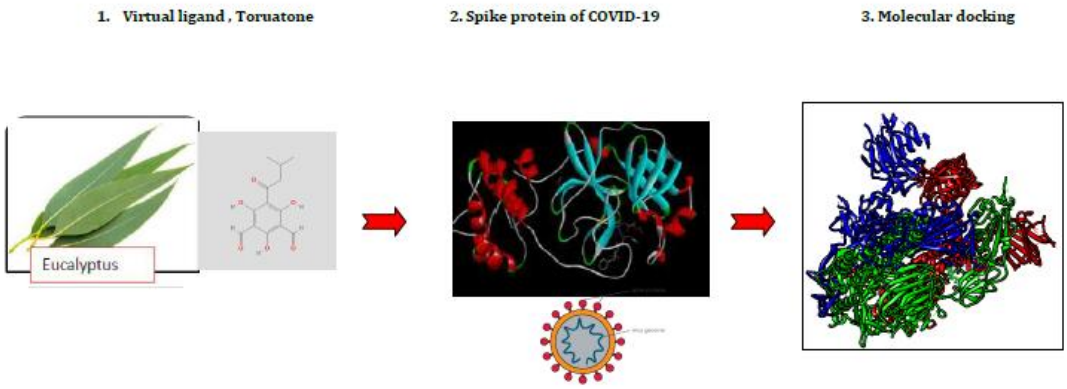
Homology modeling and molecular docking of Natural metabolites from eucalyptus essential oil against SARS-CoV-2 spike protein

Arun Dev Sharma* and Inderjeet Kaur

PG dept of Biotechnology, Lyallpur Khalsa College Jalandhar



Abstract

SARS-CoV-2 spreading its tentacles across the world as a major world pandemic. There is a race to develop suitable drug drugs and vaccines for this disease. Due to lack ofdrug drugs s at present, various anti-HIV drug drugs s have been repurposed. Due to its vital role in virus replication, Spike (S) protein has recently been regarded as a suitable target for drug drugs design. With an urgent and key need for safe and effective drugs to combat this disease, we have explored 12 bioactive compounds from eucalyptus oil against Spike (S) protein from SARS-CoV-2. SWISS-MODEL was used for homology modeling. Molecular docking studies were conducted by using Patch dock analysis. Protein Interactions Calculator was used for protein interactions. Pharmacokinetics and In-silico ADMET profile (absorption, distribution, metabolism, excretion and toxicity) was also studied. Docking score showed effective binding of all bioactive molecules especially Toruatone to COVID-19 S-protein. Interactions results indicated that Spike (S) protein / Toruatone complexes forms hydrogen and hydrophobic interactions. In-silico absorption, distribution, metabolism, excretion and toxicity (ADMET) studies provided guidelines and mechanistic scope for identification of potent anti-COVID 19 drugs. This study highlights the potential of existing bioactive components from essential oil from eucalyptus to act as anti-COVID-19drug drugs.

Keywords: COVID-19, Docking, Eucalyptus oil, *Toruatone*, Herbaldrug drugs

Corresponding author: Arun Dev Sharma, E mail: arundevsharma47@gmail.com

Introduction

A new corona virus, 2019-n-CoV, caused a pandemic of pulmonary disease (COVID-19) in the city of Wuhan, has since spread globally (Huang et al., 2020) . The virus has been named as SARS-CoV-2, because the genome (RNA) of the virus is 82% identical to the SARS corona virus (SARS-CoV). As of April, there are about 2500,000 cumulative cases globally, with a 3.7% case-fatality rate due non-availability of specific therapies for COVID-19. As a first line of treatment, along with the antiviral drug drugs s, clinicians are using hydroxychloroquine, and anti-HIV vaccines, SARS-CoV-2 and MERS-CoV neutralizing antibodies targeting the S1 domain of the SARS-CoV-2 spike glycoprotein against COVID-19 treatment (Liu and Wang, 2020). Unfortunately no specific vaccine vaccines is available to date so dire urgent need to treat COVID-19 has led the researchers on Spike proteins as potential drug drugs targets to combat this disease. CoVs belong to Corona viridae families are positive RNA genome enveloped viruses which are alienated into four genera (α, β, γ, and δ). The SARS-CoV-2 belongs to the β genus. SARS-CoV-2 contains four structural proteins: Spike (S) protein, membrane (M) protein, envelope (E) protein, and nucleocapsid (N) protein (S. Fig 1). Among them, the outer membrane spike glycoprotein, is the major host interacting protein with host cell targets (such as ACE2, CD26, Ezrin, cyclophilins and other cell adhesion factors) important for cell adhesion and virulence. Like other corona viruses,- Corona virus entry into the host cells is mediated by the transmembrane spike (S) glycoprotein that forms homo-trimer protruding from viral surface (Lu 2020). The SARS-CoV-2 S-protein contains a conserved Receptor Binding domain (RBD) which recognizes host cell receptors like ACE2, CD26, Ezrin, cyclophilins and other cell adhesion factors. It is 1200 aa long protein, belongs to class-1-viral fusion proteins and contributes to the cell receptor binding, tissue tropism and pathogenesis (Millet et al., 2012). Therefore, by virtue of its key role, SARS-CoV-2 S-protein is considered to be a suitable target for viral inhibitor development. Inhibiting the activity of SARS-CoV-2 S-protein activity would block viral replication. Since in humans no proteases with similar cleavage specifics are known, so making it an ideal antiviral target. Natural sources derived bioactive molecules that can inhibit SARS-CoV-2 spike protein are measured an additional approach to combat COVID-19 (Bhardwaj et al., 2020). A computational approach based on in silico techniques, is the capable preliminary substantiation fordrug drugs discovery. Various molecular docking studies have

recognized bioactive molecules from natural products as the potential SARS-CoV-2 spike protein inhibitors derived from natural sources (Bhardwaj et al., 2020). The discovery of newdrug drugs s as therapeutic agents is a dreary and luxurious process. Traditional methods can take about 12-14 years with lots expenditure to bring adrug drugs from discovery to market. In order to overcome these challenges, multidisciplinary approaches are required in the process ofdrug drugs discovery.

From ancient time, medicinal plants are beneficial in the field of drug drugs therapeutics as they are safer alternatives being utilized by humans for centuries. Previously, many of the new drug drugs formulations are derived from natural products. In Traditional Ayurvedic and Chinese medicines, natural plant products have been used immensely for generations as antiviral treatments. In addition, natural compounds are also a primary source for a huge number of modern drug drugs s. Chloroquine and hydroxychloroquine, important natural compounds are derived from secondary metabolites of Cinchona tree that is under clinical trial and has shown potential anti SARS-CoV-2 properties (Wu et al., 2020). Our present study focuses on the in silico analysis of essential oil present in Eucalyptus plants. It has been used in various systems of medicine for the past 2000 years (Yang et al., 2018). The leaves, flowers, fruits and roots of Eucalyptus are used to treat various diseases such as coughs, colds, asthma, skin infections, fever, bleeding piles and inflammation. Hence, immense awareness has been paid to the secondary metabolites secreted by plants that may be urbanized as medicines. Among several compounds, essential oils, from medicinal plants, have been reported to have antioxidant, antimicrobial and antiviral bioactivities, hence immensely used in cosmetic, pharmaceutical industries (Liu and Wang 2020). Essential oils are complex, volatile mixture of terpenes, oxygenated derivatives and other aromatic compounds. Empirically, eucalyptus essential oil has been used as antimicrobial agents, however, its antiviral potential is still a matter of conjuncture (Shukla et al., 2014). Over the years many approaches have used to designdrug drugs and released into the commercial market for use. Among all, structure-based drugs design (SBDD) is most commonly used , which based on 3-D structure protein target to propose a suitable ligand that can pose as its potential inhibitor (Singh et al., 2016). In SBDD,–molecular docking is a key technique that can be applied in designing drugs making process. Molecular docking has facilitated researchers to virtually monitor a collection of bio-actives against the receptor target

protein and analyze binding conformations and affinities of the compounds to the receptor (Barcellos et al., 2019). This study investigated docking of bioactive molecules from eucalyptus oil as key inhibitor candidates for SARS-CoV-2 spike (S) protein. This investigation will be helpful to other scholars with more options to find out apt drugs to combat SARS-CoV-2. We hypothesize that these bioactive components from eucalyptus essential oil has the capability to prevent infection of SARS-CoV-2. The results of the study will provide other scholars with opportunities to identify the right drug drugs to combat SARS-CoV-2. To further estimate the drug drugs gability of these bioactive molecules, the present study reports their ADMET profiles and potential biological activities using various in silico bioinformatics tools.

Materials and Methods

Homology modeling of SARS-CoV-2 Spike protein

Spike protein amino acid sequence in FASTA format was retrieved from EMBL-EBI database of protein banks with the accession number 6VXX. Afterwards, raw amino acid sequence was submitted to SWISS-MODEL web sever for modeling with sufficient sequence identity and similarity. SWISS-MODEL builds 3-D model based on Qualitative Model Energy Analysis (QMEAN) and Global Model Quality Estimation (GMQE) values. The GMQE values in the range of 0 and 1, and QMEAN value under 4.0 shows reliability of the predicted structure.

Structure Validation

SWISS-model web server itself calculates QMEAN scores to estimate global and local quality of the protein structure. This server also provides Z score between 0 and 1 based on expected values with known structure. PROCHECK tool was also used to analyze the 3D structure of Spike protein generated by SWISS-MODEL. For this 3D structure in .pdb file format was uploaded to PDBsum web servers of EMBL-EBI to obtain Ramachandran plot and the Ramachandran plot statistics.

Ligand modeling

A total of about 12 natural molecules from selected from literature as shown in Fig 1. All bioactive structures were obtained from the PubChem database

(https://pubchem.ncbi.nlm.nih.gov/). SMILES (simplified molecular-input line-entry system) were retrieved for each bioactive molecules and converted to their corresponding 3D structures by using UCSF-chimera and saved in .pdb format.

Protein receptor preparation

SARS-CoV-2 spike homology modeled 3D structure in its .pdb format was retrieved from SWISS-MODEL. 3D modeled protein structure was designated as receptors. Before the docking study, the protein structure was first prepared using the dock prep set up in chimera software. The dock preparation is an optimization part that corrects atomic and bond length, structure, charges anomalies. Original inhibitors (NAG) and water molecules were detached from the spike protein structure and any missing hydrogen atoms were added

Active sites prediction in 3D modeled receptors

CASTp (The Computed Atlas of Surface Topography of proteins) web tool was used to predict active sites residues in the spike 3D modelled protein. CASTp is an online tool used in identification and dimension of cavities on 3D protein structures.

Molecular Docking

The docking study of the compound over SARS-CoV-2- S-protein was studied using PatchDock (https://bioinfo3d.cs.tau.ac.il/PatchDock/) softwares. For this both ligands (bio-molecules) and receptor (modeled spike protein 3D structure) molecules in .pdb file formats were uploaded to PatchDock servers and job was executed. The best generated docked structure was downloaded and saved as .pdb file. After docking complex structures were studied by Biovia Discovery Studio Visualizer 2020 and Chimera tools. 3D interactions were studied using PLip web tool (https://projects.biotec.tu-dresden.de/plip-web/plip/index).

*(*In silicodrug drugs -likeness and toxicity predictions)

In silicodrug drugs -likeness and toxicity predictions were studied using admetsar2 and SWISS ADMAE tool to predicting toxicity of candidate molecules (http://lmmd.ecust.edu.cn/admetsar2/). This tool estimates the carcinogenic nature, tumorigenic mutagenic, irritant and biodegradable ability. Drugs likeliness analyzes a pharmalogicaly active

agent and they predict its oral activeness. It is usually done by Lipinski rule of five which cite that compound likely shows poor absorption when: molecule has more than 5 H-bond donors and 10 H-bond acceptors, MW more than 500 and LogP (CLogP) greater than 5.37. Swiss ADME online tool is used to find outdrug drugs -likeness and other properties like hydrophilicity (Log kp), solubility (Log S), surface area (TPSA) and blood brain barrier penetration (BBB). It also predicts the pharmacokinetic properties such as Absorption, Distribution, Metabolism, Excretion and Toxicity. Bioactivity analysis was carried out using molinspiration tools (https://www.molinspiration.com/cgi-bin/properties)

Result and discussion

Traditionally, the invention of recent therapeutic medicine could be a tedious and expensive method that typically takes 12-14 years with a great deal of cash to bring drugs into the market. With the aim of overcoming these issues a great deal several multidisciplinary approaches are accustomed discover newdrug drugs . In drugs discovery processes, in silico coming up with of drugs could be a apply of computer based mostly modeling that is extremely helpful. Within the field ofdrug drugs discovery, medicative plants are advantageous as they're used as a secure seasoning different by humans for hundreds of years. The sources of many of the active constituents of medicines and novel medicine are obtained from natural products.

Corona viruses (CoVs) are a bunch of viruses that infect animals and humans. CoV infections have an effect on animals in numerous ways that like: metabolism, fever, cold, digestive, and liver systems of humans and animals. Engulfed viruses access their host cells by a method of membrane fusion that's mediate by a selected fusion, or ‘‘spike’’ super molecule, encoded by the virus and embedded within the infective agent envelope (Wu et al., 2020) (S. Fig 1). It's the S-protein found within the CoV related to the severe acute metabolism syndrome (SARS), which might be accessed in PDB and was instructed to be a possibledrug drugs target for 2019-nCov (Millet et al., 2012). In several viruses, S-protein play essential roles in infective agent entry into host cells; so, S-proteins are usually used as supermolecule targets throughout the event of antiviral medicine. The coronavirus spike supermolecule (S) plays a key role within the early steps of virus infection, with the S1 domain liable for receptor binding and therefore the S2 domain region (aa570–aa1278), mediating membrane fusion. As cited by earlier, S-proteins

represent key targets for the inhibition virus replication, and therefore the super molecule sequences of the SARS-CoV S-protein and the 2019-nCoV S-protein are 91% identical, thence host proteases may be used as potential therapeutic targets. In several viruses, proteases and Spike proteins play essential roles in infective agent replication; so, Spike is usually used as supermolecule targets throughout the event of antiviral therapeutics (Song et al., 2018). In CoV, the Spike supermolecule is concerned in virus chemical process maturation and has been examined as a possible target supermolecule by inhibiting the cleavage of the infective agent polyprotein to forestall the unfold of infection (Li et al., 2020). Engulfed viruses get enter into their various host cells through membrane fusion method that's expedited via an explicit fusion, or ‘‘S’’ supermolecule (virus encoded), and embedded within the envelope of the virus (Millet et al., 2012). The invention of the spike proteolytic enzyme structure in COVID-19 provides a pleasant path to spot potentialdrug drugs candidates to forestall infection. We tend to follow the structural biology aspects that specialize in the supply and retrieval of an S-protein receptor structure from PDB info. In this study, we tend to perform in silico analysis of eucalyptus oil bioactive components against S-protein super molecules of COVID-19.

Homology modeling and structure validation

3D structure of SARS-CoV2 S-protein was modeled using SWISS-MODEL with GMQE and QMEAN scores 0.5 and -1.72 respectively. Also SARS-CoV spike glycoprotein bound with NAG (PDB ID: 6crx 1) was identified as **a** closet template to SARS-CoV2 S-protein with 78% sequence identity. Multiple comparison between model structure and template is shown in Fig. 2 GMQE value of 0.65 and QMEAN score -1.72 indicated that structure is reliable and good quality. Further, graphical representation of the local quality estimate is shown in Figure 3, which indicated that residues was close to 1, depicting that the local quality estimate of the residues of the predicted model is good. The modeled protein structure also lies within the range of other protein structures in PDB, which confirms its reliability (Fig. 3).

Molecular Docking

In fashionabledrug drugs discovery method, molecular arrival could be a wide used machine technique to predict the binding mode and binding affinity for ligands with the target receptor supermolecule. The effectiveness of the docked advanced was evaluated on the idea of 2

essential criteria's: The minimum energy and therefore the interaction of the matter with the site residues. A ligand undergoes either H bonding or hydrophobic interactions or each whereas arrival into the site. The results of arrival may be accustomed realize the most effective inhibitors for specific target proteins and so to style new medicine.

Molecular arrival exploitation patchDock dock tool that was accustomed verify interaction of substance bioactive compounds of oil with modeled 3D structure of SARS-CoV-2 S- protein disclosed twenty totally different poses supported the dock score and pose with highest score values indicated most binding affinity (data not shown). Among all molecules, Toruatone exhibited highest affinity with S–protein with maximum docking score followed by Phenethyl phenylacetate (Table 1). The high affinity ofdrug drugs compounds depends on the amount and the type of bonding that occur with the active site of the protein. As indicated in Figure 4, Toruatone forms many chemical bonds with S2 domain of S protein, including hydrogen bonds and hydrophobic bonds. S-protein comprises two functional subunits responsible for binding to the host cell receptor (S1 unit) and fusion of the viral and cellular membrane (S2) subunit (Song et al., 2018). According to the Plip server and Biovia studio results, the interaction of Toruatone in the binding pocket of S2-protein domain was mediated by hydrophobic interactions via PRO589 and LEU977 with S2-protein domain at atomic distances of 3.68 and 3.11 (Figure 5). Hydrogen bond interactions of Toruatone were also observed via ASP 745, ARG1000. CAST-P server also revealed the presence of PRO589, LEU977 ASP 745, and ARG1000 as active site residues in the major cavity of SARS-CoV-2 S protein. With CASTp active site prediction, a major pocket was identified with Area (SA) of 138354 and Volume (SA) of 84670195 (Figure 6). The Toruatone showed full fitness within active site amino acids of S2 domain of S-protein of SARS-CoV-2. S-protein formed complexes with other eucalyptus oil bioactive depend molecules from eucalyptus essential oil by forming various interactions as hydrophobic residues interactions, and hydrogen bond. For example, Eucalyptol formed hydrophobic interactions with ALA1016. Result of the present study shown that LEU761/864 and GLY613/667 were critical residues for binding of Jensenone to S-protein via hydrophobic and hydrogen bond interactions, respectively. GLN1002 and THR1009 participated in hydrophobic interaction with Citronellal and Geranyl acetate. It was postulated that these ligand-protein interactions might inhibit the formation of the S2 motif of S- protein. Molecular interactions of eucalyptus bio-molecules are

illustrated in Fig 7. Coronavirus trafficking into and hijacking the host cell is primarily driven by the C-terminal S2 domain of spike conjugated protein that interacts with many host cell proteins (Liu and Wang 2019). It absolutely was postulated that COVID-19 S-protein becomes closed upon binding with eucalyptol –that in turn brings conformational changes of the S-protein and stop initiation of the fusion reaction responsible for its insertion into the host cell membrane. Antiviral bioactivities of many compounds like flavonoids, terpenoids and phenolics from essential oils, have been reported (Zakaryan et al., 2017). Almost all the essential oil bioactive molecules formed hydrophobic interactions and hydrogen bonds with amino acids of S-protein of COVID-19. Hydroxy groups (-OH), ketone groups (=O) and ether groups (-O-) in eucalyptol compounds are predicted to play roles amino acid residue interactions at the active site of COVID-19 S-protein. Further studies may help to understand the role of these residues in drugs binding mechanism. Structural flexibility is one in every of the vital physical properties that have an effect on macromolecule conformation and performance. In this context, substance by binding to a macromolecule will alter its flexibility and reduce its protein activity Further studies may help to understand the role of these residues in drugs binding mechanism.

Drug-likeness and toxicity analysis

ADMET properties (absorption, distribution, metabolism, excretion and toxicity) are very important for human therapeutic use of the compounds (Wu et al., 2020). Thedrug drugs scanning results are illustrated in Table 2. All metabolites were accepted by Lipinski's rule of five. Topological polar surface area (TPSA) was good, indicating permeating nature of metabolites across cell membranes. Log BB which indicates blood–brain barrier was also good. Almost all natural metabolites were non-sensitive to efflux transporter P-glycoprotein (p-gp). Consensus Log $P_{o/w}$ (indicative of lipophilicity) was 2.67, indicating that metabolites have optimal blood-brain-barrier penetration. Molecules were non-substrate to liver and in intestine based cytochrome P series (CYP1-3) involved in detoxification. Molecular lipophilicity potential (MLP) and polar surface area (PSA) are also shown in Figure 6. MLP is key property to diminish various molecular ADME characteristics (like membrane penetration or plasma-protein binding). Analysis of 3-D distribution of hydrophobicity on molecular surface is particularly useful when explaining differences are observed in ADME properties of molecules with the same logP. Bioactivity and toxicity of best docked bioactive molecule toruatone as thedrug drugs was

calculated online by using Mol inspiration drugs likeness score. Table 3 depicts Bioactivity score of Toruatone. Ion channel property and kinase inhibitor of Calotropin were high (<0) while enzyme inhibitor and protein inhibitor activities were moderate (>0). Further organ, genome and eco-toxicity was evaluated and toxicity profile revealed that toruatone bioactive molecule was mostly non toxic as negative values were observed for organs like eyes corrosion,drug drugs -induced liver injury and human either-a-go-go-ralted gene (hERG) inhibition (Table 4). The hERG is important for cardiac repolarization that encodes rapidly activating delayed rectifier potassium channel (IKr). Dysfunction of hERG causes long QT syndrome and sudden death, which occur in patients with cardiac ischemia. Toruatone was noncancerous and eco-friendly.

Conclusion

Due to non approved drugs s at present currently, SARS-CoV-2 has emerged in the human population, in China, and is a potential threat to global health, worldwide. Currently, the main target for COVID-19 treatment primarily act on the S-protein. The aim of this study was to examine bioactive molecules from eucalyptus essential oil that may be used to inhibit the SARS-CoV-2 infection pathway. Therefore, we suggested that Toruatone may represent potential treatment options, and found in medicinal plants that may act as potential inhibitors of COVID-19 S-protein. However, further studies should be conducted for the validation of these compounds using in vitro and in vivo models to pave a way for these compounds indrug drugs discovery.

Conflict of interest
Authors declare no conflict of interest
Funding: nil
Acknowledgements: Authors want to acknowledge managing committee LKC

References

Bhardwaj VK, SinghR, Sharma J,. Rajendran V, Purohit R, Kumar S (2020) Identification of bioactive molecules from tea plant as SARS-CoV-2 main protease inhibitors Journal of Biomolecular Structure and Dynamics. 5: 1-10

Huang C, Wang Y, Li X. et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 395: 497–506

Li X, Wang Y, Agostinis P et al. (2020) Is hydroxychloroquine beneficial for COVID-19 patients? Cell Death Dis 11:512.

Lipinski CA (2004) Lead- anddrug drugs -like compounds: the rule-of-five revolution.drug drugs Discovery Today. Technologies 1 : 337-341.

Liu X, Wang XJ (2020) Potential inhibitors against 2019-nCoV coronavirus M protease from clinically 13 of 14 approved medicine. *J. Genet. Genomics*, doi: 10.1016/j.jgg.2020.02.001.

Lu H (2020)drug drugs treatment options for the 2019-new coronavirus (2019-nCoV), Biosci. Trends, doi:10.5582/bst.01020.

Millet JK, Kien F, Cheung CY et al (2012) Ezrin interacts with the SARS coronavirus spike protein and restrains infection at the entry stage. Plos One.;7, e49566.

Sharma AD, Kaur I (2020) Molecular docking studies on Jensenone from eucalyptus essential oil as a potential inhibitor of COVID 19 corona virus infection. Research and Reviews in Biotechnology and Biosciences, 1: 59-66.

Shukla A, Sharma P, Prakash O, Singh M, Kalani, K. et al (2014) QSAR and Docking Studies on Capsazepine Derivatives for Immunomodulatory and Anti- Inflammatory Activity. PLoS ONE 9: e100797.

Song WF, Gui M, Wang X. et al (2018) Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. PLoS Pathog.14: e1007236.

Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, Wang Q, et al. (2020) Analysis of therapeutic targets for SARS-CoV-2 and discovery of potentialdrug drugs s by computational methods. Acta Pharmaceutica Sinica B, 10: 766-788.

Yang L, Wen KS, Ruan X, Zhao YX, Wei F, Wang Q 2018) Response of plant secondary metabolites to environmental factors, Molecules, 23, 1–26, doi: 10.3390/molecules23040762.

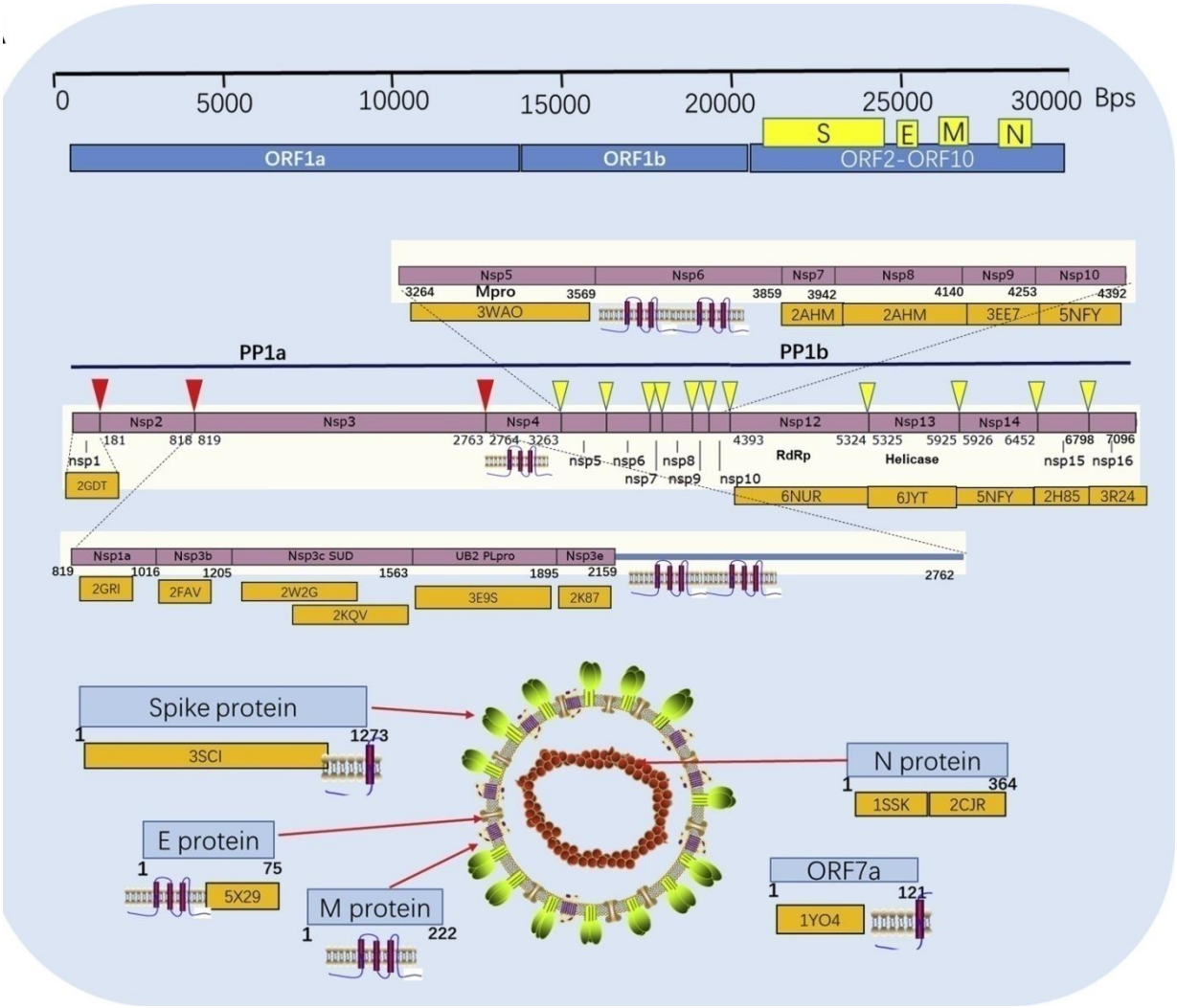
Zakaryan H, Arabyan E, Oo A, Zandi K (2017) Flavonoids: promising natural compounds against viral infections. Arch. Virol., 162: 2539–2551, , doi: 10.1007/s00705-017-3417-y.

Singh P, Kumari K, Awasthi SK, Chandra R. Virtual screening and docking studies of synthesised chalcones: potent anti-malarialdrug drugs . Int Jdrug drugs Dev Res. 2016;8:49–56.

Barcellos MP, Santos CB, Federico LB, Almeida PF, da Silva CHTP, Taft CA. Pharmacophore and structure-baseddrug drugs design, molecular dynamics and admet/tox studies to design novel potential pad4 inhibitors. J Biomol Struct Dyn. 2019;37:966–981.

Table 1: Docking analysis of essential oils molecules with spike protein

S NO	Model name				Interacting residues (4 A°)	
		Score	Area	ACE	H-bond interactions	Hydrophobic interactions
1	Eucalyptol	3122	376.90	-60.19		ALA11016,1020
2	Citronellol	3792	423.20	-85.35	GLN314, 613	PRO655, VAL772, LEU861
3	Citronellal	3664	401.10	-61.45	-	GLY1002, THR1009
4	Phenethyl phenylacetate	4846	536.40	-162.45	SER 513 GLN613	ILE312, THR768, VAL772
5	Geranyl acetate	4396	485.00	-86.73	GLN1002	GLN1002, THR1009
6	Limonene	3340	385.40	-104.30	-	ALA1016,1020
7	4-terpineol	3476	386.30	-73.80	THR1009	GLN1002, 1005
8	P-cymene	3440	376.80	-187.94	-	LEU390, 517,546, ALA522, PHE565
9	Alpha-pinene	3476	381.10	-94.02	-	ALA1016, 1020
10	Toruatone	5324	570.20	-134.81	ASP978, ARG1000	PRO589, LEU977
11	Jensenone	4162	473.70	-114.97	SER596, GLY613, 667, LYS733	PRO655, VAL772, ASP775, LEU 761,864
12	Cis-p-2-Menthen-1-ol	3576	386.50	-191.67	THR393	LEU390,517, ASN544



S. Fig 1: Overall genome and protein analysis of SARS-CoV-2

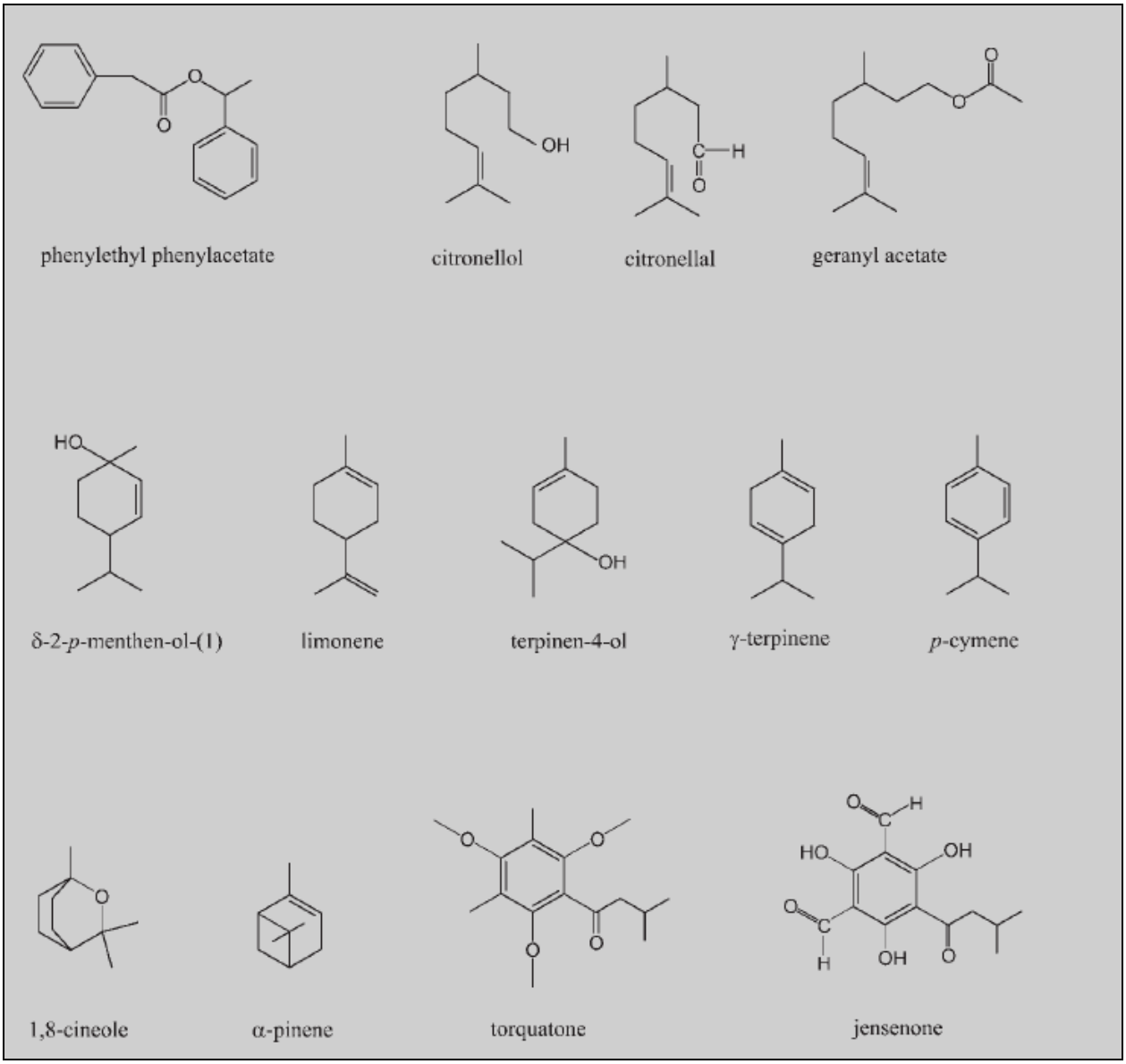


Fig. 1: Molecular structures of main substances in essential oil

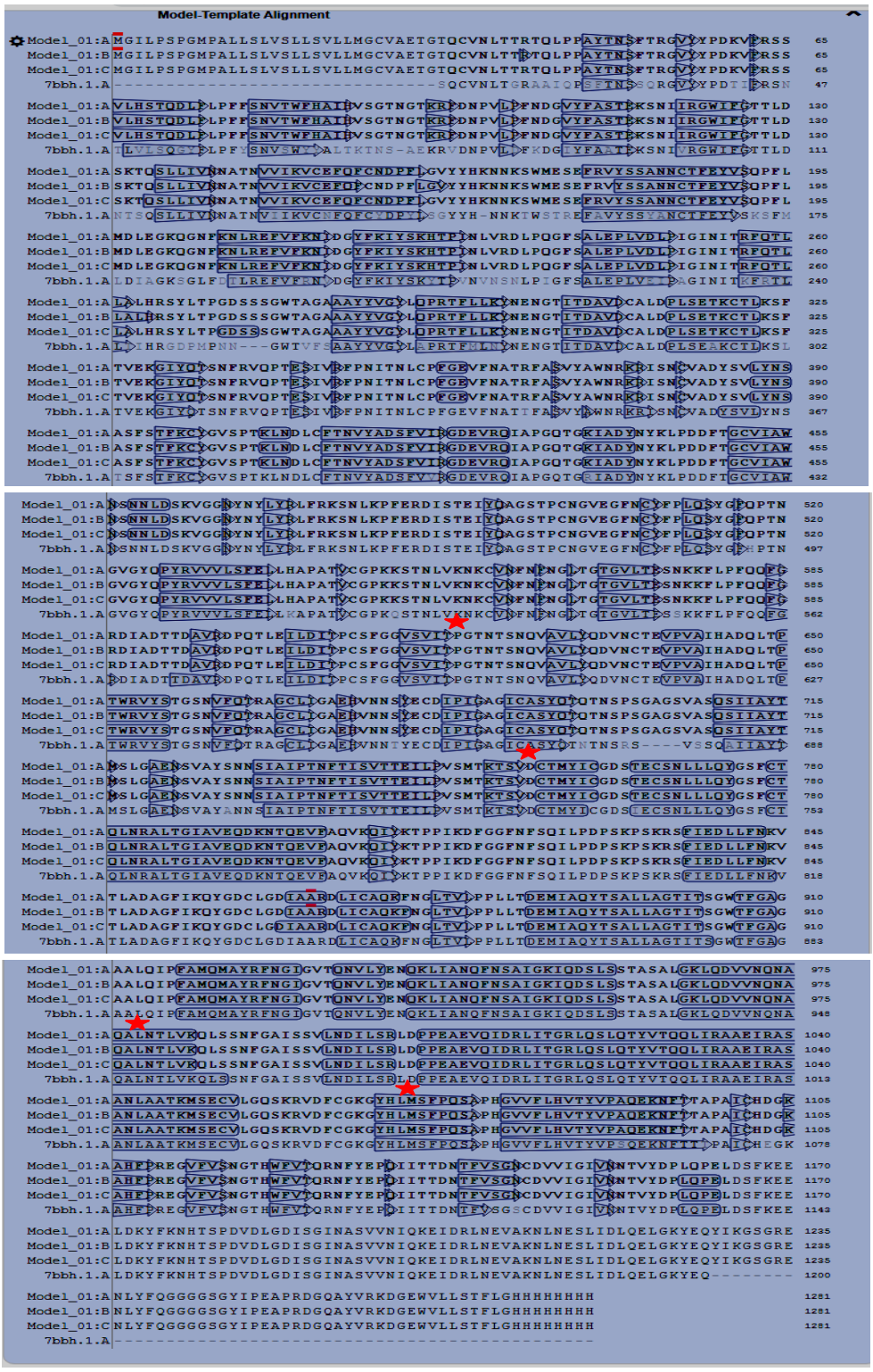
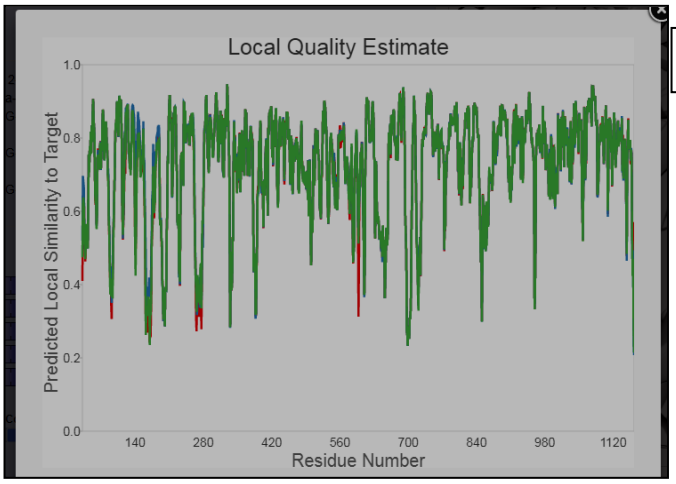
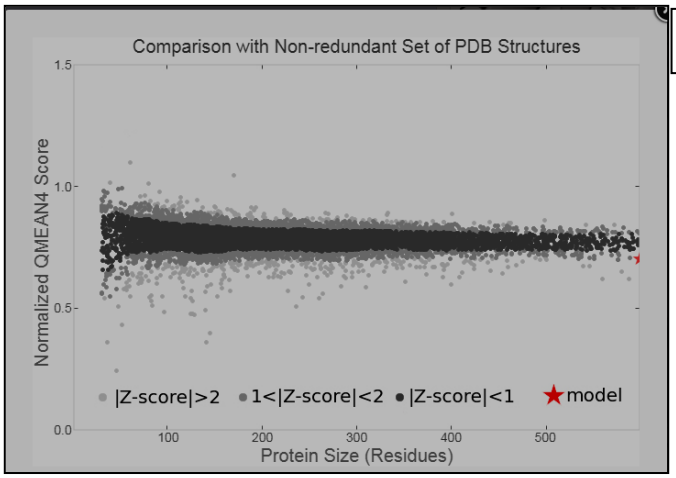


Fig. 2: Alignment of the amino acid sequences of modeled SARS-CoV 2 S protein and the crystal structure of 6crx 1 (7bbh) as generated by SWISS-MODEL web tool. ★: indicates active site residues involved in ligand interactions.


Local Quality Estimate
Predicted Local Similarity to Target
1.0
0.8
0.6
0.4
0.2
0.0
140
280
420
560
700
840
980
1120
Residue Number

A


Comparison with Non-redundant Set of PDB Structures
Normalized QMEAN4 Score
1.5
1.0
0.5
0.0
|Z-score|>2
1<|Z-score|<2
|Z-score|<1
model
100
200
300
400
500
Protein Size (Residues)

B

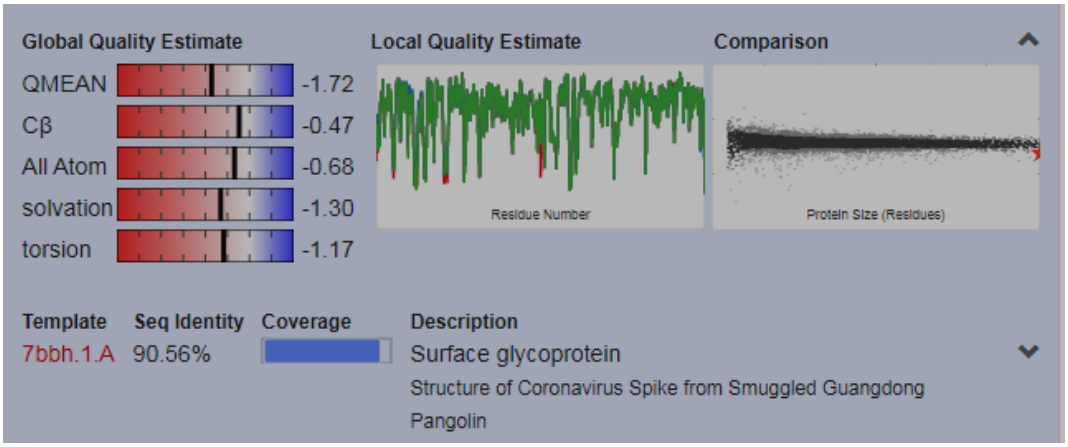
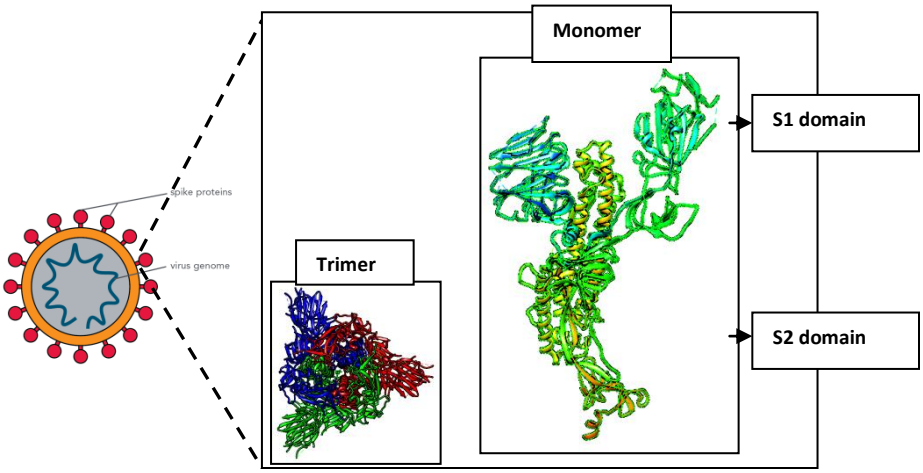

Global Quality Estimate
QMEAN -1.72
Cβ -0.47
All Atom -0.68
solvation -1.30
torsion -1.17
Local Quality Estimate
Residue Number
Comparison
Protein Size (Residues)
Template Seq Identity Coverage
7bbh.1.A 90.56%
Description
Surface glycoprotein
Structure of Coronavirus Spike from Smuggled Guangdong Pangolin

Fig 3: Validation of structure of modelled SARS-CoV-2 S protein: (A) Local quality estimate of the residues of the predicted Spike model; (B) comparison of the predicted S-protein structure with nonredundant set of PDB structures.



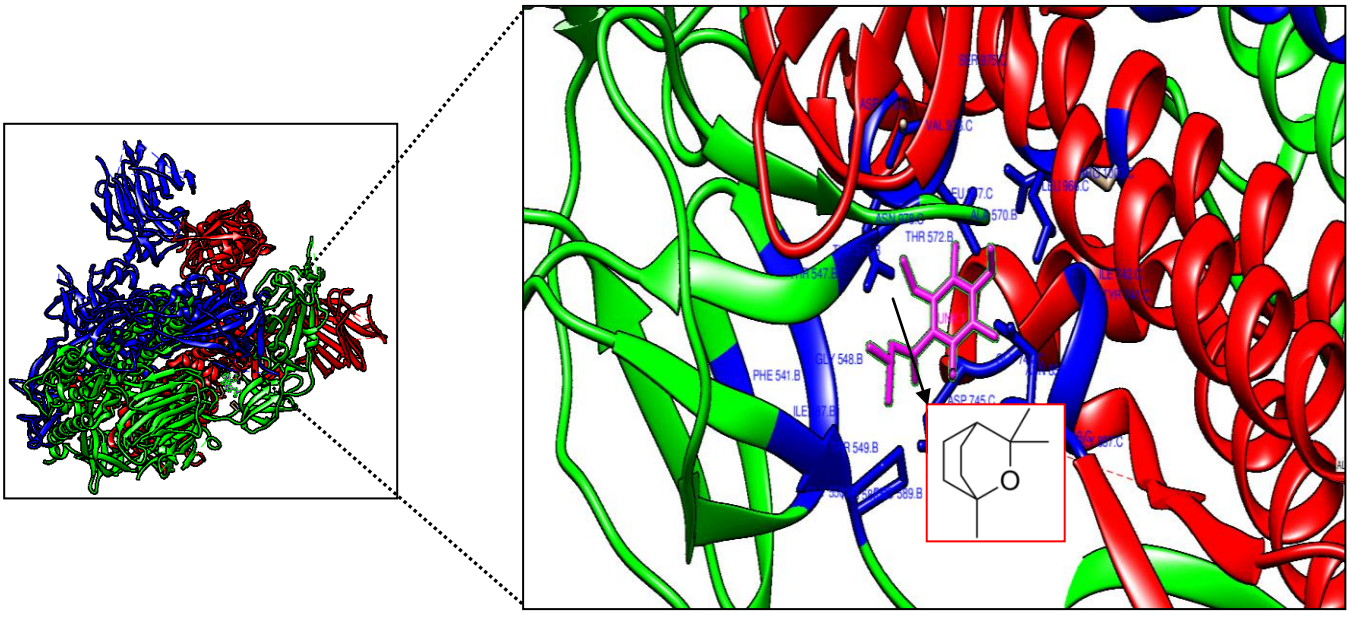
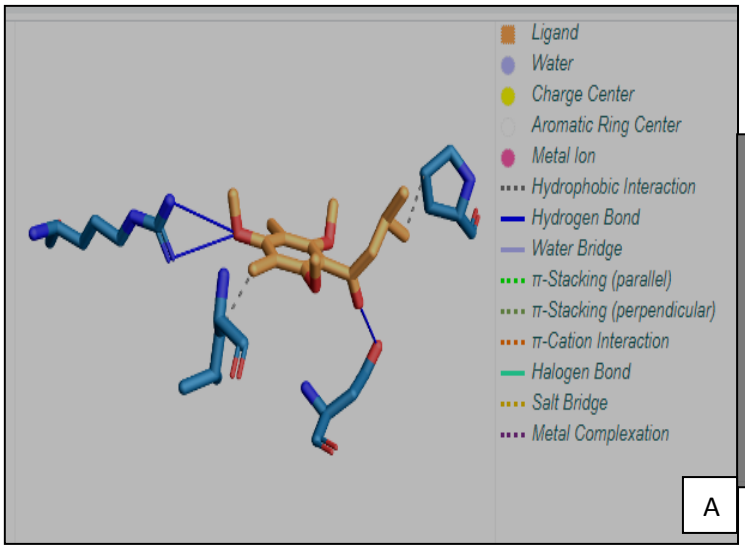


Fig. 4: Three-dimensional docked structure of SARS-CoV-2 S-protein with Toruatone. Light blue, green and the red color depicts Protomers of the trimer. Domains are labeled by S1and S2 Roman numbers.

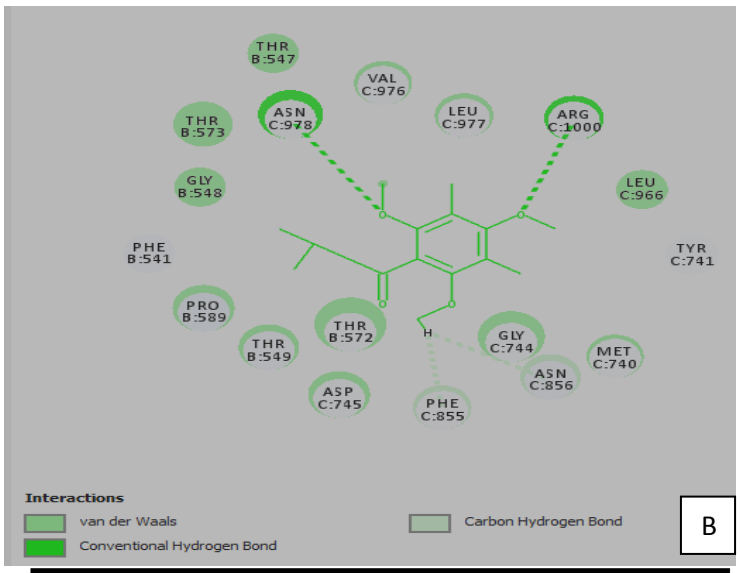


Hydrophobic Interactions

Index	Residue	AA	Distance	Ligand Atom	Protein Atom
1	589B	PRO	3.68	22829	11398
2	977C	LEU	3.11	22815	21481

Hydrogen Bonds

Index	Residue	AA	Distance H-A	Distance D-A	Donor Angle	Protein donor?	Sidechain	Donor Atom	Acceptor Atom
1	745C	ASP	2.05	2.79	134.37	✔	✔	19910 [O3]	22825 [O2]
2	1000C	ARG	2.99	3.48	112.38	✔	✔	21666 [Ng+]	22833 [O3]
3	1000C	ARG	3.66	4.07	108.10	✔	✔	21667 [Ng+]	22833 [O3]



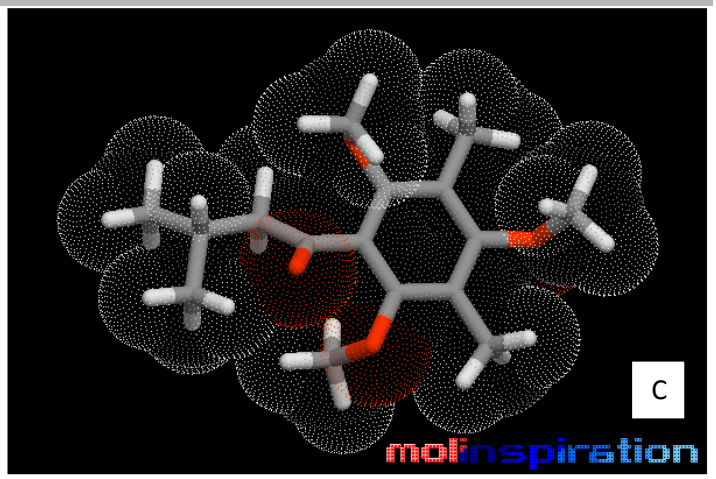


Fig. 5: Molecular docking interactions between Toruatone and SARS-CoV 2 Spike . (3-D: A, 2-D: B) and molecular lipophilicity potential (C) Hydrophobic areas: encoded by violet; Hydrophilic areas: red

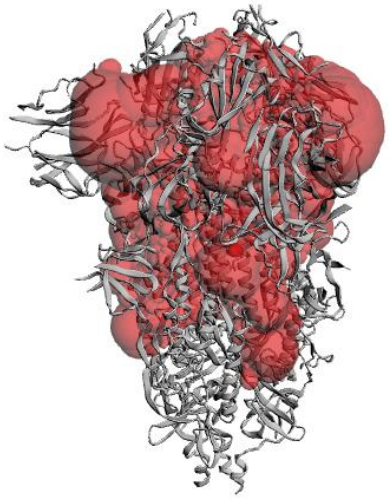


Fig. 6: Binding pocket of SARS-Cov2 S-protein as calculated by CASTp

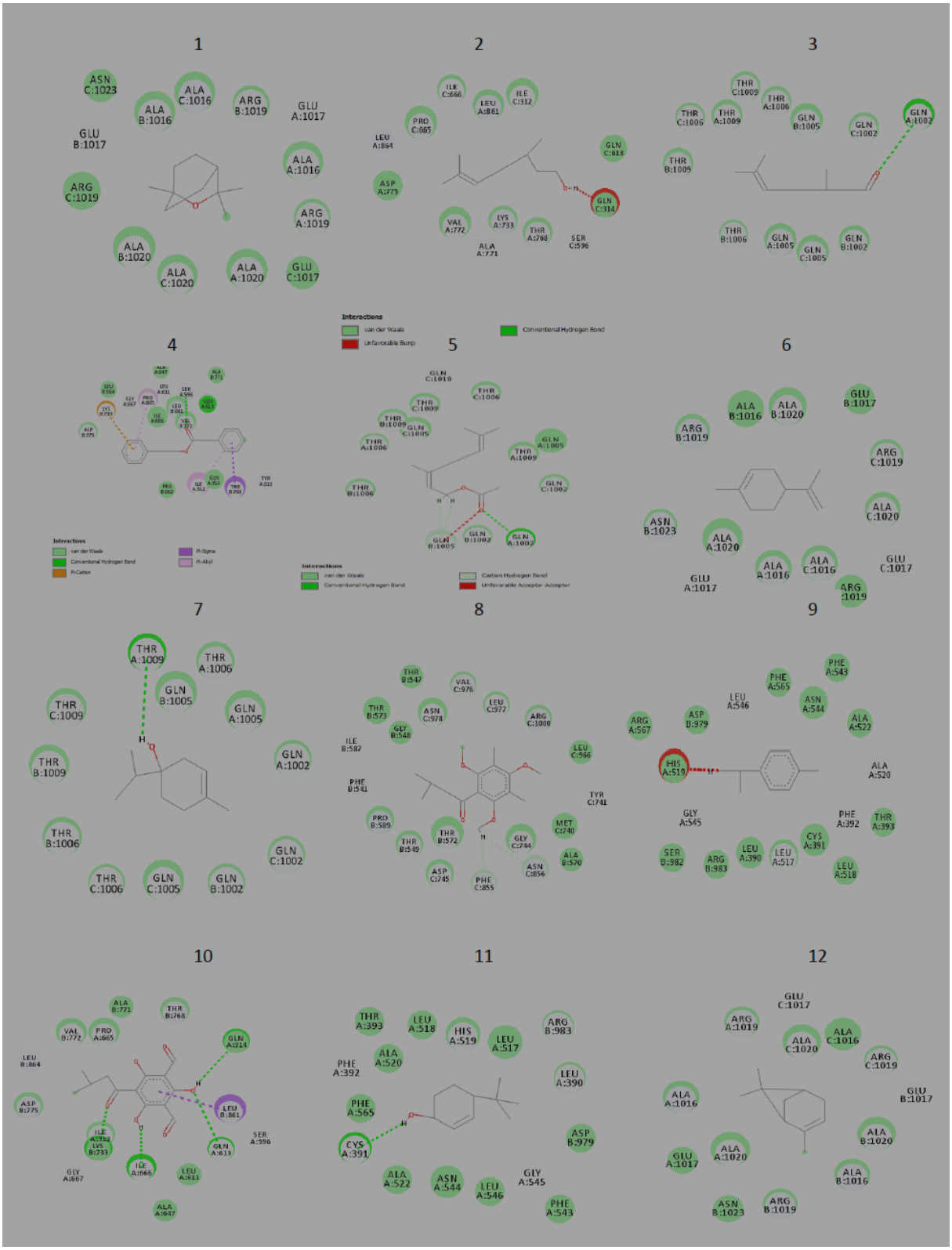


Fig 7: Molecular interactions of Eucalyptus oil bio-molecules with Spike proteins, number on each panel corresponds to bio-molecule as given in table 1

Table 2: ADMET properties of natural molecules

	Molecule	Lipinski's rule of five					GI absorption	ADMET properties						
		Molecular weight	Num. H-bond acceptors	Num. H-bond donors	Molar refractivity	lipophilicity logP		Consensus Log Po/w	TPSA Å^2	BBB permeant	P-gp substrate	CYP3A4 inhibitor	Log Kp (skin permeation)	Log S (ESOL)
1	EUCALYPTOL	154.25 g/mol	1	0	47.12	2.58	High	2.67	9.23	Yes	No	No	-5.30 cm/s	-2.52
2	CITRONELLOL	156.27 g/mol	1	1	50.87	2.72	High	2.92	20.23	Yes	No	No	-4.48 cm/s	-2.94
3	CITRONELLAL	154.25 g/mol	1	0	49.91	2.49	High	2.94	17.07	Yes	No	No	-4.52 cm/s	-2.88
4	PHENETHYL PHENYLACETATE	240.30 g/mol	2	0	71.60	2.95	High	3.50	26.30	Yes	No	No	-5.02 cm/s	-3.87
5	GERANYL ACETATE	196.29 g/mol	2	0	60.13	2.83	High	3.21	26.30	Yes	No	No	-4.63 cm/s	-2.52
6	LIMONENE	136.23 g/mol	0	0	47.12	2.72	Low	3.37	0.00	Yes	No	No	-3.89 cm/s	-3.50
7	4-TERPINEOL	154.25 g/mol	1	1	48.80	2.51	High	2.60	20.23	Yes	No	No	-4.93 cm.s	-2.78
8	P-CYMENE	134.22 g/mol	0	0	45.99	2.51	Low	3.50	0.00	Yes	No	No	-4.21 cm/s	-3.63
9	ALPHA-PINENE	136.23 g/mol	0	0	45.22	2.63	Low	3.44	0.00	Yes	No	No	-3.95 cm/s	-3.51
10	TORUATONE	280.36 g/mol	4	0	80.47	3.22	High	3.35	44.76	Yes	No	No	-5.52 cm/s	-3.62
11	JENSENONE	266.25 g/mol	6	3	67.90	0.76	High	1.30	111.9	No	No	No	-6.23 cm/s	-2.89
12	CIS-P-2-MENTHEN-1-OL	154.25 g/mol	1	1	48.50	2.46	High	2.46	20.23	Yes	No	No	-5.47 cm/s	-2.31

Table 3: Bioactivity score of TORUATONE

Molinspiration bioactivity score	
GPCR ligand	0.07
Ion channel modulator	-0.01
Kinase inhibitor	-0.29
Nuclear receptor ligand	0.52
Protease inhibitor	0.12
Enzyme inhibitor	0.88

Table 4: Toxicity profile of TORUATONE

ADMET predicted profile	Value	probability
Carcinogenicity (binary)	-	0.8173
Eye corrosion	-	0.6690
Ames mutagenesis	-	0.8600
Human either-a-go-go inhibition	-	0.6862
micronuclear	-	0.7426
Hepatotoxicity	-	0.5500
Acute Oral Toxicity (c)	III	0.5793
Estrogen receptor binding	-	0.5000
Androgen receptor binding	-	0.7063
Aromatase binding	-	0.5586
PPAR gamma	-	0.5558
Biodegradation	-	0.700