الجمـــهوريـــة الجـــزائريـــة الديـمقراطيـــــــة الشـــعبيـــــة
THE PEOPLE'S DEMOCRATIC REPUBLIC OFALGERIA
وزارة التعـلـيـم الـعـالـــي والـبـحـث الـعـلـمي
MINISTRY OF HIGHER EDUCTION AND SCIENTIFIC RESEACH
جـامـعـة عمّـــار ثـلـيـــجـــي بالأغـــواط
AMAR TELIDJI UNIVERSITY-LAGHOUAT
كـلـيـــــة الـعـلـــــوم
FACULTY OF SCIENCES
قسم علوم المادة
DEPARTMENT OF MATERIAL SCIENCES

MASTER DISSERTATION

Field :MATERIAL SCIENCES
Section: Chemistry
Option: Applied Organic Chemistry
Presentedby :

CHEGMA Taha and ADI Nawal

THEME

Molecular docking analysis of selected plant-based compounds for inhibition of SARS-CoV-2 main protease

Sustained in front of the jury composed of:

Mr. DJERIDANE Amar	*professor*	*Chairman*
M^{me}. MAHFOUDI Reguia	*M.C.A*	*Examiner*
Mr.YOUSFI Mohammed	*Professor*	*Supervisor*
Mr. LINANI Abderahmane	*Doc*	*Co- Supervisor*

University year2020 -2021

Dedication

I dedicate this work to:

My parents: to my mother, who's always pushing me to look forward, who's sacrificed everything for me and was by my side in my entire life.

To my father. Who is my hero and role model in life, who has always helped me and held my hand.

To my brother and my sister, who are a part of me, who had supported me and put their faith on me.

To my grandmother and my grandfather.

To all my uncles and my aunts.

To my best friend especially and also to all my friends.

Chegma Taha

Dedication

I dedicate this work to ***My parents.***

To ***my mother*** *who 's the dearest and most precious person in my life who stirred my path with her advice, and was a clear sea flowing with an abundance of love, and smile to the one who decorated my life with the full moon and candles of joy.*

To ***my father*** *who gave me strength to continue the path,*

they are the reason to continue my studies, who taught me patience and diligence, to the one dear to my heart.

To the ***soul of my dear grandmother*** *who always accompanies me with her kind prayers.*

To my beloved ***sisters*** *and my* ***brother*** *, To all my* ***uncles and my aunts.***

To my ***best friends*** *who spent an hours of studying and laughing together.*

Adi Nawal

Acknowledgments

First of all, we would like to thank ALLAH, our creators for given us the strength to accomplish this work.

We would like to thank our professor YOUSFI Mohammed and Mr.LINANI Abderahmane For their tremendous efforts with us, for every information they gave us, and for every letter they taught us

We would like to thank Mr. DJERIDANE Amar and M^{me}. MAHFOUDI Reguia for agreeing to examine this work.

We also thank all the teachers of the material science department for their contribution to this training. Thanks to all those who contributed to us kindly from near or far to the development from Thesis.

LIST OF FIGURES:

LIST OF TABLES:

LIST OF ABREVIATIONS

RCSB: Research Collaboratory for Structural Bioinformatics.

PDB: Protein Data Bank.

WHO: World Health Organization.

MERS –CoV: Middle East Respiratory Syndrome CoronaVirus.

SARS –CoV: Severe Acute Respiratory Syndrome CoronaVirus.

nCoV: A new strain novel corona virus.

ACE2: Angiotensin Converting Enzyme 2.

M^{pro}: Main Protease.

EO: Essential Oil.

GC: Gas Chromatography.

MS: Mass Spectroscopic.

Table of contents:

GENERAL

INTRODUCTION

Coronaviruses vary significantly in risk factor. Some can kill more than 30% of those infected, such as MERS-CoV, and some are relatively harmless, such as the common cold.(1)Coronaviruses can cause colds with major symptoms, such as fever, and a sore throat from swollen adenoids.(2)Coronaviruses can cause pneumonia and bronchitis. The human coronavirus discovered in 2003, SARS-CoV, which causes severe acute respiratory syndrome (SARS), has a unique pathogenesis because it causes both upper and lower respiratory tract infections.(3)

Six species of human coronaviruses are known, with one species subdivided into two different strains, making seven strains of human coronaviruses altogether.

Three human coronaviruses produce symptoms that are potentially severe:

- Severe acute respiratory syndrome coronavirus (SARS-CoV), β-CoV (identified in 2003)
- Middle East respiratory syndrome-related coronavirus (MERS-CoV), β-CoV (identified in 2012)
- Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), β-CoV (identified in 2019).

Coronavirus disease 2019 (COVID-19), also known as the coronavirus and COVID, is a contagious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The first known case was identified in Wuhan, China, in December 2019.(4)

The disease has since spread worldwide. The virus that causes COVID-19 spreads mainly when an infected person is in close contact with another person small droplets and aerosols containing the virus can spread from an infected person's nose and mouth as they breathe, cough, sneeze, sing, or speak. Other people are infected if the virus gets into their mouth, nose or eyes. The virus may also spread via contaminated surfaces, although this is not thought to be the main route of transmission. The exact route of transmission is rarely proven conclusively, but infection mainly happens when people are near each other for long enough. People who are infected can transmit the virus to another person up to two days before they themselves show symptoms, as can people who do not experience symptoms. People remain infectious for up to ten days after the onset of symptoms in moderate cases and up to 20 days in severe cases.(5)

On3 July 2021 the WHO stated that COVID-19 was characterized as a pandemic. To date, more than 185 million people have contracted this disease, and more than 4 million have died because of it in the world.

The virus's protein spikes attach to a protein on the surface of cells, called ACE2. Normally, ACE2 plays a role in regulating blood pressure. But when the coronavirus binds to it, it sets off chemical changes that effectively fuse the membranes around the cell and the virus together, allowing the virus's RNA to enter the cell.

The virus hacks the host cell's protein-making machinery to translate its RNA into new copies of the virus. In just hours, a single cell can be forced to produce tens of thousands of new viruses, which then infect other healthy cells.(6)**Figure 1** shows how the virus attacks the cell.

Parts of the virus's RNA also code for proteins that stay in the host cell. At least three are known. One prevents the host cell from sending out signals to the immune system that it's under attack. Another encourages the host cell to release the newly created viruses. And another helps the virus resist the host cell's innate immunity.

Infection is a race between the virus and the immune system. The outcome of that race depends on where it starts: the milder the initial dose, the more chance the immune system has of overcoming the infection before the virus multiplies out of control. The relationship between symptoms and the number of viruses in the body, though, remains unclear.

If an infection sufficiently damages the lungs, they will be unable to deliver oxygen to the rest of the body, and a patient will require a ventilator. The CDC estimates that this happens to between 3% and 17% percent of all covid-19 patients. Secondary infections that take advantage of weakened immune systems are another major cause of death.(7)

Sometimes it is the body's response that is most damaging. Fevers are intended to cook the virus to death, but prolonged fevers also degrade the body's own proteins. (7)

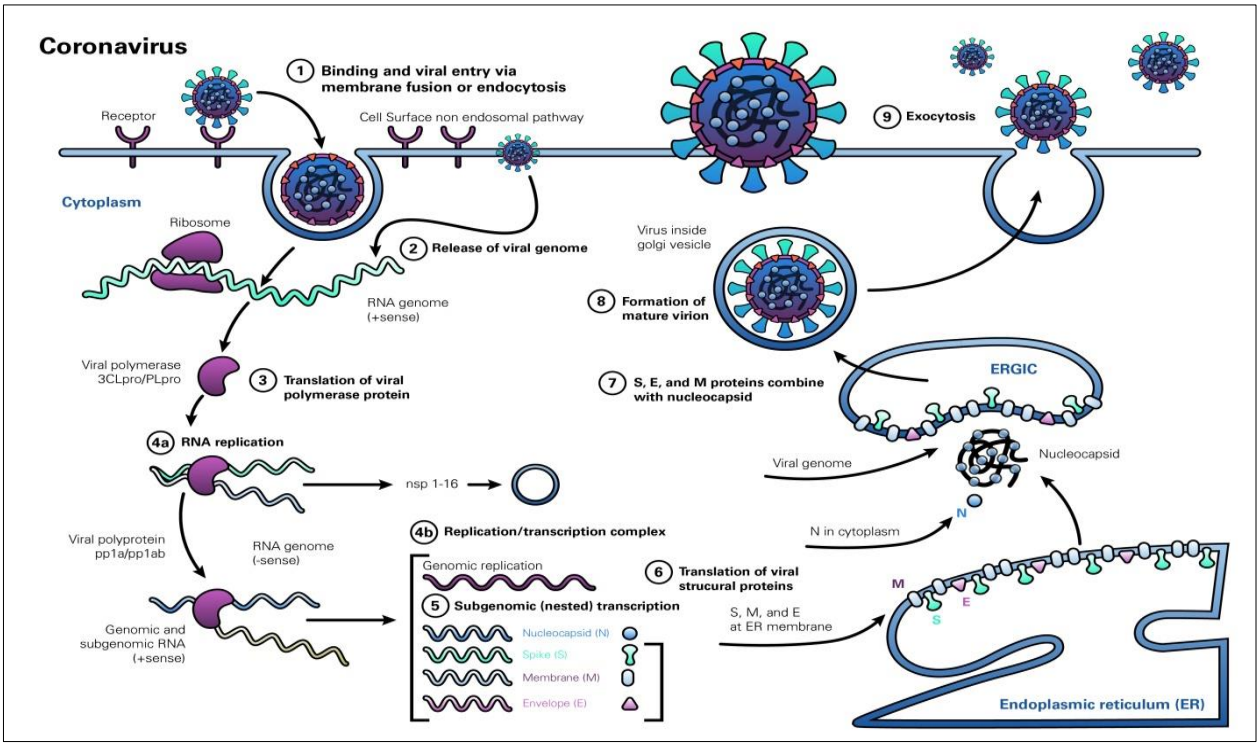


Figure 1: Life cycle of SARS-CoV-2.

For thousands of years, mankind has used various plants found in its environment as food and also to treat and care for all kinds of diseases. These plants represent an immense reservoir of compounds which have the advantage of having a great diversity of chemical structure. The choice of plants is based on the one hand on their richness in metabolites secondary (polyphenols, flavonoids, essential oils.) while knowing that the latter are produced exclusively by plants and on the other hand on their traditional use known and frequent in our populations.

Odorous product, generally of complex composition, obtained from a material first botanically defined plant, either by steam entrainment or otherwise process. It is a mixture of various molecules, including in particular terpenes, and oxygenated compounds (alcohols, aldehydes, ketones). Obtaining essential oils by cold expression (zest). In the latter case, a certain ambiguity exists on the denomination of essential oil. For this type of extract, the term gasoline is also used. Essential oils can be extracted from different parts of the plant: - Flowers (rose petals), - Fruit peel (lemon, bergamot, orange), - Seeds (anise), - Leaves (eucalyptus), - Berries (juniper), - Floral buds (clove), - Fruits (parsley), - Wood (sandalwood, cinchona bark). The essential oil content of plants is low, around 1 to 3% with the exception of the nail cloves (14 to 19%), mace (10 to 13%), nutmeg (8 to 9%), and cardamom (4 to 10%). EOs are usually classified according to the chemical nature of the major active ingredients, more rarely on the method of extraction, or the biological effects. At retains 8 main classes: - Sesquiterpene carbides - Terpene carbides, - Alcohols, - Esters and alcohols, - Aldehydes, - Ketones, - Phenols, - ethers and peroxides).(8)

Docking in silico aims to predict the structure of a molecular complex from isolated molecules, which is considerably easier to implement, less expensive and faster than using one of the experimental methods mentioned above. Docking software is therefore very useful tools in biology, pharmacy and medicine, because most of the active ingredients are small molecules (ligand) which interact with a biological target of therapeutic interest, generally protein (receptor), in order to influence the mechanism in which this protein is involved.(9)The **Table 1**presentedsome of docking software.This thesis work ends with applications of docking in the design of new active molecules.

Table 1: Examples of software available for protein-ligand docking and their search algorithm.

Name	Search algorithm	Type	References
AUTODOCK4	Lamarckian genetic algorithm	Academic	Morris et al., 2009
DOCK	Shape matching	Academic	Allen et al., 2015
OEDOCKING	Shape matching	Academic	Kelley, Brown, Warren &Muchmore, 2015; Mcgann, 2011
FLEKSY	Ensemble-based	Commercial	Nabuurs, Wagener & De Vlieg, 2007; Wagener, De Vlieg&Nabuurs, 2012
SWITSSDOCK	Evolutionary optimization	Academic	A. Grosdidier, Zoete&Michielin, 2011
GOLD	Genetic algorithm	Commercial	Jones, Willett, Glen, Leach & Taylor, 1997
GLIDE	Hybrid	Commercial	Friesner et al., 2004
VINA	Local optimization	Academic	Trott& Olson, 2009
RDOCK	Hybrid	Academic	Ruiz-Carmona et al., 2014
LEDOCK	Simulated annealing	Academic	Unzue et al., 2016
PLANTS	Ant colony optimization	Academic	Korb, Stützle&Exner, 2009
HADDOCK	Hybrid	Academic	Dominguez, Boelens&Bonvin, 2003
SURFLX-DOCK	Shape matching	Commercial	Spitzer & Jain, 2012
MOE	Hybrid	Commercial	Vilar, Cozza& Moro, 2008

FLAXX	Shape matching	Commercial	Vilar, Cozza& Moro, 2008
FITTED	Hybrid	Commercial	Corbeil &Moitessier, 2009; De Cesco, Kurian, Dufresne, Mittermaier&Moitessier, 2017; Englebienne&Moitessier, 2009; Moitessier et al., 2016
LIGANDFIT	Shape matching	Commercial	Venkatachalam, Jiang, Oldfield & Waldman, 2003
ICM	Hybrid	Commercial	Neves, Totrov&Abagyan, 2012
IGEMDOCK	Evolutionary algorithm	Academic	Yang & Chen, 2004

In this study we have used molecular docking which is a computational method to predict the interaction of two molecules generating a binding model. In many drug discovery applications, docking is done between a small molecule and a macromolecule for example, protein-ligand docking. More recently, docking is also applied to predict the binding mode between two macromolecules, for instance protein-protein docking.(10)

On the basis of this study, we has just confirmed the benefaction effects of plants on the treatment of covid-19, a choice was made on 13 local medicinal aromatic plants. Therefore, we started this work by the extraction of essential oils and the identification of structural character of volatile compounds. Then, the effect of essential oils on the treatment of covid-19 is made by a theoretical method based on the inhibition of protease enzyme responsible for the virus.

MATERIALS AND METHODS

2.1 Materials:

2.1.1 Medicinal Plant:

Thirteen local plants were recolted in April 2021 and used in this study for extraction of the essential oils. The **Table 2** summarized some information of those plants.

Table 2: The medicinal plant that been chosen.

Systematic and uses	Photos
Scoparia deverra: Kingdom : *Plantae* Class : *Magnoliopsida* Family : *Apiaceae* Genus : *Deverra* DC Activities: Anti oxydant.(11) Arabic label : الكزيح Location :Sebseb –metlili Ghardaia Part : Twigs and flowers	
***Artemisia herba alba*:** Kingdom : *Plantae* Class : *Magnoliophyta* Family : *Asteraceae (Compositae)* Genus : *Artemisia* Activities: Widely used as medicinal plants in folk medicin analgesic, antibacterial, antispasmodic, and hemostatic agents.(12) Arabic label : الشيح العادي Location :Sebseb –metlili Ghardaia Part : Twigs	

***Citrus*(lemon)** Kingdom : *Plantae* Class : *Dicotyledonae* Family : *Rutaceae* Genus: *Citrus* Activities: Traditionally used to treat scurvy, sore throats, fevers, rheumatism, high blood pressure, and chest pain-antimicrobial.(13) Arabic label :الليمون Location :Laghouat Part : Peel	
***Citrus sinensis*(orang)** Kingdom : *Plantae* Class : *Dicotyledonae* Family : *Rutaceae* Genus: *Citrus* Activities: Especially during cold-anti-inflammatory. Arabic label : البرتقال Location :Laghouat Part : Peel	
Artemisia campestris Kingdom : *Plantae* Class : *Magnoliopsida.* Family: *Asteraceae.* Genus: *Artemisia.* Activities: Lost appetite, suffer from weaknesses in the digestive system, rheumatism Arabic label :الدقفت Location :Milak- laghouat Part :Twigs	

***Juniperus oxycedrus*:** Kingdom : *Plantae* Class :*Pinopsida* Family : *Cupressaceae* Genus: *Juniperus* Activities: Antioxidant. Arabic label :العرعار Location:Madena –oud mora –Aflou Laghouat Part :Twigs and flowers	
***Cupressuss empervirens* (Cypress):** Kingdom: *Plantae* Class: *Pinopsida* Family : *Cupressaceae* Genus : *Cupressus* Activity: Antileishmanial(14) Arabic label :السرو Location :Laghouat Part : Fruits	
***Artemisia annua*:** Kingdom :*Plantae* Class: *Magnoliopsida* Family : *Asteraceae* Genus : *Artemisia* Activity: Antioxidant-antimicrobial-antibacterial-antifungal.(15) Arabic label : الشيح الحولي Location:Laghouat Part :Leaves	

***Salvia rosmarinus*(rosemary):** Kingdom : *Plantae* Class: *Asterids* Family : *Lamiaceae* Genus: *Salvia* Activity: Antioxidant, Antimicrobial activity was tested against 13 bacterial strains and 6 fungi.(16) Arabic label :اكليل الجبل Location :Madena –oudmora –Aflou Laghouat Part : Leaves	
***Syzygium aromaticum*(Clove):** Kingdom : *Plantae* Class: *Dicotyledonae* Family: *Myrtaceae*. Genus : *Syzygium aromaticum* Activity: Antimicrobial and antioxidant(17) Arabic label :القرنفل Location :Bayed- laghouat Part : Fruits	
***Saccocalyx satureioides*:** Kingdom : *Plantas* Class: *Equisetopsida* Family : *Lamiaceae* Genus : *Saccocalyx* Activity: The antimicrobial activity of the oils against bacteria.(18) Arabic label :زعترالرمل Location : Bayed- laghouat Part : Twigs	

***Rhanterium adpressum*:** Kingdom : *Plantae* Class: *Magnoliopsida* Family :*Asteraceae* Genus : *Rhanterium* Activity: The essential oils have been used due to their anticancer, antinociceptive, antiphlogistic, antiviral, antibacterial, and antioxidant.(19) Arabic label : العرفج Location :Zelfana –Ghardaia Part : Twigs and flowers	
***Origanum vulgare*:** Kingdom: *Plantae* Class: *Magnoliopsida* Family: *Lamiaceae.* Genus: *Origanum.* Activity: Great potential of antimicrobial activity against bacteria-Antioxidant.(20) Arabic label : زعيترة Location :Djelfa Part : Twigs	

2.1.2 Bioinformatic (The selected proteases):

We have download the Main protease (M^{pro}) from RCSB PDB website that give as the structure of which has been crystallized in previse work than try to improve the result by funding a molecule that befitting more in active site and gives lowest affinity and distance(hydrogen bond).

- **5RUD:**

The X-Ray diffraction with resolution: 1.00 A°, shows that the protease 5RUD contain 3 unique types of molecules **Figure 2** shows the composition of 5RUD M^{pro}.

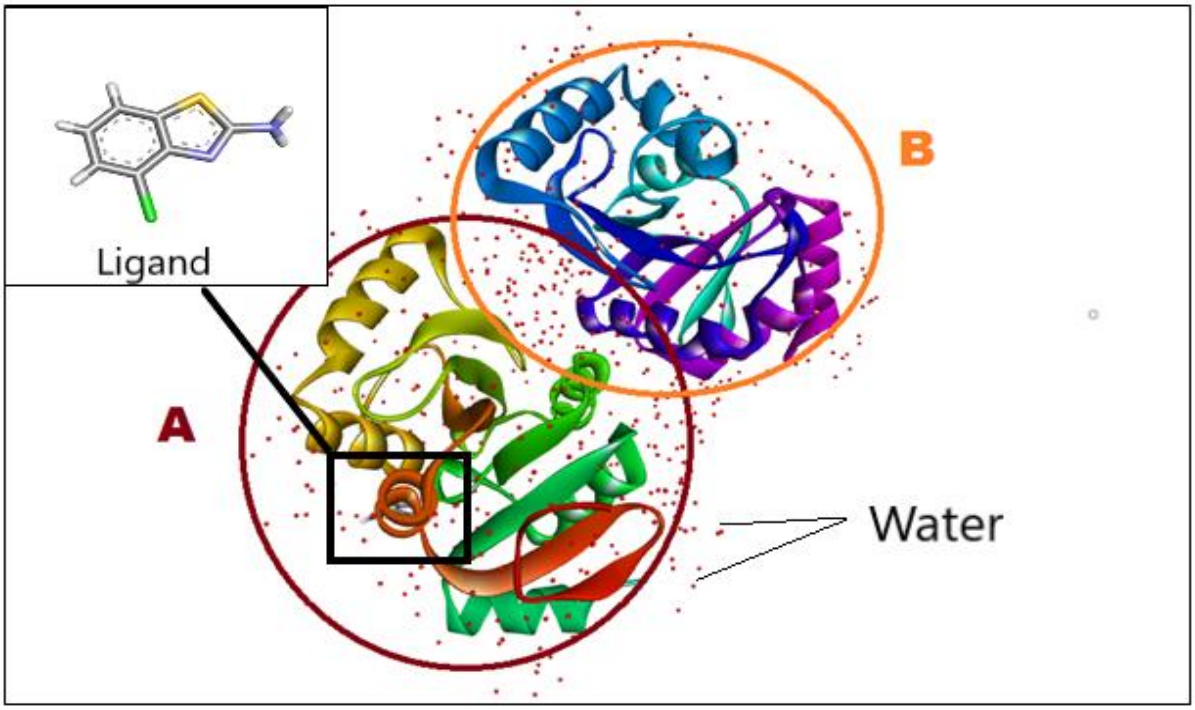


Figure 2: Simulation structure on 3D of M^{pro} [PDB:5RUD]

Table 3: Composition of 5RUD M^{pro}.

	Molecule	Residuse	Atoms
Protein	A	167	Total C H N O S 2558 810 1279 220 245 4
	B	169	Total C H N O S 2602 822 1305 221 249 5
Ligand	C7H5ClN2S	1	Total C Cl H N S 16 7 1 5 2 1
Water	H2O	521	Total O H 1563 521 1042

- **5RUN:**

By X-Ray diffraction with resolution: 1.00 A°, there are 3 unique types of molecules in this entry **Figure 3** shows the composition of 5RUN M^{pro}.

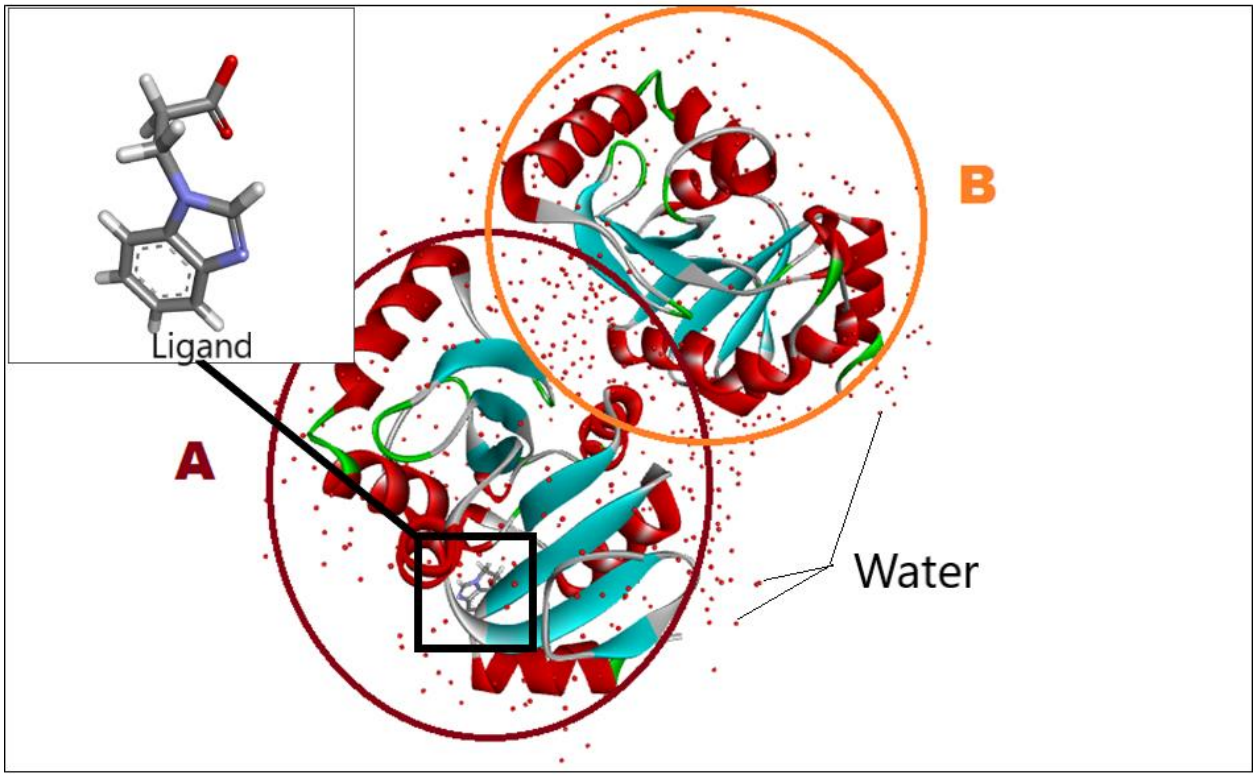


Figure 3: Simulation structure on 3D of M^{pro}[PDB:5RUN]

Table 4: Composition of 5RUN M^{pro}.

	Molecule	Residuse	Atoms
Protein	A	167	Total C H N O S 2558 810 1279 220 245 4
	B	169	Total C H N O S 2602 822 1305 221 249 5
Ligand	C10H10N2O2	1	Total C H N O 23 10 9 2 2
Water	H2O	502	Total O H 1506 502 1004

- **5RUE:**

By X-Ray diffraction with resolution: 1.02 A°, there are 3 unique types of molecules in this entry **Figure 4** shows the composition of 5RUE M^{pro}.

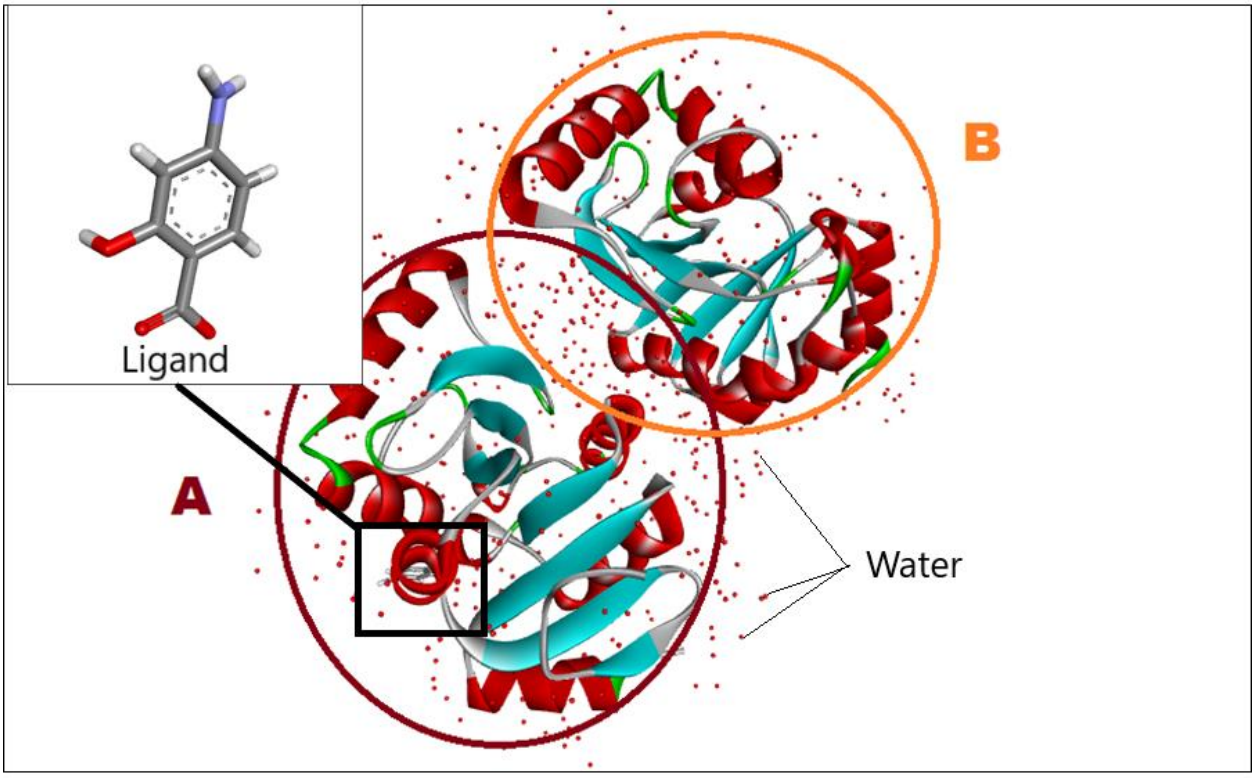


Figure4: Simulation structure on 3D of M^{pro} [PDB:5RUE]

Table 5: Composition of 5RUEM^{pro}..

	Molecule	Residuse	Atoms
Protein	1	167	Total C H N O S 2558 810 1279 220 245 4
	1	169	Total C H N O S 2602 822 1305 221 249 5
Ligand	C7H7NO3	1	Total C H N O 17 7 6 1 3
Water	H2O	489	Total O H 1467 489 978

2.2 Methods:

2.2.1 Extraction of essential oils:

We prepared the plants by drying them to get rid of moisture then we crushed100g and put it in balloon, add 800 ml of water, then start the Clevenger with in addition the temperature, the vapors charged with essential oils pass through the vertical tube, then in the refrigeration where the condensation will take place.

The droplets thus produced accumulate in the tube previously filled with distilled water then we add Sodium Anhydrate Sulfate to pull out all the water particles.

2.2.2The identification of essential oils compounds by GC:

The analysis of the samples was carried out in the Technical Platform of Physico-Chemical Analysis (PTAPC-CRAPC)-Laghouat-Algeria, using a SHIMADZU GCMS-QP2020 Instruments, equipped with a fused Rxi-5ms capillary column (Phase: Crossbond 5% diphenyl/ 95% dimethyl polysiloxane) its dimensions are: 30 m × 0.25 mm and 0.25 µm film thickness, this column has similar phase to the following columns: HP-1ms, HP-1msUI, DB-1ms, DB-5ms, DB-1msUI, Ultra-1, VF-1ms, ZB-1, ZB-1ms and considered also equivalent to USP G1, G2, G38 phases. A volume of 0.5 µL solution prepared by 10% vol. of the sample dilution in n-hexane, was injected in split mode (80:1). Injector and detector temperatures were maintained at 250°C and 310°C, respectively the column temperature was programmed at: 60°C fixed for 3min then increased to 310°C with an increase increment of 2°C/min, and then maintained at 310°C for 10 min. The carrier gas used was helium (99.995% purity) with a flow rate of 1 mL/min. The mass spectrometer conditions were as follow: ionization voltage 70 eV, ion source temperature 200°C, and electron ionization mass spectra were acquired over the mass range of 45-600 m/z.

Linear retention indices (LRI) were calculated for separate compounds relative to a homologous n-alkanes serial (n-C7-C20). Components were identified by comparison of their calculated (LRI) with those of literature (Babushok et al., 2011: https://doi.org/10.1063/1.3653552), (book of Adams Robert P, 2017, ed.4.1) as well as their mass spectra with those recorded by the NIST (National Institute of Standards and Technology) and Wiley libraries “NIST17.lib, W11N17MAI and FFNSC1.2.lib”.

2.2.3 Docking Molecular:

For the preparation of protease of covid-19: we have used PDB (protein data banc), which contain a million of proteases. **Figure 5** present RCSB Protein Data Bank or PDB website which is a global collection of structural data three-dimensional (3D structure) of biological macromolecules:

mainly proteins, and nucleic acids. These structures are essentially determined by crystallography X-ray or NMR spectroscopy.

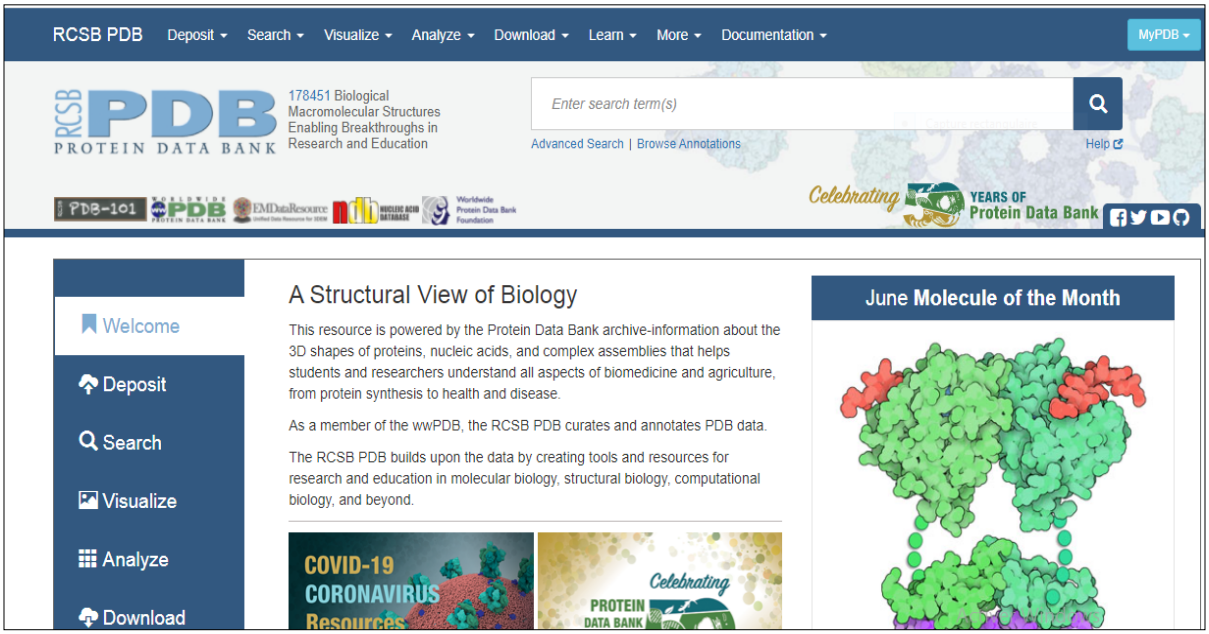


Figure 5: The web site RCSB PDB.

2.2.3.1 Preparation ofM^{pro}:

After download those tree $M^{pro}$from RCSD PDB we use **Discovery Studio 2020 client** program to disperse between ligand and protein. We downloaded the proteases that been chosen on the PDB version with their receptor and we separated the ligand from the protein we use software is discovery studio 2020 client, so that water molecules will be deleted, ligand and protein must be separated separately. After identification of molecules from essential oils, we use Chem-Spider as it shows in **Figure 6**. To download those molecules in 3D geometry.

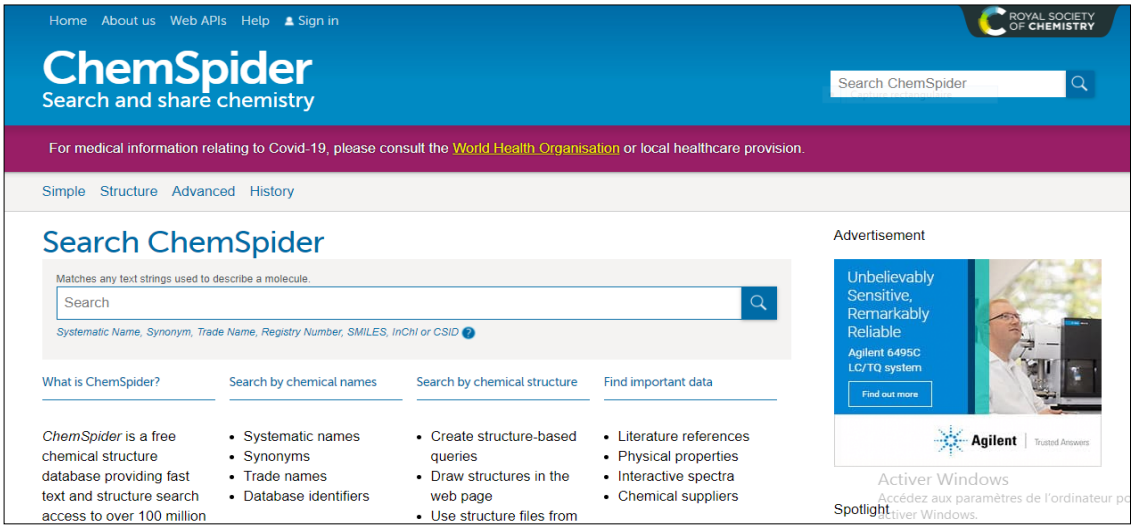


Figure 6: The web site chem-spider for molecules geometry.

2.2.3.2AMDock analysis:

In this part we start process of docking using this second software (**AMDock**) **Figure 7** which can find the decent position for ligand in active site and calculate the energy of this combination, the best suits position gives the lowest affinity as well as distance of hydrogen bonds.

For the origin ligand that has been crystallized with the protease will be redocked so can be taken as the reference for all our molecules of essential oils, tarrying to find which will give lest than the previous affinity, than among them we going to select small hydrogen bond.

We've tested the three M^{pro} with 13 plants by using the autudock4 in **AMDock** program with the majority of compounds in each essential oils. The tables shows the results of docking molecules:

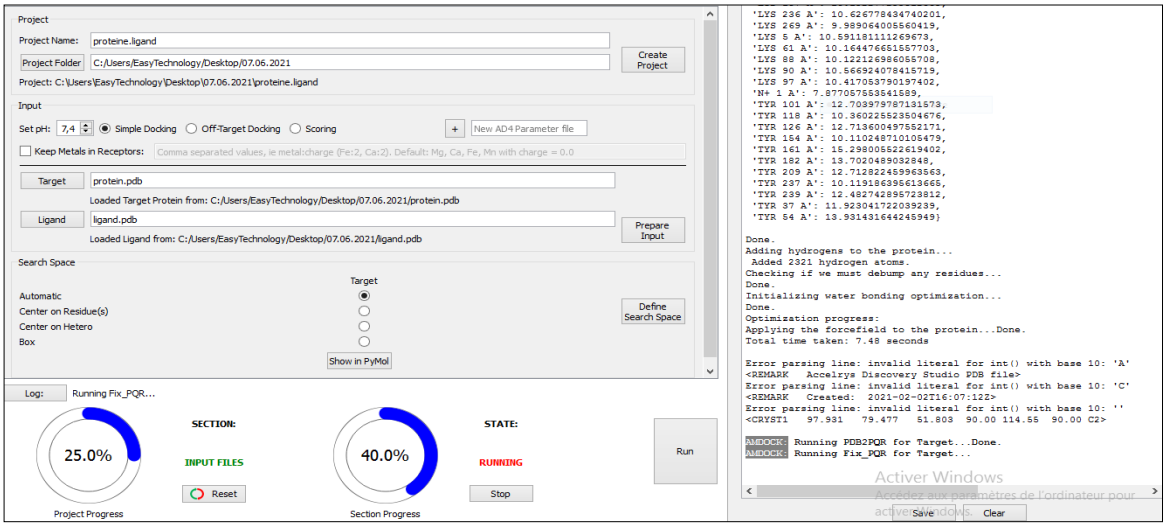


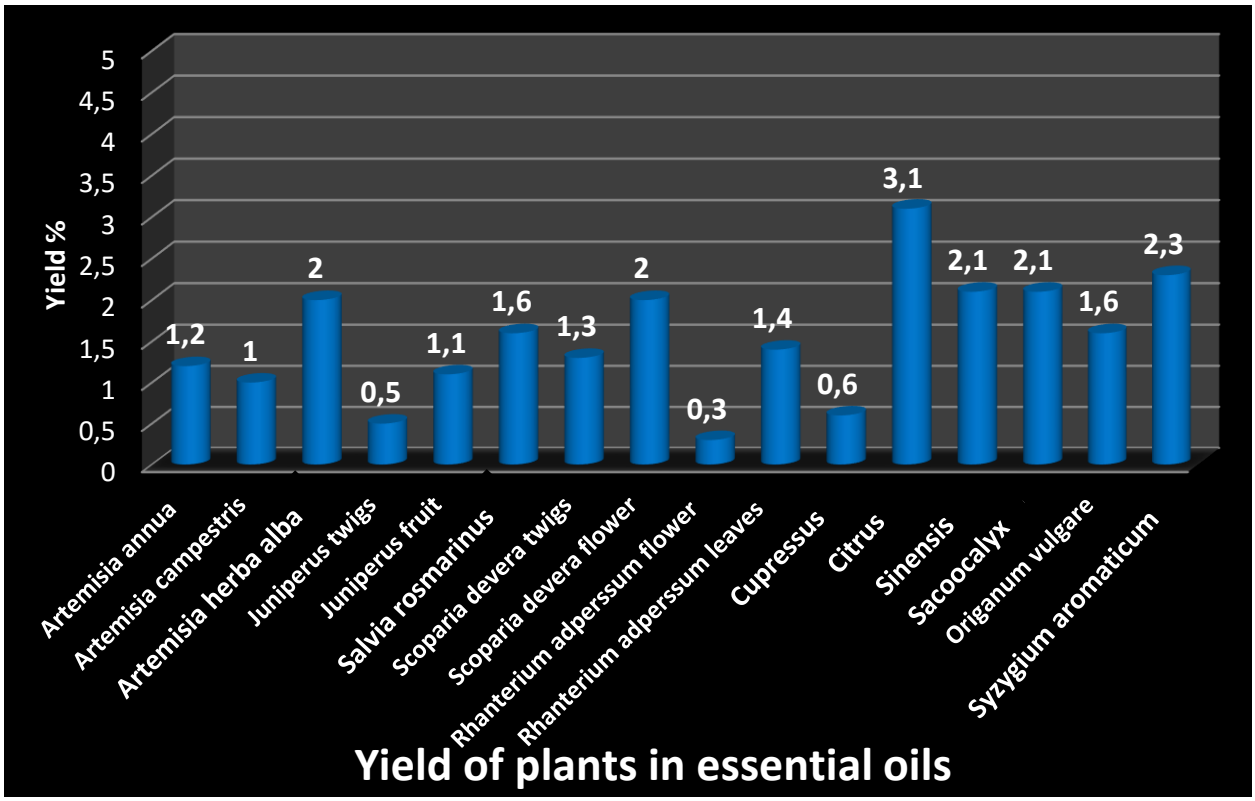
Figure 7: The interface of AMDock program.

RESULTS AND DISCUSSION

3. RESULTSANDDISCUSSION

3.1 Essential oils:

The usual method of extracting essential oils is still the Clevenger method, because it is easy to use and simple. A sufficient quantity of plant material (stems, leaves, and flowers).in this **diagram** we showing the yield that we get from each plant.



3.2Identification of majority compounds by GC:

In this **Table 6** we can see the different compounds that been identify by GC-MS.

Table 6: The constitution form essential oils of plants.

Constituent	Kovalt Index (IK)	Plants(Yield % in EOV/W)	Relative Area of total pics (%)
α-pinene	932	*Salvia rosmarinus*(1.6)	11.07
		Scoparia devera(1.3)	6.64
		Rhanterium adpressum(1.4)	23.50
		Cupressus(0.6)	32.65
		Juniperus(1.1)	39.37
		Artemisia campestris(1.0)	7.15

Camphene	949	*Salvia rosmarinus*	12.19
		Rhanterium adperssum	15.70
Camphor	1142	*Salvia rosmarinus*	48.77
		Artemisia alba(2.0)	7.63
Thujone	1105	*Artemisia alba*	69.99
3-thujanone	1114	*Artemisia alba*	15.13
Alpha phellandrene	1004	*Scoparia devera*	8.40
Limonene	1028	*Scoparia devera*	17.05
		Sinensis(2.1)	91.79
		Citrus(3.1)	61.14
		Artemisia campestris	18.2
Myristicin	1523	*Scoparia devera*	52.65
Ocimen Z- β	1027	*Scoparia devera*	7.79
Apiol	1625	*Scoparia devera*	21.02
Ligustillid	1731	*Scoparia devera*	7.31
Myrcene	990	*Rhanterium adperssum*	37.90
		Artemisia campestris	6.66
Sabinene	971	*Rhanterium adpressum*	6.71
		Artemisia campestris	6.66
β–pinene	975	*Citrus*	8.13
γ-terpinene	1058	*Citrus*	12.25
		Artemisia campestris	9.42
δ- 3 carene	1010	*Cupressus*	15.14
		Juniperus	10.79
Cedrol	1597	*Cupressus*	10.54
Terpinolene	1087	*Cupressus*	5.29
Terpinolene acetate	1349	*Cupressus*	8.88
Germacrene D	1478	*Juniperus*	7.48
Thymol	1297	*Saccocalyx*	37.58

		Origanum vulgare	13.52
α-terpineol	1192	*Saccocalyx(2.1)*	18.39
Borneol	1164	*Saccocalyx*	9.7
γ-terpene	1057	*Saccocalyx*	6.66
P-cymene	1023	*Saccocalyx*	5.34
Eugenol	1364	*Syzygium aromaticum(2.3)*	69.68
		Origanum vulgare(1.6)	5.91
β-caryophyllene	1420	*Syzygium aromaticum*	20.33
		Origanum vulgare	7.69
Eugenol acetate	1529	*Syzygium aromaticum*	6.32
Linalool	1107	*Origanum vulgare*	54.67
Linalyl acetate	1258	*Origanum vulgare*	5.02
Sabinyl acetate	1293	*Artemisia annua(1.2)*	78.43

3.3Docking:

The successful results that we have gotten from the docking process are summarized in **Table7.**

Table 7: Results of docking with tree Main proteases.

	Molecule	Energy(kcal/mol)	Estimated Ki	Ki units	Distance(Å)
5RUD	Origin Ligand	-5.48	96.19	uM	2.04 (H)
	Ligustilid	-6.39	20.71	uM	1.89 (H)
5RUN	Origin ligand	-5,45	0,1	mM	1,95545
	Cedrol	-6,69	12,48	uM	1,87454
	α-Terpineol	-5,73	63,08	uM	1,7955
5RUE	Origin ligand	-4,04	1,09	mM	1,88377

	α-Terpineol	-5,48	96,19	uM	1,70255
	Borneol	-4,49	0,51	mM	1,82651
	Eugenol	-4,9	0,26	mM	1,85963

- For 5RUD**:** we find that there are one compound which given the lowest binding energy and Hydrogen bond distance **Figure8**define the combination of Ligustillid in active site.

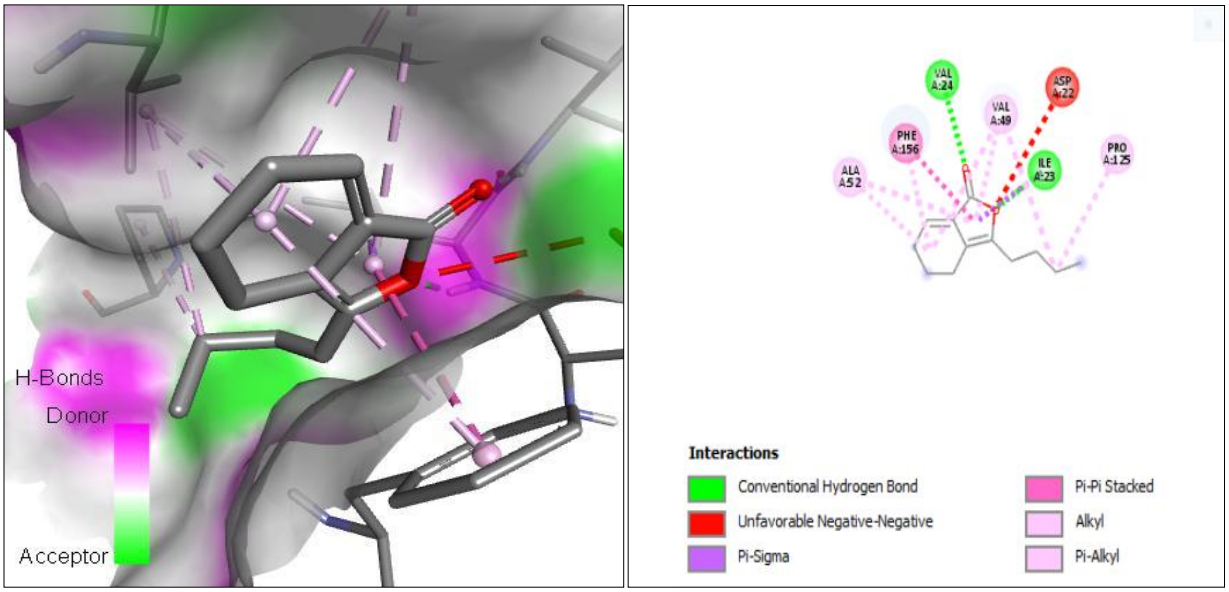


Figure 8: The docking interactions between Ligustillid molecule and active site of 5RUD.

Table 8: Result bond between Ligustillid and active site.

Numbers	Name	Color	Distance(Å)	Types
1	A:ILE23:HN		1,86	Conventional Hydrogen Bond
2	A:VAL24:HN		2,28	Conventional Hydrogen Bond

We note that there are several different bonds between the active site in the A chain of the protease and (4chloro-1.3-benzothiazol 2-amine) that was crystallized with it, this combination equal E= -5.48 Kcal/mol which including 2 hydrogen bonds between NH and the amino acid ALA 21 because it contains a carboxylic function of length of this bond is 2.04987 Å, we also have the second hydrogen bond between H in the amino acid ILE 23 and the electronegativity of N at

(4chloro-1.3-benzothiazol 2-amine) has a length of 2.09 Å, plus many weak interactions between the amino acids VAL 49, PHE156, ALA52, and VAL24 with this ligand.

Among all the majority compounds that we found in essential oils, only one has been successed in docking and gives us the lowest energy as well as hydrogen bonding:

The molecule of Ligustillid who shows a good result with E = -6.39 Kcal/mol , it formed two hydrogen bonds between ILE23 and VAL24 with two different oxygens, their length respectively are d = 1.86A°, d = 2,28A°.

- For 5RUN: we find that there are two compounds which given the lowest binding energy and Hydrogen bond distance. **Figure 9** define the combination of α-Terpineol with active site.

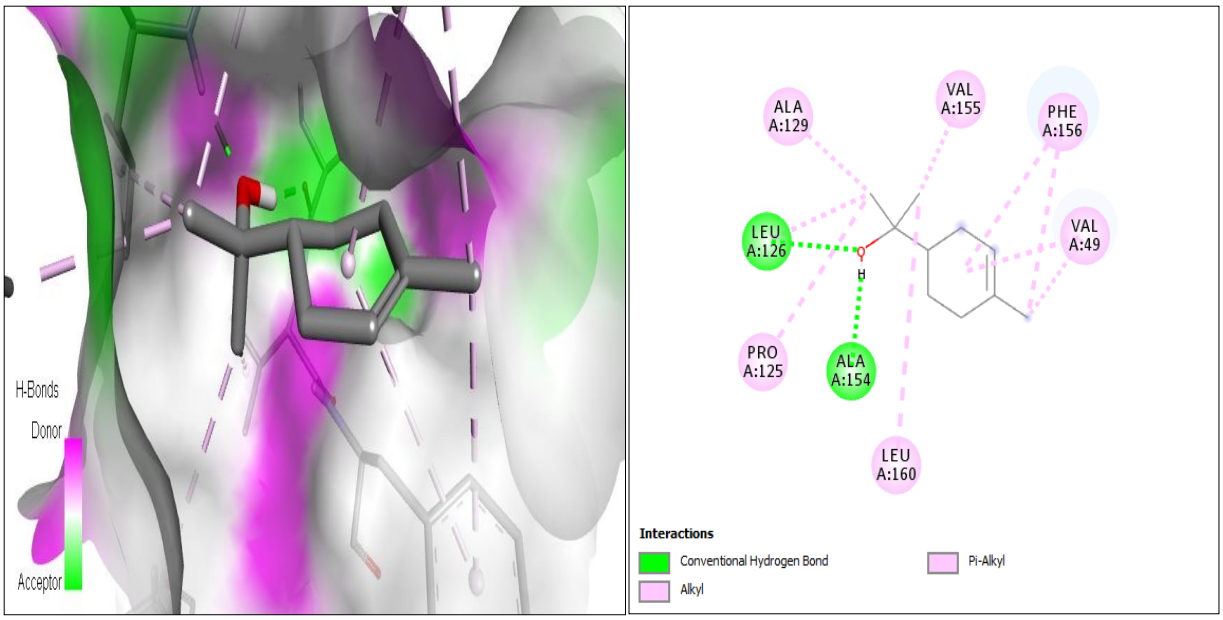


Figure 9: The docking interactions between α-Terpineol molecule and active site of 5RUN.

Table 9: Result bond between α-Terpineol molecule and active site.

Numbers	Name	Color	Distance(Å)	Types
1	A:ALA154:O		1,7955	Hydrogen Bond
2	A:LEU126:HN		1,9032	Hydrogen Bond

2- The second molecule is Cedrol combination **Figure 10** demonstrate that.

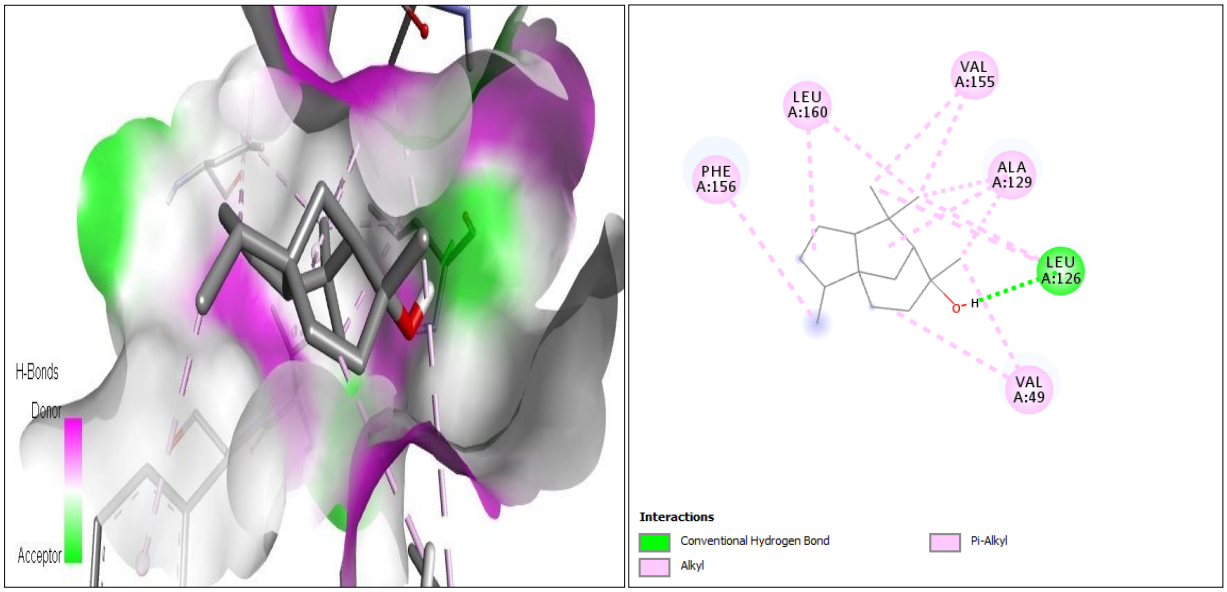


Figure 10: The docking interactions between Cedrol molecule and active site of 5RUN.

Table 10: Result bond between Cedrol molecule and active site.

Numbers	Name	Color	Distance(Å)	Types
1	A:LEU126:O		1,87454	Hydrogen Bond

In this second protease we have also several different bonds between the active site that exist in chain A of the protease and (3-(1Hbenzimidazol1yl) propanoic acid) that was crystallized with it, this combination equal E= -5.45 Kcal/mol which including 2 hydrogen bonds between the oxygen in carboxylic function and the amino acid ILE131 length of this bond is 1.95545Å,than we have the second hydrogen bond between the negative charge of oxygen and the amino acid VAL49 has a length of 2.06 Å, plus many weak interactions between the amino acids PHE156, LEU126, ALA129, and VAL155 with this ligand.

The successful compound on this protease are:

1-the first is Cedrol molecule which found in *Cupressus*, with an energy E = -6.69 kcal/mol, we see that there is one hydrogen bond between the hydrogen and LEU126 with a distance 1.87A°.

2-the second is α-Terpineol from the plant *Saccocalyx*, with an energy E = -5.73 kcal/mol, in atom oxygen of ligand which has a negative charge we see that it formed two hydrogen bonds with LEU126 and ALA154 , their distances are 1.90A° , 1.79A° respectively.

- For 5RUEwe find that there are tree compounds which given the lowest binding energy and Hydrogen bond distance **Figure 11** present the binding between α-Terpineol and active site.

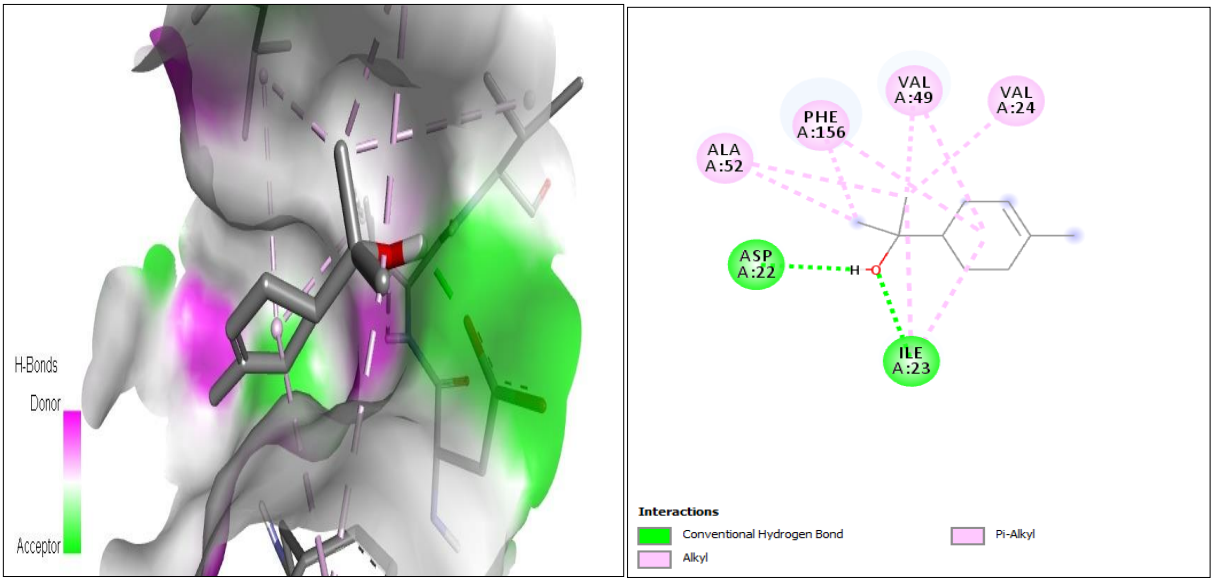


Figure11: The docking interactions between α-Terpineol molecule and active site of 5RUE.

Table 11: Result bond between α-Terpineol molecule and active site.

Numbers	Name	Color	Distance(Å)	Types
1	- A:ASP22:O		1,70255	Hydrogen Bond
2	A:ILE23:HN		2,29651	Hydrogen Bond

2-The second molecule is Borneol have interaction with active site as it look like in **Figure 12.**

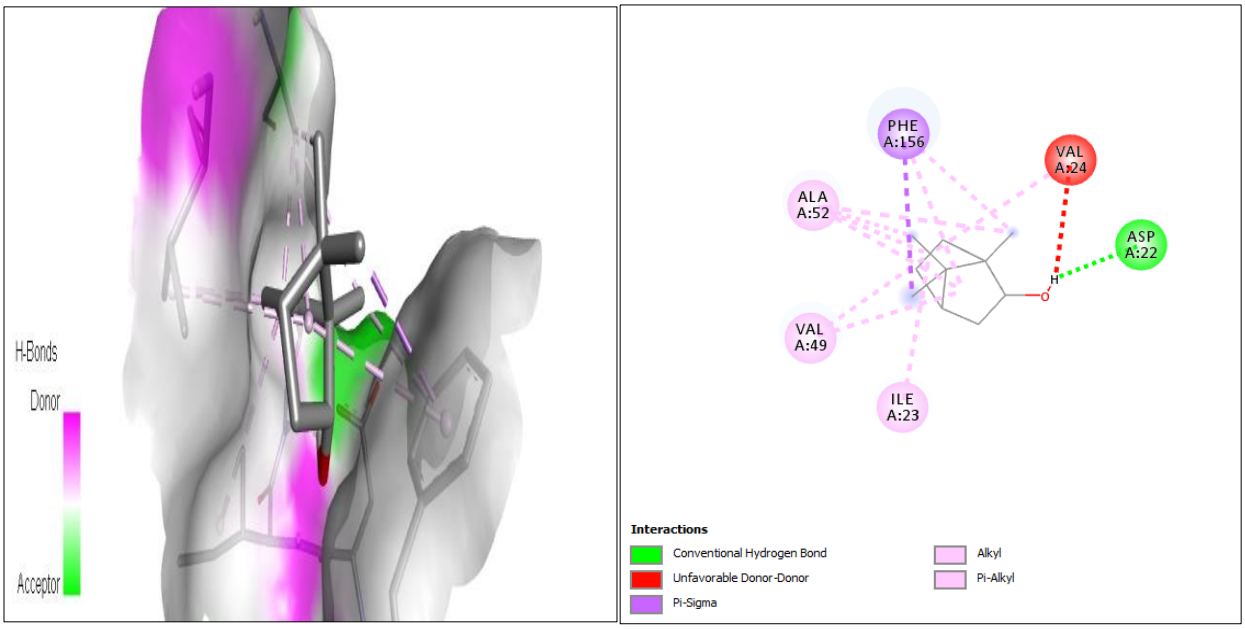


Figure12: The docking interactions between Borneol molecule and active site of 5RUE.

Table 12: Result bond between Borneol molecule and active site.

Numbers	Name	Color	Distance(Å)	Types
1	A:ASP22:OD1		1.82651	Hydrogen Bond

2-The third molecule is Eugenol also have a combination with active site **Figure 13** showing those combinations.

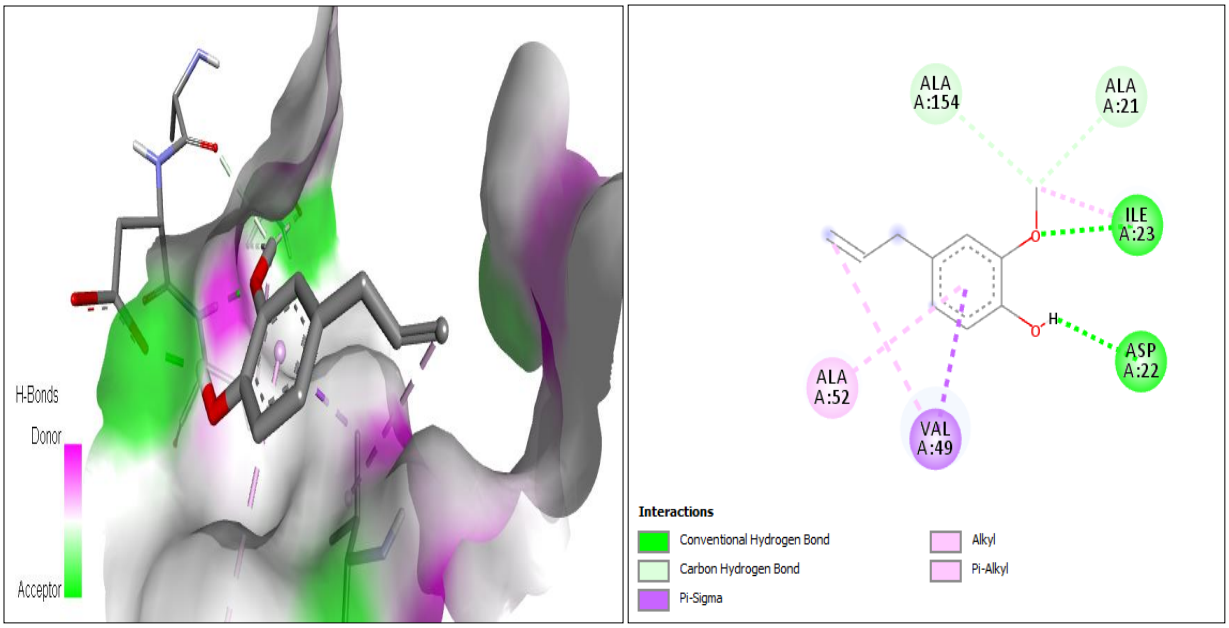


Figure13: The docking interactions between Eugenol molecule and active site of 5RUE.

Table 13: Result bond between Eugenol molecule and active site.

Numbers	Name	Color	Distance(Å)	Types
1	A:ASP22:O		1.85963	Hydrogen Bond
2	A:ILE23:HN		1.98982	Hydrogen Bond

The third protease like the others we can see that there is several different bonds between the active site that exist in chain A of the protease and (2-hydroxy-4-aminobenzoic acid) that was crystallized with it, this combination equal E= - 4.04 Kcal/mol which including three hydrogen bonds, first one between the negative charge of oxygen in carboxylate function and the amino acid PHE156 length bond is 1.88Å,than we have the second hydrogen bond between the oxygen carboxylic in ASP22 amino acid and HN in ligand has a length of 1.98Å,the third one is the amino acid ILE23 with the oxygen in ligand which equal 2.07Å. plus three weak interactions between the amino acids ILE23,VAL49,ALA52 and this ligand.

The successful compound on this protease are:

1-the first is α-Terpineol, in *Saccocalyx* plant with E = -5.48 kcal/mol we can see two hydrogen bonds, the first is madden between oxygen atom of ligand and ILE23 and the second between the hydrogen atom of ligand and ASP22 , their distances are 2.29A°, 1.70A°respectively.

2-the second is Eugenol exist in plant *Syzygium aromaticum* and *Origanum vulgare*, it's energy E = -4.90 kcal/mol we found two hydrogen bonds with different oxygen of ligand withILE23and ASP22, their distances are1.85A°, 1.98A°.

3-the third is Borneol molecule in *Saccocalyx* plant the energy equal E = -4.49 kcal/mol, there is one hydrogen bond between the hydrogen of alcohol function in ligand and ASP22 with distance1.82A°.

GENERAL

CONCLUSION

Among the thirteen plants studied, we selected five plants namely (*Saccocalyx, Scoparia deverra*, *Syzygium aromaticum*, *Origanum vulgare* and *Cupressus*) for their significant inhibitory ability. The simulation treatment showed that α-Terpineol (exist in *Saccocalyx*) inhibits two protease 5RUN and 5RUE with energy:-5.73and-5.48 kcal/mol the length hydrogen bonds: 1.79 and 1.90 A °, 1.70 and 2.29 A° respectively.

Ligustillid (from *Scoparia deverra*) inhibit the 5RUD protease, whose active site inhibitor stabilization energy is -6.39 kcal / mol and two hydrogen bonds of length from 1.89 to 2.28A °. Similarly to others, the Cedrol found in *Cupressus* has also the ability of inhibition for the 5RUN protease with energy equal:-6.69 kcal/mol and one hydrogen bond1.87 A°.

The 5RUE protease is inhibited by two other molecules eugneol (in *Syzygium aromaticum* and *Origanum vulgare*), borneol (in *Saccocalyx*) with the energies -4.90 kcal / mol and -4.49 kcal / mol respectively, the first form two hydrogen bonds while the second has formed a single hydrogen bond.

This all work shows that the essential oils of these five plants could inhibit the three proteases, which justifies the traditional use of the local population of herbal teas from these plants as a remedy against COVID-19.

In the end, there are prospects that we would hope to achieve in the future like:

- ✓ Tested the same experiment on tannins, phenols, alkaloids for those plants and compared the results.
- ✓ Confirmation of those results by testing in vivo.
- ✓ Studying the dynamic molecular of these compounds.

REFERENCES

1. Pedersen, S. F., & Ho, Y. C. (2020). SARS-CoV-2: a storm is raging. The Journal of clinical investigation, 130(5).

2. Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The architecture of SARS-CoV-2 transcriptome. Cell, 181(4), 914-921.

3. Yang, Y., Peng, F., Wang, R., Guan, K., Jiang, T., Xu, G., ... & Chang, C. (2020). The deadly coronaviruses: The 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China. Journal of autoimmunity, 109, 102434.

4. Luo, H., Tang, Q. L., Shang, Y. X., Liang, S. B., Yang, M., Robinson, N., & Liu, J. P. (2020). Can Chinese medicine be used for prevention of corona virus disease 2019 (COVID-19)? A review of historical classics, research evidence and current prevention p.

5. Rockx, B., Kuiken, T., Herfst, S., Bestebroer, T., Lamers, M. M., Munnink, B. B. O., ... & Haagmans, B. L. (2020). Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model. Science, 368(6494), 1012-1015.

6. *https://www.youtube.com/watch?v=Odltcv6q96I.* 27 November 2020.

7. *How does the coronavirus work? What it is, where it comes from, how it hurts us, and how we fight it.* Patel, Neel V. April 15, 2020.

8. Zhai, H., Liu, H., Wang, S., Wu, J., & Kluenter, A. M. (2018). Potential of essential oils for poultry and pigs. Animal Nutrition, 4(2), 179-186.

9. Takaya, D., Takeda-Shitaka, M., Terashi, G., Kanou, K., Iwadate, M., & Umeyama, H. (2008). Bioinformatics based Ligand-Docking and in-silico screening. Chemical and Pharmaceutical Bulletin, 56(5), 742-744.

10. Ballester, P. J., & Mitchell, J. B. (2010). A machine learning approach to predicting protein–ligand binding affinity with applications to molecular docking. Bioinformatics, 26(9), 1169-1175.

11. Benarfa, A., Gourine, N., Mahfoudi, R., Harrat, M., & Yousfi, M. (2019). Effect of seasonal and regional variations on phenolic compounds of Deverra scoparia (flowers/seeds) methanolic extract and the evaluation of its in vitro antioxidant activity. Chemi.

12. Mohamed, A. E. H. H., El-Sayed, M., Hegazy, M. E., Helaly, S. E., Esmail, A. M., & Mohamed, N. S. (2010). Chemical constituents and biological activities of Artemisia herba-alba. Records of Natural Products, 4(1)..

13. Dhanavade, M. J., Jalkute, C. B., Ghosh, J. S., & Sonawane, K. D. (2011). Study antimicrobial activity of lemon (Citrus lemon L.) peel extract. British Journal of pharmacology and Toxicology, 2(3), 119-122..

14. Naman, C. B., Gromovsky, A. D., Vela, C. M., Fletcher, J. N., Gupta, G., Varikuti, S., ... & Kinghorn, A. D. (2016). Antileishmanial and cytotoxic activity of some highly oxidized abietane diterpenoids from the bald cypress, Taxodium distichum. Journal of.

15. Ćavar, S., Maksimović, M., Vidic, D., & Parić, A. (2012). Chemical composition and antioxidant and antimicrobial activity of essential oil of Artemisia annua L. from Bosnia. Industrial Crops and Products, 37(1), 479-485.

16. Bozin, B., Mimica-Dukic, N., Samojlik, I., & Jovin, E. (2007). Antimicrobial and antioxidant properties of rosemary and sage (Rosmarinus officinalis L. and Salvia officinalis L., Lamiaceae) essential oils. Journal of agricultural and food chemistry, 55(19.

17. Radünz, M., da Trindade, M. L. M., Camargo, T. M., Radünz, A. L., Borges, C. D., Gandra, E. A., & Helbig, E. (2019). Antimicrobial and antioxidant activity of unencapsulated and encapsulated clove (Syzygium aromaticum, L.) essential oil. Food chemistry, 2.

18. Tantaoui-Elaraki, A., Lattaoui, N., Errifi, A., & Benjilali, B. (1993). Composition and antimicrobial activity of the essential oils of Thymus broussonettii, T. zygis and T. satureioides. Journal of Essential Oil Research, 5(1), 45-53.

19. Awad, M., & Abdelwahab, A. (2016). Chemical diversity of essential oils from flowers, leaves, and stems of Rhanterium epapposum Oliv. growing in northern border region of Saudi Arabia. Asian Pacific Journal of Tropical Biomedicine, 6(9), 767-770.

20. Şahin, F., Güllüce, M., Daferera, D., Sökmen, A., Sökmen, M., Polissiou, M., ... & Özer, H. (2004). Biological activities of the essential oils and methanol extract of Origanum vulgare ssp. vulgare in the Eastern Anatolia region of Turkey. Food control, 1.

21. zeraib, azzedine. *Les huiles essentiells* . 2010.

22. Perricone, M., Arace, E., Corbo, M. R., Sinigaglia, M., & Bevilacqua, A. (2015). Bioactivity of essential oils: a review on their interaction with food components. Frontiers in microbiology, 6, 76.

23. Wolford, R. W., Kesterson, J. W., & Attaway, J. A. (1971). Physicochemical properties of citrus essential oils from Florida. Journal of Agricultural and Food Chemistry, 19(6), 1097-1105.

24. *Handayani, D., Hadi, D. R., Isbaniah, F., Burhan, E., & Agustin, H. (2020). Corona virus disease 2019. Jurnal Respirologi Indonesia, 40(2), 119-129.*

الملخص

يعتبر فيروس كورونا المستجد هو جائحة القرن الحادي والعشرين، حيث يصيب الرئتين و يسبب متلازمة الالتهاب الرئوي الحاد. كما انه سريع الانتشار من خلال أي اتصال قريب مع شخص مصاب. يهدف هذا العمل الى دراسة امكانية تثبيط ثلاثة بروتياز 5رید، 5ران و 5ري بواسطة الزيوت العطرية المستخلصة لثلاثة عشر نبتة محلية . باستخدام الالتحام الجزيئي من خلال برامج اوتودوك4 و ام دوك. و لفهم آلية التثبيط والتفاعل المتضمن داخل الموقع الفعال استخدمنا ديسكوفري ستوديو و لمعالجة النتائج تم استخدام الحد الأدنى من قيمة الطاقة للتحقق من نتيجة الإرساء. تظهرالنتائج التي تم الحصول عليها أن ليجيستيليد، الفا تربينيول،سيدرول،ايجينول، الفا كادينول، بيتا ايدسمول من بين المثبطات الجيدة.

الكلمات المفتاحية

سارس-كوف، مارس-كوف، ان-كوف، بروتين داتا بنك. دوكينغ، بروتين،ليجاند، احماض نووي

Résumé

Le nouveau corona virus est la pandémie du 21émé siècle qui infect les poumons et provoque le syndrome respiratoire aigu sévère. Il se propage rapidement par tout contact proche avec une personne ou une surface infectée. L'intérêt de ce travail, est étudier l'inhibiter trois protéases : 5RUD, 5RUN et 5RUE par des huiles essentielles in silico. En utilisant le docking moléculaire, par les programmes Autodock4 , pour comprendre le mécanisme d'inhibition et l'interaction impliquée à l'intérieur du site actif, nous avons utilisé le discovery studio pour le traitement des résultats, la valeur d'énergie minimale est utilisée pour valider les résultats d'amarrage. Les résultats obtenus montrent que les composés: le Ligustillid, le α-Terpineol, le Cedrol, l'Eugenol, le α-Cadinol, le β-Eudesmol sont des meilleurs inhibiteurs.

Les mots clés : in silico, Autodock4, amarrage, protéine, ligand

Abstract

The novel coronavirus is the pandemic of the 21st century, infects the lungs and causes the severe acute respiratory syndrome. It spreads quickly through any near contact with an infected person or surface. The aim of this work is to inhibit three protease: 5RUD, 5RUN and 5RUE by essential oils in silico. Using molecular docking, byAutodock4 programs, to understand the inhibition mechanism and the involved interactions inside the active site we used discovery studio for the results treatment, the minimum energy value was used to validate the docking results. The obtained results show that the Ligustillid, the α-Terpineol, the Cedrol, the Eugenol, the α-Cadinol ,the β-Eudesmol best inhibitors to proteases covid-19

The Key words: in silico, Autodock4,docking, protein, ligand