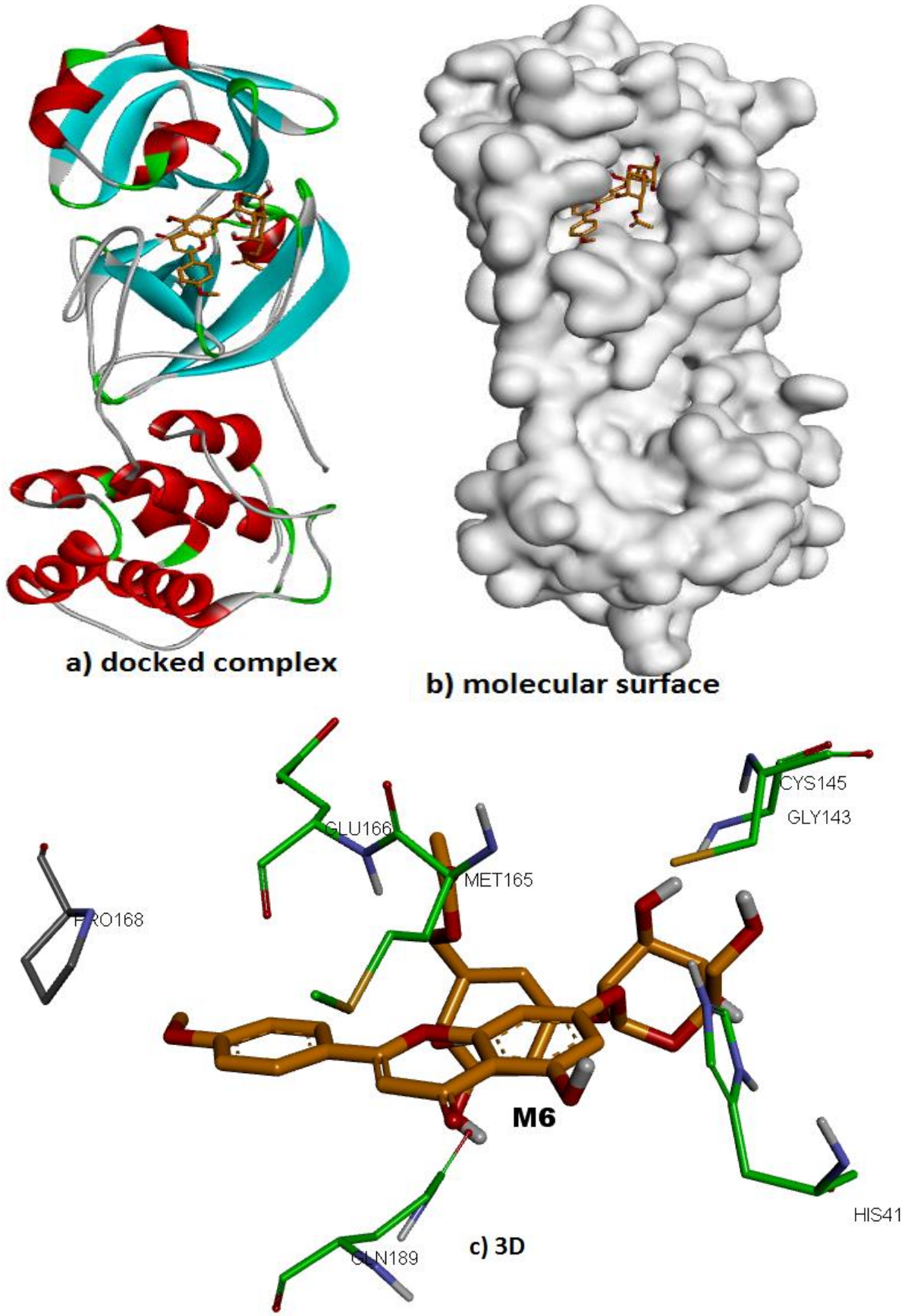
Appendix 3. Molecular docking interaction of Acacetin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside with Mpro (6LU7) (Catalytic amino acids in green, ligand is in blue).



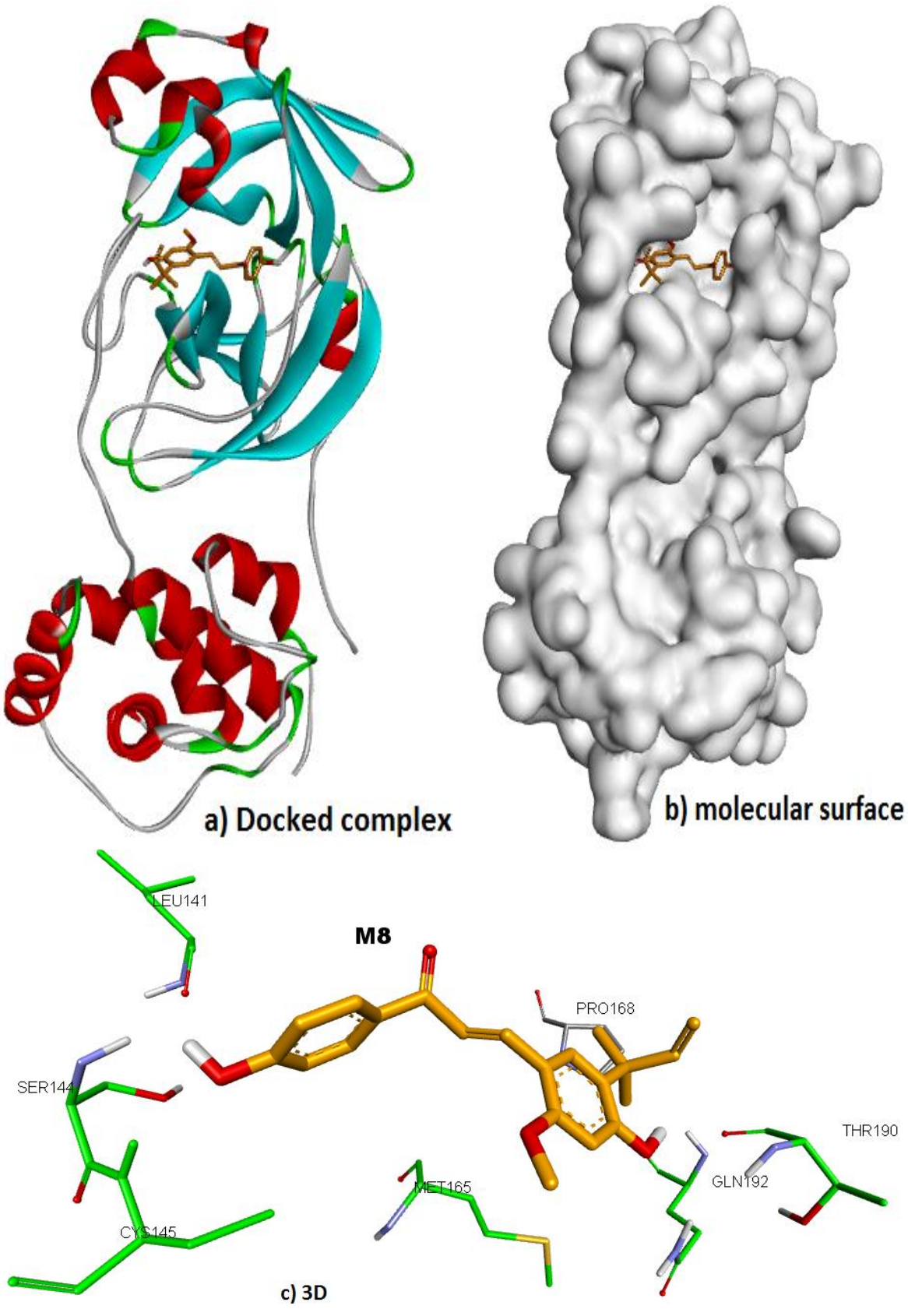
Appendix 4. figure represent molecular docking interaction of Apigenin 7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>3)]-b-Dxylopyranoside with Mpro (6LU7) (Catalytic amino acids in green, ligand in white).



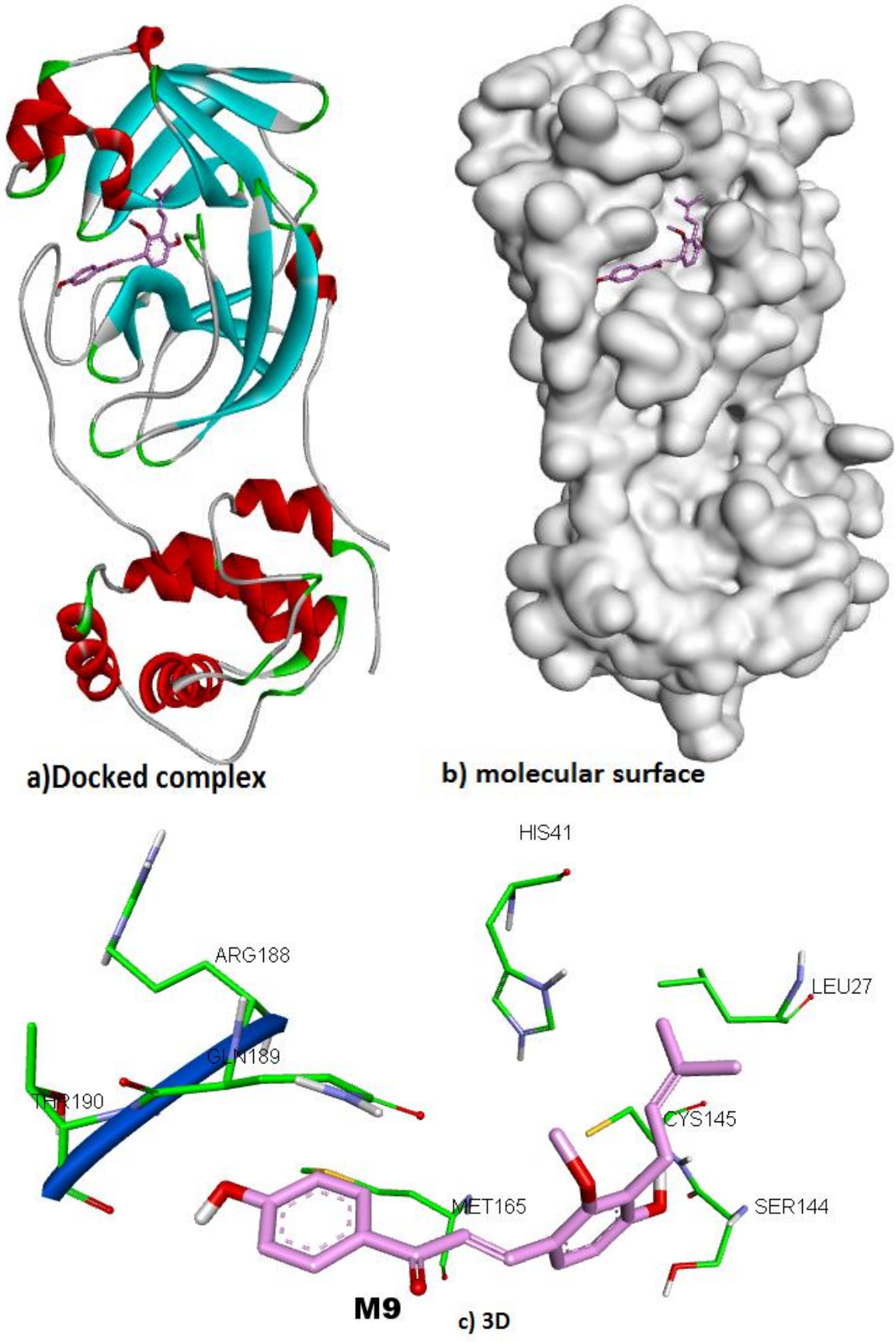
Appendix 5. Molecular docking interaction of Acacetin-7-O-[6’’’-O-acetyl-b-D-galactopyranosyl-(1=>2)]-bD-glucopyranoside with Mpro (6LU7) (Catalytic amino acids in green, ligand in orange).



Appendix 6. Molecular docking interaction of Licochalcone A with Mpro (6LU7) (Catalytic amino acids in green, ligand in yellow).



Appendix 7. Molecular docking interaction Licochalcone A with Mpro (6LU7) (Catalytic amino acids in green, ligand in purple).



Abstract

SARS-CoV-2, a rapidly spreading new strain of coronavirus, has affected more than 170 countries and received worldwide attention. Several clinical trials are underway to identify specific drugs for the treatment of this novel virus. The crystallized form of SARS-CoV-2 main protease (Mpro) was demonstrated by a Chinese researcher, which is a therapeutic drug target, its inhibition is necessary for the blockage of the viral replication. The present study aimed to evaluate the efficacy of ten medicinal plant-based phenolic compounds as SARS-CoV-2 Mpro (PDB ID: 6LU7, Resolution 2.16 Å) inhibitors, using a molecular docking with ADMET analysis. The modeling of these polyphenols were performed using Autodock Vina, Audotock tool and Discovery Studio Visualizer programs. ADMET properties were calculated using PreADMET prediction online server. The results show that Lonchocarpol A and 3-Methoxycarpachromene exhibited high binding affinity than others, toward SARS-CoV-2 Mpro, and expressed good pharmacokinetic properties. Therefore, they may represent a potential treatment option for COVID-19. Further researches are urgently required to investigate their potential uses in COVID-19 treatment.
Keywords: SARS-CoV-2, COVID-19, polyphenols, Molecular docking, ADMET Analysis, Mpro.

Résumé

Le SRAS-CoV-2, une nouvelle souche de coronavirus à propagation rapide, a touché plus de 170 pays et a attiré l'attention du monde entier. Plusieurs essais cliniques sont en cours pour identifier des médicaments spécifiques pour le traitement de ce nouveau virus. La forme cristallisée de la protéase principale du SRAS-CoV-2 (Mpro) a été démontrée par un chercheur chinois, qui est une cible thérapeutique médicamenteuse, son inhibition est nécessaire au blocage de la réplication virale. La présente étude visait à évaluer l'efficacité de dix composés phénoliques à base de plantes médicinales en tant qu'inhibiteurs de Mpro du SRAS-CoV-2 (PDB ID : 6LU7, Résolution 2.16 Å), en utilisant un amarrage moléculaire. L'amarrage moléculaire de ces polyphénols a été effectué à l'aide des programmes Autodock Vina, Audoock tool et Discovery studio Visualizer. Les propriétés d'ADMET ont été calculées à l'aide d'un site Web de prédiction PreADMET. Les résultats de l'analyse d'amarrage, il montre que le Lonchocarpol A et le 3-méthoxycarpachromène présentent une affinité de liaison élevée par rapport aux autres inhibiteurs vis-à-vis de Mpro du SRAS-CoV-2 Mpro et expriment de bonnes propriétés pharmacocinétiques. Par conséquent, ils peuvent représenter une option de traitement potentielle pour COVID-19. Des recherches supplémentaires sont instamment nécessaires pour étudier leurs utilisations potentielles dans le traitement du COVID-19.
Mot clés : SARS-CoV-2, COVID-19, Amarrage moléculaire, polyphénols, Analyse ADMET, Mpro

.

ملخص

اثرت السلالة جديدة من فيروس كورونا "سارس كوفيد 2" سريع الانتشار، على أكثر من 170 دولة واكتسب اهتمامًا عالميًا. تجرى العديد من التجارب السريرية لتحديد أدوية معينة لعلاج هذا الفيروس الجديد. تم توضيح الشكل المتبلور للبروتياز الرئيسي لسارس كوفيد 2 من قبل باحث صيني ، وهو هدف دوائي علاجي ، وتثبيطه ضروري لمنع تكاثر هذا الفيروس. هدفت الدراسة الحالية إلى تقييم فعالية عشر مركبات فينولية عشبية كمثبطات للبروتياز الرئيسي لسارس كوفيد2 ، باستخدام الارساء الجزيئي. تم إجراء الارساء الجزيئي لهذه البوليفينول باستخدام اوتودوك فينا ، اوتودوك تولس و ديسكوفري ستوديو فيزيباليزور. تم حساب خصائص الامتصاص ، التوزيع ، التمثيل الغذائي ، الإخراج والسمية باستخدام موقع بري ادمت. نتائج تحليل الارساء ، يبدو واعدًا أن لونشوكاربول ا ، و 3 ميثوكسي كارباكرومن ، يظهران خاصية عاليّة للربط مقارنةً بالمثبطات الأخرى للبروتياز الرئيسي ويعبران عن خصائص حركية دوائية جيدة. لذلك قد يمثلون خيارًا علاجيًا محتملاً لكوفيد 19 ، هناك حاجة ماسة إلى مزيد من البحث لدراسة استخداماتهم المحتملة في علاج كوفيد19 .
الكلمات المفتاحية : سارس كوفيد 2 ، كوفيد 19 ، الارساء الجزيئي ، مركبات فينولية تحليل الامتصاص ، التوزيع ، التمثيل الغذائي ، الإخراج والسمية ، للبروتياز الرئيسي.