TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2020) 44: 242-264
© TÜBİTAK
doi:10.3906/biy-2005-46

Interaction of certain monoterpenoid hydrocarbons with the receptor binding domain of 2019 novel coronavirus (2019-nCoV), transmembrane serine protease 2 (TMPRSS2), cathepsin B, and cathepsin L (CatB/L) and their pharmacokinetic properties

Erman Salih İSTİFLİ[1], Arzuhan ŞIHOĞLU TEPE[2], Cengiz SARIKÜRKCÜ[3], Bektaş TEPE[4,*]

[1]Department of Biology, Faculty of Science and Literature, Çukurova University, Adana, Turkey
[2]Department of Biology, Faculty of Science and Literature, Gaziantep University, Gaziantep, Turkey
[3]Department of Analytical Chemistry, Faculty of Pharmacy, Afyonkarahisar Health Sciences University, Afyonkarahisar, Turkey
[4]Department of Molecular Biology and Genetics, Faculty of Science and Literature, Kilis 7 Aralık University, Kilis, Turkey

Received: 16.05.2020 • **Accepted/Published Online:** 09.06.2020 • **Final Version:** 21.06.2020

Abstract: As of June 2020, the coronavirus disease 19 (COVID-19) caused by the 2019 new type coronavirus (2019-nCoV) infected more than 7,000,000 people worldwide and caused the death of more than 400,000 people. The aim of this study was to investigate the molecular interactions between monoterpenoids and spike protein of 2019-nCoV together with the cellular proteases [transmembrane serine protease 2 (TMPRSS2), cathepsin B (CatB), and cathepsin L (CatL)]. As a result of the relative binding capacity index (RBCI) analysis, carvone was found to be the most effective molecule against all targets when binding energy and predicted (theoretical) IC_{50} data were evaluated together. It was found to exhibit drug-likeness property according to the Lipinski's rule-of-five. Carvone has also been determined to be able to cross the blood-brain barrier (BBB) effectively, not a substrate for P-glycoprotein (P-gp), not to inhibit any of the cytochrome P molecules, and to have no toxic effects even on liver cells. In addition, the LD_{50} dose of carvone in rats was 1.707 mol/kg. Due to its interaction profile with target proteins and excellent pharmacokinetic properties, it has been concluded that carvone can be considered as an alternative agent in drug development studies against 2019-nCoV.

Keywords: 2019-nCoV, ACE2, TMPRSS2, cathepsin, monoterpene, docking, pharmacokinetic

1. Introduction

Some Coronaviridae viruses are in circulation among people and are known to cause mild respiratory infections (Corman et al., 2019). However, it has been established that 2 important members of this family are transmitted from animals to humans and cause serious infections. These are severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV). Both types have led to severe respiratory infections in humans and the death of some individuals, especially those with chronic conditions (Fehr et al., 2017). SARS was first seen in Guangdong, China in 2002 and then quickly spread to other countries. This variant caused 8,096 people to become infected and 774 of these individuals died[1] (De Wit et al., 2016). The main source of SARS-CoV has been found to be Chinese horseshoe bats (Lau et al., 2005; Li et al., 2005). It has been reported to be transferred to humans through civet cats and raccoon dogs sold as food at Chinese wet markets (Guan et al., 2003). There is no approved antiviral agent or vaccine used in the treatment of SARS, either at the time of its first appearance or now. The spread of the pandemic that emerged in the period of 2002–2003 has been prevented by using traditional methods such as restricting people's travel and isolation of sick individuals, just like today (Hoffmann et al., 2020).

After the epidemics of SARS and MERS in the early millennium, a new and highly contagious respiratory disease was detected in Wuhan (China) towards the end of 2019 (Huang et al., 2020; Wang et al., 2020; Zhu et al., 2020). It was reported that the first infected were people who came into contact with animals in the Huanan seafood market. Later it became clear that the virus could spread among humans (Chan et al., 2020). The so-called coronavirus disease 19 (COVID-19) spread rapidly to all parts of China. A new variant of SARS virus was found

[1]WHO (2003). Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003 [online]. Website https://www.who.int/csr/sars/country/table2004_04_21/en/ [accessed 00 Month Year].

* Correspondence: bektastepe@yahoo.com

in the analysis of samples taken from sick individuals. The pathogen that caused the disease was called SARS-coronavirus 2 (SARS-CoV-2) or 2019 new coronavirus (2019-nCoV) because it is from the same family as SARS-CoV (Zhu et al., 2020). In February 2020, nearly 45,000 cases were detected in China. It has been announced that about 8,000 of these cases are in critical condition and over 1,000 people have died[2]. The virus has spread to about 2 dozen countries, primarily through people traveling from China to other countries. As of June 2020, the number of COVID-19 cases worldwide has exceeded 7,000,000. More than 400,000 of these patients died as of early June 2020[3]. At present, it is not possible to make any inference about the sequence similarities between the SARS-CoV-2 and the SARS-CoV on the pandemic properties of these 2 variants (Munster et al., 2020).

In both coronavirus variants, spike protein plays an important role in the entry of the pathogen into the cell. Entry occurs as a result of the interaction between the S1 subunit of the spike protein and the receptor on the surface of the target cell. However, for entry, priming of the spike protein by cellular proteases is required. In this process, the spike protein is cut at the S1/S2 point and the S2'unit provides the junction between the virus and the cellular membrane. The virus needs angiotensin-converting enzyme 2 (ACE2) to bind to the receptor on the target cell surface (Li et al., 2003). Research has shown that transmembrane serine protease 2 (TMPRSS2) is involved in the priming of spike proteins of SARS viruses (Matsuyama et al., 2010; Glowacka et al., 2011; Shulla et al., 2011). Since TMPRSS2 is actively involved in the priming process of the 2019-nCoV spike protein, there is some evidence that camostat mesylate, which has an inhibitory effect on this protease, prevents infection in lung cells. Cathepsin B and cathepsin L (CatB/L), which are endosomal cysteine proteases, are also thought to be involved in priming of the spike protein of 2019-nCoV. Thus, it has been shown that inhibition of these proteases may also prevent the virus from entering the cell (Hoffmann et al., 2020).

Plant secondary metabolites are synthesized by many species such as vegetables, fruits, medicinal and aromatic plants (Prakash et al., 2007; Singh et al., 2009; Singh et al., 2009; Singh et al., 2009; Singh et al., 2010). Some phytochemicals have been reported to have significant antiviral activity. Thus, great attention has been paid to plant secondary metabolites in the treatment of some viral infections. Various studies have been conducted on the potential of some of these phytochemicals to inhibit the receptor binding domain (RBD) of the spike protein of 2019-nCoV, cellular proteases or endoribonucleases (Meneguzzo et al., 2020; Sampangi-Ramaiah et al., 2020; Tallei et al., 2020; Thuy et al., 2020).

The aim of this study was to investigate the molecular interactions between monoterpenoid hydrocarbons (Figure 1 and Table 1), which constitute an important group of plant essential oils, and i) RBD of the spike protein of 2019-nCoV and ii) cellular proteases (TMPRSS2, CatB and CatL). In addition, drug-likeness properties and ADMET profiles of monoterpenoids were presented. Using the binding free energy (kcal/mol) and predicted (theoretical) IC_{50} (mM) values of the monoterpenoids, the relative binding capacity index (RBCI) values were also statistically calculated and 'hit' compounds were determined.

2. Materials and methods

2.1. Computational capacity

Two Dell laptops (Windows 8.1 and Windows 10 operating system) were used to perform the molecular docking analyses in the current study. One of the computers had the Windows 10 operating system and was equipped with the Intel Core i5-7200U CPU 2.50 GHz and 2.70 GHz processors. The other computer with the Windows 8 operating system had Intel Core i5-5200U CPU 2.20 GHz and 2.20 GHz processor power.

2.2. Structural optimization of ligands

The protein data bank (.pdb) files of all the ligands have been downloaded from PubChem (https://pubchem.ncbi.nlm.nih.gov) via the download module of Vega ZZ 3.2.0.9 software. In the Vega ZZ, the atom types and electrical charges of the ligands were fixed with MMFF94 force field and Gasteiger-Marsili parameters (Pedretti et al., 2004). The ligands were energetically minimized by the conjugate gradient minimization method. For this purpose, the minimization steps and tolerance were set to 1000 and 0.01, respectively.

2.3. Energy minimization of 2019-nCoV ACE2-RBD, TMPRSS2, CatB/L using nanoscale molecular dynamics (NAMD)

Firstly, in the Vega ZZ environment, the structure of the spike glycoprotein was gained by removing the ACE2 subunit from the angiotensin-converting enzyme 2 - 2019-nCoV receptor binding domain (RBD) complex which was downloaded from the url: https://swissmodel.expasy.org/interactive/HLkhkP/models/03 (PDB ID: model_03.pdb) (Camacho et al., 2009; Remmert et al., 2011). Since the structure of the spike glycoprotein in model_03 shows

[2]WHO (2020). Novel Coronavirus (2019-nCoV): situation report, 3 [online]. Website https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200212-sitrep-20200223-ncov.pdf?sfvrsn=20200241e20200219fb20200278_20200214 [accessed 08 May 2020].

[3]Worldometer (2020). COVID-19 coronavirus pandemic [online]. Website https://www.worldometers.info/coronavirus/ [accessed 08 May 2020].

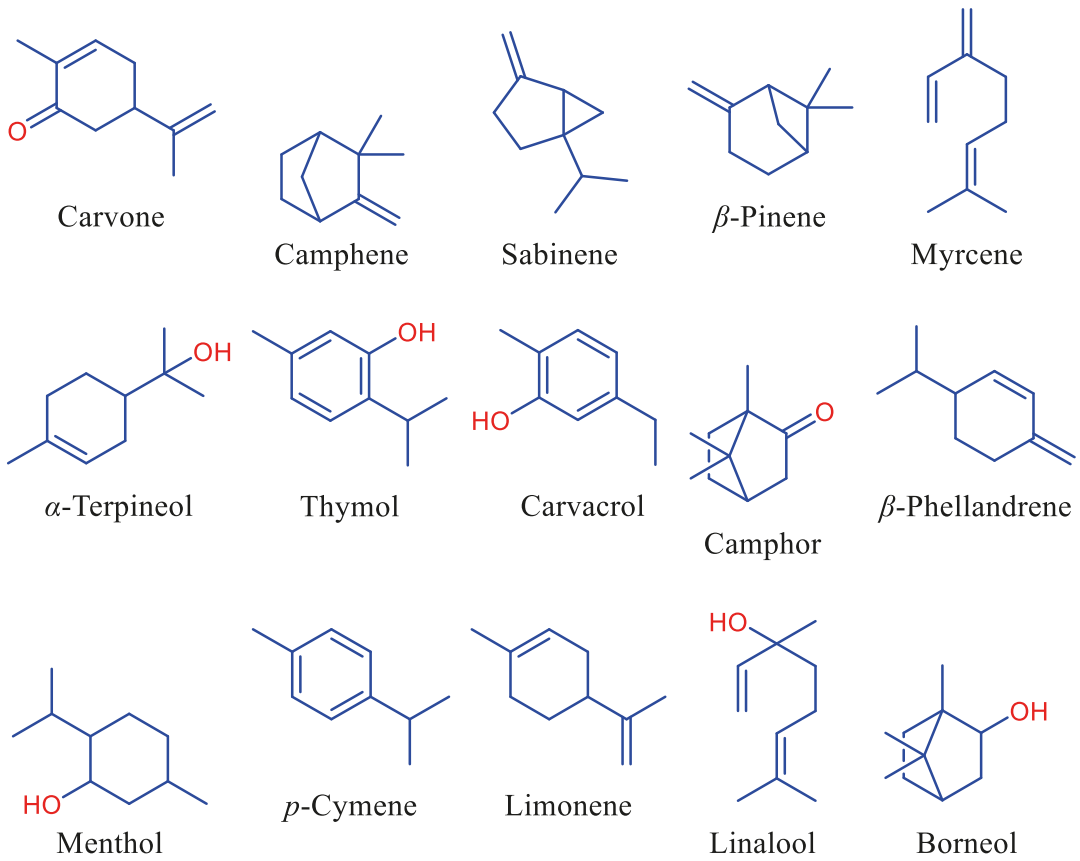

Figure 1. Chemical structures of the monoterpenoids.

a sequence identity of 99.88% to the 2019-nCoV ACE2-binding domain, this model was chosen as the appropriate 3D structure in the docking analyses. During the protein preparation step, the atom types and electrical charges of the spike glycoprotein were fixed using CHARMM22_PROT force field and Gasteiger-Marsili charges. Next, for the energy minimization of the spike glycoprotein with NAMD, each parameter was loaded from a template file. The number of time steps (number of minimization steps) were set to 10,000 and CHARMM22_PROT was set as the force field. When the energy minimization was completed, the 3D structure corresponding to the last minimization step was saved as the lowest energy conformation. Also, to keep the spike glycoprotein structurally closer to the original crystallographic data, atom constraints were also applied to the protein backbone. In the energy minimization of TMPRSS2, CatB, and CatL, the same steps described for the 2019-nCoV spike glycoprotein were applied.

2.4. Homology modeling of TMPRSS2

Since the crystallographic data of TMPRSS2 enzyme structure has not been resolved until today, we generated

Table 1. PubChem CID, molecular weight and molecular formula of the compounds.

No	Compound	PubChem CID	Molecular weight (g/mol)	Molecular formula
1	Carvone	7439	150.22	$C_{10}H_{14}O$
2	Camphene	6616	136.23	$C_{10}H_{16}$
3	Sabinene	18818	136.23	$C_{10}H_{16}$
4	β-Pinene	14896	136.23	$C_{10}H_{16}$
5	Myrcene	31253	136.23	$C_{10}H_{16}$
6	α-Terpineol	17100	154.25	$C_{10}H_{18}O$
7	Thymol	6989	150.22	$C_{10}H_{14}O$
8	Carvacrol	10364	150.22	$C_{10}H_{14}O$
9	Camphor	2537	152.23	$C_{10}H_{16}O$
10	β-Phellandrene	11142	136.23	$C_{10}H_{16}$
11	Menthol	1254	156.26	$C_{10}H_{20}O$
12	*p*-Cymene	7463	134.22	$C_{10}H_{14}$
13	Limonene	22311	136.23	$C_{10}H_{16}$
14	Linalool	6549	154.25	$C_{10}H_{18}O$
15	Borneol	64685	154.25	$C_{10}H_{18}O$

a homology model to use in docking analyses with this enzyme. The amino acid sequence of TMPRSS2 was downloaded from UniProtKB (https://www.uniprot.org/uniprot/O15393). Template search for TMPRSS2 catalytic domain was performed against the SWISS-MODEL template library with BLAST and HHBlits. BLAST was used to search the TMPRSS2 catalytic domain target sequence against the primary amino acid sequence in the SMTL. As a result of the BLAST search, a total of 788 templates were found. An initial HHblits profile has been built using the procedure as described in (Remmert et al., 2011). This procedure was followed by 1 iteration of HHblits against NR20. The obtained profile was then searched against all profiles of the SMTL and, finally, a total of 1167 templates were found.

ProMod3 was used to carry out model building for TMPRSS2 catalytic domain based on the target-template alignment. The coordinates preserved between the target structure and the template were copied from the template to the model. Insertion and deletions were remodeled based on the fragment library. Subsequently, the side chains were rebuilt. Finally, using the CHARMM27 force field, the geometry of the resulting TMPRSS2 model was optimized. In cases where ProMod3 failed in loop modeling, an alternative model was developed with PROMOD-II (Guex et al., 2009).

The model quality (global and per-residue) of TMPRSS2 obtained was evaluated with the QMEAN scoring function (Studer et al., 2020). The near-zero QMEAN score reflects a good agreement between the model structure and the experimental structure, although scores of –4.0 and below indicate that the model is of low quality. Therefore, among the top 5 TMPRSS2 templates obtained as a result of homology modeling, the 5ce1.1.A (model 06) template with the QMEAN score closest to zero (QMEAN = –1.43) was selected as the target in the docking analysis.

In addition, whether our model has an energetically favorable conformation was analyzed by generating a Ramachandran plot in the PROCHECK (Laskowski et al., 1993) web-based tool. ERRAT (Colovos and Yeates 1993) online web-based tool was also deployed to calculate the overall quality factor (OQF) for nonbonded atomic interactions.

2.5. Molecular docking analyses

Molecular docking analyses was performed using AutoDock 4.2 to predict the binding affinities of carvone, camphene, sabinene, beta-pinene, myrcene, alpha-terpineol, thymol, carvacrol, camphor, β-phellandrene, menthol, *p*-cymene, limonene, linalool, borneol with 2019-nCoV RBD (PDB ID: model_03, https://swissmodel.expasy.org/interactive/HLkhkP/models/03),TMPRSS2 (model_06,https://swissmodel.expasy.org/interactive/HMKd4q/models/), cathepsin B (PDB ID: 1GMY), and cathepsin L (PDB ID: 2YJ9). AutoDockTools-1.5.6 was used to prepare the target and ligand molecules and also the parameters prior to initiating the docking analysis using AutoDock 4.2 (Sanner, 1999). In this study, the grid box coordinates used in molecular docking analyzes were adjusted to ensure that all the tested phytochemicals interact with amino acids in the active sites of the enzymes in question (Greenspan et al., 2001; Wilson et al., 2005; Hardegger et al., 2011; Andersen et al., 2020).

In the molecular docking analyzes, polar hydrogen atoms in the receptor and the ligand molecules were retained while nonpolar hydrogens were merged and then, the Gasteiger charges of the ligands were calculated with AutoDockTools as previously described (Ricci and Netz 2009; Nasab et al., 2017). In addition, the Kollmann charges were added for the receptor. During the docking experiments, all the rotatable bonds of the ligands were allowed to rotate and then the optimized protein (rigid) and ligand (flexible) structures were saved in PDBQT format. Grid box coordinates were adjusted as: a) 80 × 90 × 40 Å points for the spike glycoprotein; b) 60 × 110 × 86 Å points for TMPRSS2; c) 86 × 84 × 44 Å points for CatB; and d) 54 × 52 × 60 Å points for CatL. Prior to docking analyses, these grid box sizes were determined to include the active amino acid residues of these enzymes.

In all docking analyses, 50 genetic algorithm (GA) runs using an initial population of 150 individuals, maximum number of 2,500,000 energy evaluations, and a maximum number of 27,000 generations were selected. The values of 0.02 and 0.8 were chosen as the default parameters for mutation and crossover rates, respectively. After 50 independent docking runs, all the possible binding modes (conformations) of the ligands were clustered by the program and were ranked based on the lowest RMSD (root mean square deviation) and the binding free energy (kcal/mol) of the ligand conformation. The best docking poses obtained using the AutoDock 4.2 between the ligand and receptor structures was analyzed with the BIOVIA Discovery Studio Visualizer 2016.

2.6. Success criteria set in docking analysis

In the current study, the RMSD (root mean square deviation) value of the docking results obtained for each phytochemical analyzed was considered successful when it was less than 2 angstroms (<2 Å). The criterion considered after the RMSD value was the binding energy (deltaG) of the ligand in the most efficient docked complex. Briefly, the closeness of all the phytochemicals tested in this study to the ACE2-RBD, TMPRSS2, CatB and CatL, and then the energy of that binding in these zones were determined (Morris and Lim-Wilby, 2008). The calculated inhibition constants (K_i) obtained with AutoDock 4.2 for each docked phytochemical were also given.

2.7. Drug-likeness prediction, ADMET profile and target prediction

The drug-likeness, ADMET and target profiles of potential hit compounds are very important in terms of reducing side effects in the pharmaceutical industry. In our study, web-based SwissADME, pkCSM and Swiss Target Prediction online tools were used to determine such effects of monoterpene hydrocarbons analyzed (Pires et al., 2015; Daina et al., 2017; Daina et al., 2019).

2.8. Calculation of RBCI values

A new analysis method called RBCI was applied to statistically rank the activity potentials of phytochemicals by using binding energy and IC_{50} values obtained from the parameters given above. Through this analysis, it is possible to compare data, each of which has different scientific meanings, statistically with each other. If sorting based on the interaction of molecules with proteins is performed according to only 1 of these parameters (eg binding energy or IC_{50} value only), the molecules can only be sorted in terms of their potential in that parameter. However, sorting using only 1 of these parameters cannot represent the activity potential of these molecules from all parameters.

The most commonly used method to determine the interaction between the receptor-ligand in multiple measurements is the central tendency, in which the components are ranked-based on the mean value for each component (Zar, 1996). However, since the units and scales of the data obtained from each parameter are different, it is not possible to obtain an average value for all components.If the values in each data set (binding energy and IC_{50}) are converted to standard scores, it is possible to compare them with each other.

In order to calculate the arithmetic mean values, first of all, binding energy and IC_{50} data of each phytochemical were used regardless of their units and raw values were obtained. These raw values calculated for each component were subtracted from the arithmetic mean and divided by standard deviation, and standard scores were obtained (see equation given below) (Sharma, 1996). RBCI values of each phytochemical was calculated by averaging these standard scores obtained separately for each protein target.

$$Standart\ score = \frac{(x - \mu)}{\sigma}$$

where 'x' is the raw data, 'μ' is the mean, and 'σ' is the standard deviation.

Although RBCI is a relative index and does not represent the specific binding capacities of the components, it makes it possible to rank components reasonably based on their binding energy and IC_{50} values. Therefore, it can be used as an integrated approach to evaluate the molecular interaction of the components, considering all parameters.

3. Results

3.1. Homology modeling

Of the 5 models created by ProMod3, model 06, which had the QMEAN scoring function closest to zero (QMEAN = –1.43) was chosen. The sequence identity of TMPRSS2 (model 06) created by ProMod3 was 33.82% and sequence similarity was 38%. If the target and template sequences show an amino acid identity over 30% and above, homology modeling is accepted as being reliable and successful (Xiang, 2006). To further verify model used in this study, a Ramachandran plot was created through the PROCHECK web-based tool to evaluate the energetically allowed regions of TMPRSS2 (Figure 2). According to plot statistics, 84.6% of residues were in most favored regions, 14.7% was in additional allowed regions, 0.3% was in generously allowed regions, and 0.3% of the residues were in disallowed regions. ERRAT web-based tool was also used to calculate the overall quality factor (OQF) of nonbonded atomic interactions of the model. According to this online web-based tool, the OQF of a high quality model should be above 91%. The OQF calculated by the ERRAT server of TMPRSS2, which was created as the homology model, was 92.92% (Figure 3).

3.2. RMSD values

RMSD values of atomic positions obtained as a result of computer-based interactions of monoterpenoid hydrocarbons with spike, TMPRSS2, CatB and CatL are given in Tables 1, 2, 3, 4, and 5, respectively. RMSD values obtained as a result of the interaction of monoterpenoids with the proteins in question were in the range of 0.02–1.97, 0.01–20.103, 0.00–13.569, and 0.01–1.14 Å, respectively. The compound showing the lowest RMSD value in interaction with spike, TMPRSS2 and CatB was camphene (0.02, 0.01, and 0.00 Å, respectively). On the other hand, RMSD values of borneol, sabinene, β-pinene, camphor, and carvone were also found to be quite low. With some exceptions, the RMSD values of monoterpenoids on all proteins were below 2.0 Å. The RMSD values obtained as a result of the interactions of menthol with TMPRSS2 and *p*-cymene with Cat B were found to be quite high (20.103 and 13.569 Å, respectively).

3.3. Binding energies

The binding energies of monoterpenoids to target molecules ranged from –5.01 to –2.38 kcal/mol. The binding energies between monoterpenoids and spike, TMPRSS2, CatB and CatL were in the ranges of –5.0/–3.6, –4.45/–2.38, –5.53/–3.78, and –5.03/–3.8 kcal/mol, respectively. The binding energy of carvone against all proteins was found to be significantly lower than the others. In addition, the binding energies of menthol, camphor, and α-terpineol were also found to be promising. Linalool and myrcene were the molecules with the highest binding energy against all proteins.

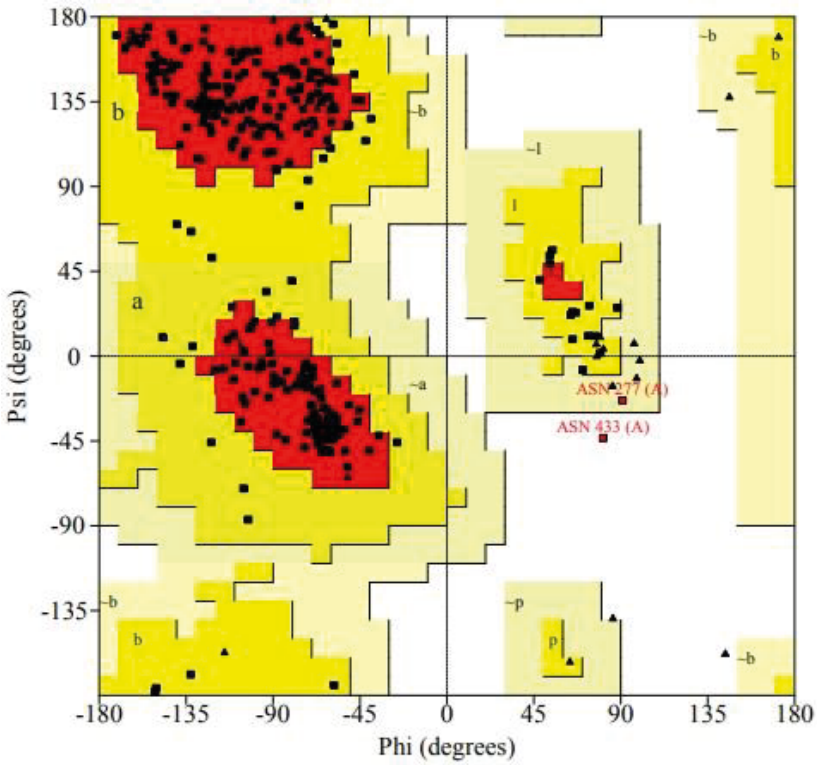

Figure 2. Ramachandran plot of TMPRSS2 model.

3.4. Predicted IC_{50} values

The predicted IC_{50} values of monoterpenoids against spike, TMPRSS2, CatB, and CatL were in the range of 0.213–2.29, 0.547–17.89, 0.087–1.7 and 0.205–1.65 mM, respectively. As with the binding energies of the molecules, the IC_{50} value of carvone was found to be significantly lower than other monoterpenoids. In addition, the IC_{50} values of camphor, menthol, and α-terpineol against all protein targets were also low. IC_{50} values of myrcene and linalool were quite high compared to other monoterpenoids. In particular, IC_{50} value of linalool against TMPRSS2 was determined as 17.89 mM.

3.5. Bonding interactions

The bonds between the monoterpenoids and contact residues in the active regions of the target proteins are detailed in Tables 1, 2, 3, and 4. Van der Waals and hydrophobic interactions (both p-p and mixed p/alkyl) played an important role in the interaction between monoterpenoids and RBD of the spike protein of 2019-nCoV. Van der Waals forces and hydrophobic interactions also played a decisive role in the molecular interactions between monoterpenoids and TMPRSS2, CatB, and CatL. However, it was found that hydrophobic effects between monoterpenoids and target proteins were rather established through mixed p/alkyl interactions.

In general, no interaction has been formed between monoterpenoids and Leu455, Phe486, and Gln493, amino acids involved in binding ACE2 in the RBD of the spike protein of 2019-nCoV (Figure 4). However, it was found that molecular interaction was formed between other active amino acids (Asn501 and Tyr505) and monoterpenoids. In addition, monoterpenoids were found to interact with other amino acids (Arg403, Tyr453, Tyr495, Gly496 and Phe497) of the spike protein.

As understood from heatmap, which shows the molecular interaction between monoterpenoids and TMPRSS2 (Figure 5), carvone and carvacrol were the most intensely interacting molecules with the protein in question. On the other hand, the number of bonds between TMPRSS2 and camphene, α-terpineol, thymol, camphor, and linalool was less than others. While molecular interaction was formed between monoterpenoids and the active amino acids of TMPRSS2, His296 and Ser441, no interaction with the other active amino acid, Asp345, was achieved. Monoterpenoids also showed several interactions with some other amino acids of TMPRSS2 (His279, Val280, Cys281, Thr393, Gln438, and Gly439).

The vast majority of the active amino acids of CatB (Gln23, Gly24, Cys26, Gly27, Cys29, His110, His111, His199 and Trp221) were found to have formed

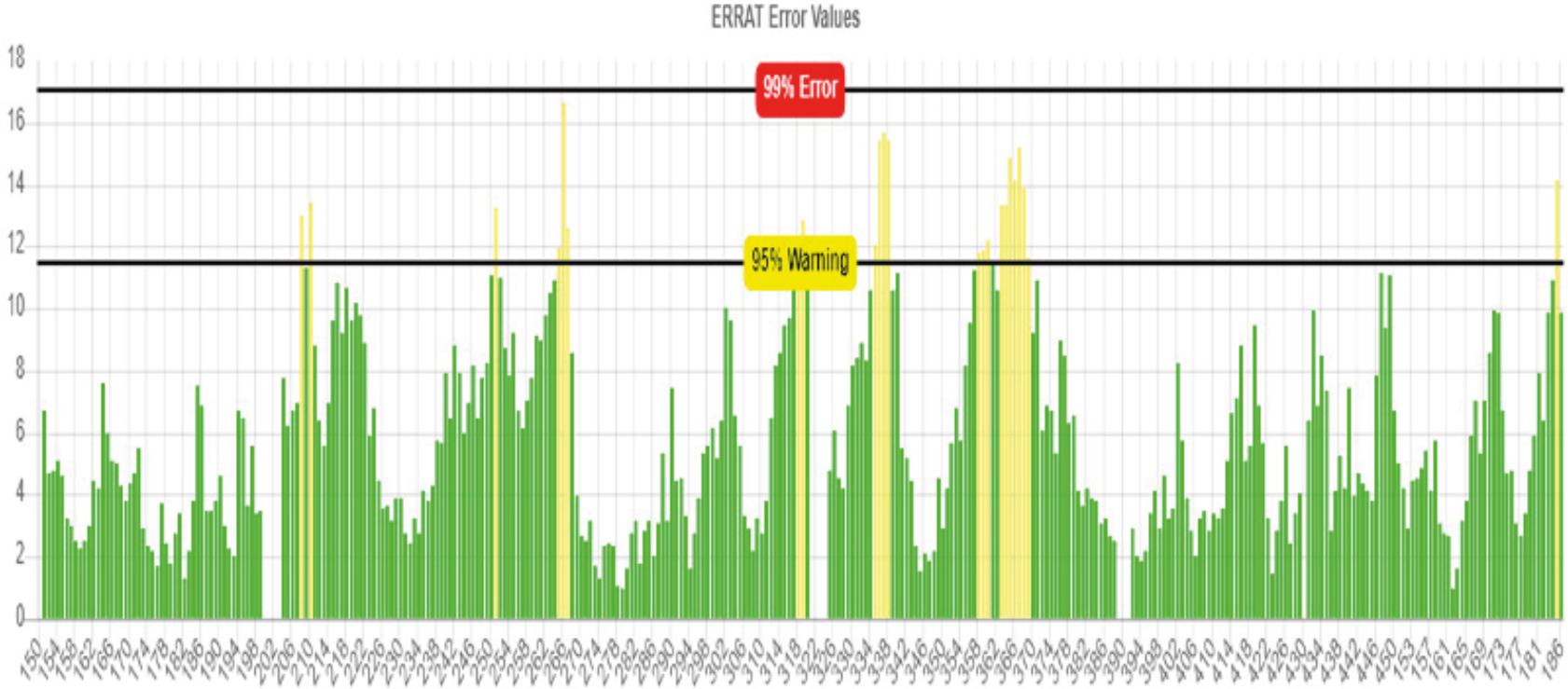

Figure 3. ERRAT error values for TMPRSS2 model. (Protein regions show misfolding at 95% confidence level are indicated with yellow bars. Green bars, on the other hand, point to regions that show correct folding.)

Table 2. Molecular interactions between the monoterpenoids and RBD of the spike protein of 2019-nCoV.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
1	Carvone	0.15	-4.70	0.360	Asn501[2]	Ser494[2], Gly496, Phe497	-	Arg403, Tyr453, Tyr495, Tyr505[2]	-	-	-
2	Camphene	0.02	-4.77	0.320	-	Ser494[2], Gly496, Asn501[2]	-	Arg403, Tyr453, Tyr495, Phe497, Tyr505[2]	-	-	-
3	Sabinene	0.03	-4.54	0.469	-	Ser494[2], Gly496, Asn501[2]	-	Arg403, Tyr453, Tyr495, Phe497, Tyr505[2]	-	-	-
4	β-Pinene	0.12	-4.54	0.469	-	Glu406, Ser494[2], Gly496, Asn501[2]	-	Arg403, Tyr453, Tyr495, Phe497, Tyr505[2]	-	-	-
5	Myrcene	0.87	-3.85	1.510	-	Glu406, Ser494[2], Gly496, Asn501[2]	-	Tyr453, Tyr495, Arg403, Phe497, Tyr505[2]	-	-	-
6	α-Terpineol	0.08	-4.61	0.417	Gly496, Ser494[2]	Arg403, Glu406	-	Tyr453, Tyr495, Tyr505[2]	-	-	-
7	Thymol	0.21	-4.69	0.363	Gly496, Ser494[2]	Arg403, Asn501[2], Tyr505[2]	-	Phe497, Tyr495	Tyr453	-	-
8	Carvacrol	0.05	-4.56	0.457	Tyr505[2]	Arg403, Ser494[2], Gly496, Phe497, Asn501[2]	-	Tyr505[2]	Tyr453, Tyr495, Tyr505[2]	-	-
9	Camphor	1.88	-4.28	0.727	Gly496	Arg403	-	-	Tyr453, Tyr495, Tyr505[2]	-	-
10	β-Phellandrene	0.37	-4.56	0.453	-	Arg403, Ser494[2], Gly496, Phe497, Asn501[2]	-	-	Tyr453, Tyr495, Tyr505[2]	-	-

Table 2. Continued.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
12	*p*-Cymene	1.97	-4.17	0.875	-	Arg403, Ser494[2], Gly496, Phe497, Asn501[2]	-	Tyr505[2]	Tyr453, Tyr495, Tyr505[2]	-	-
13	Limonene	0.03	-4.59	0.433	-	Gln493[2], Ser494[2], Gly496, Asn501[2]	-	-	Arg403, Tyr453, Tyr495, Phe497, Tyr505[2]	-	-
14	Linalool	1.70	-3.60	2.290	-	Glu406, Ile418, Gly496, Asn501[2]	Arg403	-	Tyr453, Lys417, Arg403, Tyr453, Tyr495, Phe497, Tyr505[2]	-	-
15	Borneol	0.02	-4.49	0.507	Gly496, Ser494[2]	Arg403	-	-	Tyr453, Tyr495, Tyr505[2]	-	-

[1]RMSD: Root mean square deviation (Angstrom = Å).
[2]Amino acid residues involved in binding to ACE2 in the receptor binding domain of 2019-nCoV (Leu455, Phe486, Gln493, Ser494, Asn501, and Tyr505).

Table 3. Molecular interactions between the monoterpenoids and TMPRSS2.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
1	Carvone	0.07	-4.45	0.547	His279	Cys297, Thr393, Gln438, Gly439, Ser441[2]	Val280	-	Val280, Cys281, His279, His296[2]	-	-
2	Camphene	0.01	-3.79	1.660	-	Cys297, Glu299, Trp308	-	-	Pro301, Leu302, Val280, His296[2]	-	-
3	Sabinene	1.42	-3.48	2.800	-	His279, Cys281, His296[2], Thr393, Cys437, Gln438, Gly439, Ser441[2]	-	-	Val280	-	-
4	β-Pinene	0.04	-3.84	1.520	-	Ser436, Cys437, Gln438, Ser441[2], Ser460, Trp461, Gly462, Gly464	-	-	His296[2]	-	-
5	Myrcene	0.80	-3.27	3.990	-	Trp384, Thr393, Gln438, Gly439, Ser441[2]	-	-	Val280, Cys281, Cys297, His279, His296[2]	-	-
6	α-Terpineol	1.83	-4.17	0.882	Glu299	Asp338, Ser339, Thr341	-	-	Tyr337, Lys342, Lys340	-	-
7	Thymol	1.04	-3.84	1.540	His296[2]	Ala295, Cys297, Val298, Glu299	-	-	Val280, Pro301, Leu302	-	-
8	Carvacrol	0.57	-3.52	2.650	His279	Val278, Cys281, His296[2], Thr393, Gln438, Gly439, Ser441[2]	-	-	Val280, His279	Val280	-
9	Camphor	0.04	-4.27	0.737	Lys340	Thr341	-	-	Lys342, Leu419, Trp461	-	-

Table 3. Continued.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
12	*p*-Cymene	0.86	-3.29	3.880	-	Val278, Cys281, His296[2], Thr393, Gln438, Gly439, Ser441[2]	-	-	Val280	His279, Val280	-
13	Limonene	0.16	-3.59	2.330	-	His279, Cys297, Thr393, Gln438, Gly439, Ser441[2]	-	-	Val280, Cys281, His296[2]	-	-
14	Linalool	1.71	-2.38	17.890	Glu299	Asp338, Ser339, Lys340, Thr341	-	-	Lys342, Tyr337	-	-
15	Borneol	0.09	-4.02	1.120	Gly439, Ser441[2]	Gly282, Cys437, Gln438, Asp440	-	-	Val280, Cys281, His296[2]	-	-

[1] RMSD: Root mean square deviation (Angstrom = Å).
[2] The active amino acid residues of TMPRSS2 (His296, Asp345, Ser441).

Table 4. Molecular interactions between the monoterpenoids and CatB.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
1	Carvone	0.01	-5.10	0.182	His110[2], His111[2]	Gln23[2], Cys26[2], Gly27[2], Gly121, Glu122	-	-	Cys119, Cys29[2], His111[2], His199[2], Trp221[2]	-	-
2	Camphene	0.00	-4.67	0.379	-	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], Cys108, Gly121, Glu122	-	-	Cys119, His110[2], His111[2], Trp221[2]	-	-
3	Sabinene	0.08	-4.33	0.664	-	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], Cys108, Cys119, Gly121, Glu122	-	-	His110[2], His111[2], His199[2], Trp221[2]	-	-
4	β-Pinene	0.21	-4.56	0.456	-	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], His111[2], Gly121, Glu122	-	-	Cys108, Cys119, Cys29[2], His110[2], His199[2], Trp221[2]	-	-
5	Myrcene	1.27	-3.78	1.700	-	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], Gly121	-	-	Cys108, Cys119, Cys29[2], His110[2], His111[2], His199[2], Trp221[2]	-	-
6	α-Terpineol	1.00	-4.73	0.339	Gly24[2], Gly121	Gln23[2], Ser25, Cys26[2], Gly27[2], Cys108, Thr120, Glu122, Trp221[2]	-	-	Cys29[2], Cys119, His110[2], His111[2], His199[2]	-	-
7	Thymol	0.32	-4.20	0.836	Cys29[2], Gly198	Gln23[2], Cys26[2], Gly27[2], Gly197, Trp221[2]	-	His199[2]	Val176, Met196, Cys29[2], His199[2]	-	Cys29[2]

Table 4. Continued.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
11	Menthol	0.05	-4.71	0.355	Glu122	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], Asn72, Cys108, His110[2], Gly121, Trp221[2]	-	-	Cys29[2], Cys119, His199[2]	-	-
12	*p*-Cymene	13.569	-3.93	1.320	-	Gln23[2], Gly24[2], Ser25, Gly27[2], His111[2], Gly121, Glu122, Trp221[2]	-	-	Cys29[2], Cys108, Cys119, His110[2], His199[2]	Cys26[2]	Cys29[2]
13	Limonene	0.88	-4.26	0.754	-	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], Gly121, Glu122	-	-	Cys29[2], Cys119, Cys108, His110[2], His111[2], His199[2], Trp221[2]	-	-
14	Linalool	1.15	-3.91	1.370	His111[2]	Gln23[2], Gly24[2], Ser25, Cys26[2], Gly27[2], Glu109, Gly121	-	-	Cys29[2], Cys108, Cys119, His110[2], His111[2], His199[2], Trp221[2]	-	-
15	Borneol	0.14	-4.94	0.240	Gly24[2], Cys26[2], Gly121	Gln23[2], Ser25, Gly27[2], Thr120, Glu122	-	-	Cys119, His110[2], His111[2], Trp221[2]	-	-

[1] RMSD: Root mean square deviation (Angstrom = Å).
[2] The active amino acid residues of CatB (Gln23, Gly24, Cys26, Gly27, Ser28, Cys29, Trp30, Gly73, Gly74, His110, His111, His199, Trp221).

Table 5. Molecular interactions between the monoterpenoids and CatL.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
1	Carvone	0.08	-5.03	0.205	Cys25[2], Trp26[2]	Gly23[2], Asn66[2], Gly68[2], Val134, Asp162[2], Gly164	Trp26[2], Gly67[2]	-	Ala135[2], Ala214, Leu69[2], Met161[2], Met70[2]	-	-
2	Camphene	0.12	-3.95	1.280	-	Trp26[2], Asn66[2], Gly67[2], Gly68[2], Met161[2], Asp162[2], Gly164	-	-	Cys25[2], Leu69[2]	-	-
3	Sabinene	0.02	-4.29	0.712	-	Trp26[2], Gly67[2], Gly68[2], Met70[2], Val134, Ala135[2], Met161[2], Asp162[2], Gly164, Ala214	-	-	Cys25[2], Leu69[2]	-	-
4	β-Pinene	0.03	-3.91	1.360	-	Glu63, Asn66[2], Gly67[2],Gly68[2], Leu69[2], Met161[2], Asp162[2], Gly164	-	-	Cys25[2], Trp26[2]	-	-
5	Myrcene	0.26	-3.92	1.340	-	Gly23[2], Trp26[2], Asn66[2], Gly67[2], Gly68[2], Val134, Gly164	-	-	Cys25[2], Ala135[2], Ala214, Leu69[2], Met70[2], Cys25[2]	-	-
6	α-Terpineol	1.09	-4.96	0.233	Gly68[2]	Asn66[2], Gly67[2], Val134, Met161[2], Asp162[2], Gly164	-	-	Cys25[2], Leu69[2], Ala135[2], Ala214, Met70[2], Cys25[2], Trp26[2]	-	-
7	Thymol	0.95	-4.25	0.761	Cys25[2], Trp26[2]	Asn66[2], Gly67[2], Val134, Met161[2], Gly164	Trp26[2]	-	Ala135[2], Ala214, Leu69[2], Met70[2], Cys25[2], Leu69[2]	Gly68[2], Asp162[2]	-

Table 5. Continued.

No	Compound	RMSD (Å)[1]	Binding energy (kcal/mol)	Predicted IC_{50} value (mM)	Classical H-bond	Van der Waals	Nonclassical H-bond (C-H, Pi-Donor)	Hydrophobic interaction		Electrostatic	Miscellaneous (Lone pairs)
								p-p interaction	Mixed p/Alkyl		
9	Camphor	0.06	-3.93	1.320	Ser213	Asp71, Asp114, Lys117, Ala215, Ser216	Ala214	-	Ala214, Leu69[2], Met70[2]	-	-
10	β-Phellandrene	0.08	-4.80	0.302	-	Trp26[2], Asn66[2], Gly67[2], Gly68[2], Val134, Met161[2], Asp162[2], Gly164	-	-	Cys25[2], Leu69[2], Ala135[2], Ala214, Met70[2]	-	-
11	Menthol	0.06	-4.80	0.303	Trp26[2]	Cys25[2], Gly67[2], Gly68[2], Val134, Met161[2], Asp162[2], Gly164	-	-	Leu69[2], Ala135[2], Ala214, Met70[2]	-	-
12	*p*-Cymene	0.09	-4.36	0.632	-	Trp26[2], Asn66[2], Gly67[2], Val134, Met161[2], Gly164	-	-	Ala135[2], Ala214, Leu69[2], Met70[2], Cys25[2], Ala135[2]	Gly68[2], Asp162[2]	-
13	Limonene	0.05	-4.62	0.411	-	Asn66[2], Gly67[2], Gly68[2], Val134, Met161[2], Asp162[2], Gly164	-	-	Cys25[2], Leu69[2], Ala135[2], Ala214, Met70[2], Trp26[2]	-	-
14	Linalool	1.14	-3.80	1.650	Gly68[2]	Asn66[2], Gly67[2], Val134, Met161[2], Asp162[2], Gly164	-	-	Ala135[2], Ala214, Cys25[2], Leu69[2], Met70[2], Trp26[2]	-	-
15	Borneol	0.01	-4.02	1.130	Asp114	Asp71, Lys117, Ser213, Ala215, Ser216	-	-	Ala214, Leu69[2], Met70[2]	-	-

[1] RMSD: Root mean square deviation (Angstrom = Å).

[2]The active amino acid residues of CatL (Gln19, Gly20, Gln21, Cys22, Gly23, Ser24, Cys25, Trp26, Gly61, Asn66, Gly67, Gly68, Leu69, Met70, Ala135, Met161, Asp162, Trp189).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Arg403	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
Glu406	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0
Lys417	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Ile418	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Tyr453	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
Leu455[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Phe486[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gln493[1]	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Ser494[1]	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1
Tyr495	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gly496	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Phe497	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0
Asn501[1]	1	1	1	1	1	0	1	1	0	1	1	1	1	1	0
Tyr505[1]	1	1	1	1	1	1	1	3	1	1	1	2	1	1	1

Figure 4. Heatmap of monoterpenoid/RBD of the spike protein of 2019-nCoV interaction.
[1]Amino acid residues involved in binding to ACE2 in the receptor binding domain of 2019-nCoV (1: Carvone, 2: Camphene, 3: Sabinene, 4: β-Pinene, 5: Myrcene, 6: α-Terpineol, 7: Thymol, 8: Carvacrol, 9: Camphor, 10: β-Phellandrene, 11: Menthol, 12: *p*-Cymene, 13: Limonene, 14: Linalool, 15: Borneol).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Val278	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
His279	2	0	1	0	1	0	0	2	0	1	0	1	1	0	0
Val280	2	1	1	0	1	0	1	2	0	1	0	2	1	0	1
Cys281	1	0	1	0	1	0	0	1	0	1	0	1	1	0	1
Gly282	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Ala295	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
His296[1]	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1
Cys297	1	1	0	0	1	0	1	0	0	1	0	0	1	0	0
Val298	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Glu299	0	1	0	0	0	1	1	0	0	0	0	0	0	1	0
Pro301	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
Leu302	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
Trp308	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Tyr337	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
Asp338	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
Ser339	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
Lys340	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0
Thr341	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0
Lys342	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0
Asp345[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Trp384	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Thr393	1	0	1	0	1	0	0	1	0	1	0	1	1	0	0
Leu419	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Ser436	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Cys437	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1
Gln438	1	0	1	1	1	0	0	1	0	1	1	1	1	0	1
Gly439	1	0	1	0	1	0	0	1	0	1	1	1	1	0	1
Asp440	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Ser441[1]	1	0	1	1	1	0	0	1	0	1	1	1	1	0	1
Thr459	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Ser460	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
Trp461	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0
Gly462	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
Gly464	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0

Figure 5. Heatmap of monoterpenoid/TMPRSS2 interaction.
[1] The active amino acid residues of TMPRSS2 (1: Carvone, 2: Camphene, 3: Sabinene, 4: β-Pinene, 5: Myrcene, 6: α-Terpineol, 7: Thymol, 8: Carvacrol, 9: Camphor, 10: β-Phellandrene, 11: Menthol, 12: *p*-Cymene, 13: Limonene, 14: Linalool, 15: Borneol).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Gln23[1]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gly24[1]	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1
Ser25	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1
Cys26[1]	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1
Gly27[1]	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
Ser28[1]	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Cys29[1]	1	0	0	1	1	1	3	2	0	1	1	2	1	1	0
Trp30[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Asn72	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
Gly73[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gly74[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cys108	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0
Glu109	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
His110[1]	1	1	1	1	1	1	0	1	2	1	1	1	1	1	1
His111[1]	2	1	1	1	1	1	0	1	1	1	0	1	1	2	1
Cys119	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
Thr120	0	0	0	0	0	1	0	0	1	1	0	0	0	0	1
Gly121	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
Glu122	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1
Val176	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Met196	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Gly197	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Gly198	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
His199[1]	1	0	1	1	1	1	2	1	0	1	1	1	1	1	0
Trp221[1]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 6. Heatmap of monoterpenoid/CatB interaction.
[1] The active amino acid residues of CatB (1: Carvone, 2: Camphene, 3: Sabinene, 4: β-Pinene, 5: Myrcene, 6: α-Terpineol, 7: Thymol, 8: Carvacrol, 9: Camphor, 10: β-Phellandrene, 11: Menthol, 12: *p*-Cymene, 13: Limonene, 14: Linalool, 15: Borneol).

interactions with monoterpenoids (Figure 6). However, there was no interaction occurred between some active amino acids (Ser28, Trp30, Gly73, and Gly74) and phytochemicals. Intermolecular bonds were also formed between monoterpenoids and other amino acids of CatB (Ser25, Cys108, Cys119, Gly121, and Glu122). The monoterpenoids that interacted most with the amino acids of CatB were β-phellandrene, α-terpineol, *p*-cymene, and linalool.

The heatmap showing the interaction of monoterpenoids with CatL is given in Figure 7. Monoterpenoids formed several bonds with about 60% of the active amino acids of CatL (Cys25, Trp26, Asn66, Gly67, Gly68, Leu69, Met70, Ala135, Met161, and Asp162). However, the monoterpenoids could not form any interactions with other active amino acids (Gln19, Gly20, Gln21, Cys21, Gly23, Ser24, Gly61, and Trp189) or the interaction was rather weak. Monoterpenoids also formed bonds with Val134, Gly164 and Ala214, which are nonactive residues of CatL. The number of bonds established between CatL and thymol, carvone, α-terpineol and *p*-cymene was found to be higher than other monoterpenoids.

3.6. RBCI values of monoterpenoids

As known, in this study, the molecular interaction between 15 different monoterpenoids and 4 different protein targets was analyzed. Binding energy and IC_{50} data were obtained in the interaction analysis for each protein-ligand system. In order to detect 'hit' monoterpenes, a new analysis method called RBCI, the details of which are given in the 'Material and methods' section, has been applied. In order to calculate the RBCI values of phytochemicals, the binding energy and IC_{50} values of each monoterpene against all proteins were used together. It is not possible to use all data sets, each of which has different scientific meanings, in order to rank about the effectiveness of phytochemicals on target proteins. However, with RBCI analysis, the effectiveness of phytochemicals on all proteins can be determined by using different data sets at the same time. Thus, the most effective protein on all 4 proteins can be documented statistically (Figure 8).

RBCI analysis resulted in the superiority of carvone. Top ranked conformation of carvone in the RBD of the spike protein of 2019-nCoV, TMPRSS2, CatB, and CatL are presented in Figure 9. RBCI analysis also confirmed

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Asn18	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Gln19[1]	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Gly20[1]	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Gln21[1]	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Cys22[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gly23[1]	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Ser24[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cys25[1]	1	1	1	1	2	2	2	0	0	1	1	1	1	1	0
Trp26[1]	2	1	1	1	1	1	2	0	0	1	1	1	1	1	0
Gly61[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Glu63	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Asn66[1]	1	1	0	1	1	1	1	0	0	1	0	1	1	1	0
Gly67[1]	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0
Gly68[1]	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0
Leu69[1]	1	1	1	1	1	1	2	0	1	1	1	1	1	1	1
Met70[1]	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1
Asp71	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Asp114	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Lys117	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Val134	1	0	1	0	1	1	1	0	0	1	1	1	1	1	0
Ala135[1]	1	0	1	0	1	1	1	0	0	1	1	2	1	1	0
Met161[1]	1	1	1	1	0	1	1	0	0	1	1	1	1	1	0
Asp162[1]	1	1	1	1	0	1	1	0	0	1	1	1	1	1	0
Gly164	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0
Trp189[1]	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
Trp193	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Ser213	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Ala214	1	0	1	0	1	1	1	0	2	1	1	1	1	1	1
Ala215	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
Ser216	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1

Figure 7. Heatmap of monoterpenoid/CatL interaction.
[1] The active amino acid residues of CatL (1: Carvone, 2: Camphene, 3: Sabinene, 4: β-Pinene, 5: Myrcene, 6: α-Terpineol, 7: Thymol, 8: Carvacrol, 9: Camphor, 10: β-Phellandrene, 11: Menthol, 12: *p*-Cymene, 13: Limonene, 14: Linalool, 15: Borneol).

that *p*-cymene, myrcene, and linalool are the weakest components against protein targets in question.

3.7. Pharmacokinetic properties of monoterpenoids

Drug-likeness properties of docked monoterpene hydrocarbons against spike glycoprotein of 2019-nCoV, TMPRSS2, CatB and CatL and their ADMET profiles are given in Tables 6 and 7, respectively. Carvone, myrcene, α-terpineol, thymol, carvacrol, β-phellandrene, α-terpinene, limonene, linalool, and borneol were determined to be molecules in accordance with Lipinski's rule-of-five (Lipinski et al., 1997). However, the Moriguchi Log P (MLOGP) values of camphene, sabinene, β-pinene, α-thujene, and *p*-cymene have been found to be higher than 4.15. It has been determined that all molecules can cross the blood-brain barrier (BBB). None of the molecules were substrates of P-glycoprotein (P-gp). While camphene, β-pinene, thymol, carvacrol, *p*-cymene, and limonene inhibited CYP2C9, CYP1A2, and CYP2D6, other monoterpenoids had no inhibitory effect on cytochromes. None of the monoterpenoids exhibited toxic effect. The LD_{50} doses of the molecules in rats were in the range of 1.549–2.074 mol/kg. While thymol and carvacrol were monoterpenoids with the highest LD_{50} value, the LD_{50} values of sabinene and camphene were found to be lower than others.

4. Discussion

RMSD is a value commonly used in molecular docking analysis to measure the reproduction quality of a known binding pose of the molecule couples. This value is functional when the ligand molecule shows different poses in the binding site of the protein. Researchers suggest that RMSD values less than 2.0 Å are an important indicator of the quality of binding poses (Ramírez and Caballero, 2018). Accordingly, almost all of the analyzed monoterpenoids were found to have RMSD values below

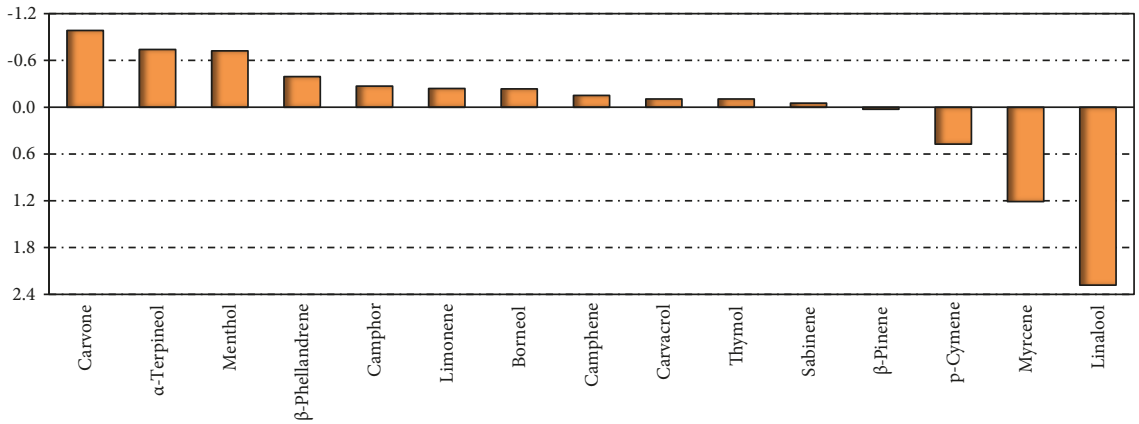

Figure 8. RBCI values of monoterpenoids.

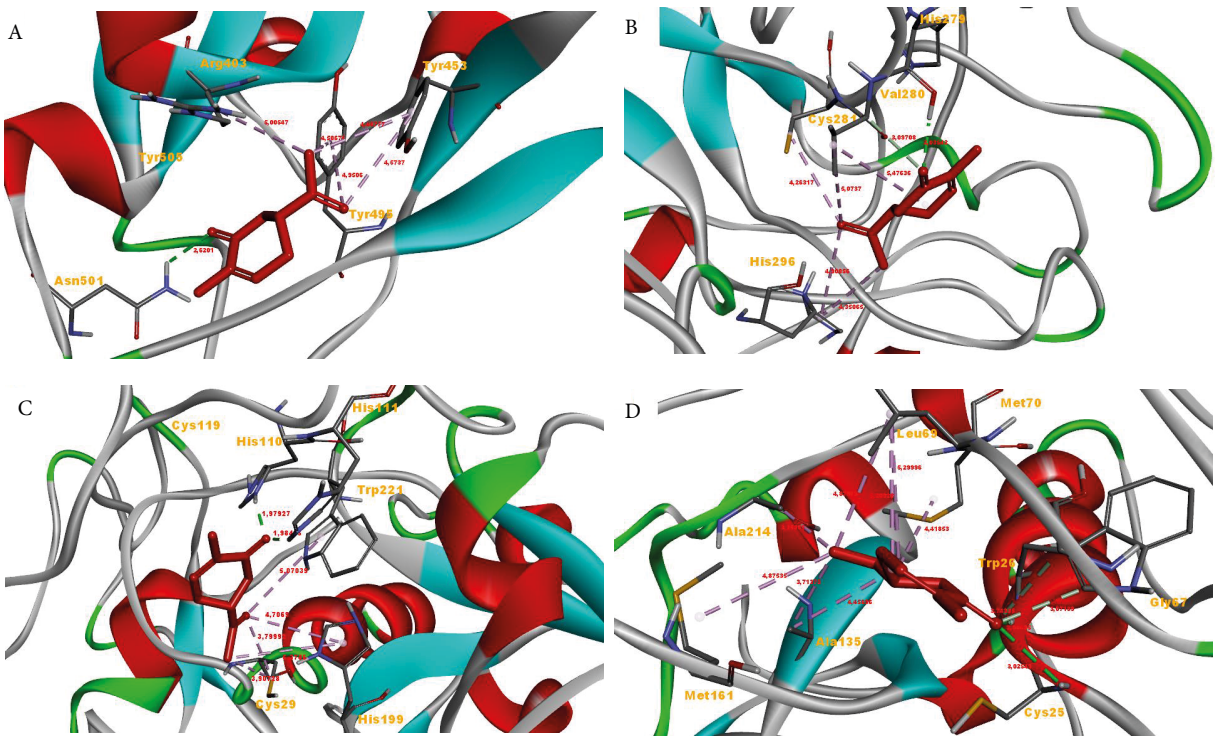

Figure 9. Top ranked conformation of carvone in the A) RBD of the spike protein of 2019-nCoV, B) TMPRSS2, C) CatB, and D) CatL.

2.0 Å against protein targets. It was found that RMSD values obtained only from menthol-TMPRSS2 and *p*-cymene-CatB interactions were above this critical limit and therefore could not exhibit sufficient binding activity on these molecules.

There was no correlation between the RMSD values of the molecules and binding energies and predicted IC_{50} values. However, a high correlation was found between binding energies obtained against all protein targets and IC_{50} values.

As a result of docking analysis, monoterpenoids interacted with more than half of the active amino acids of protein targets. Monoterpenoids mostly interacted with the active amino acids of CatB and TMPRSS2 (69% and 66% of active amino acids, respectively). Monoterpenoids bound with 55% of active amino acids of CatL, while this ratio has decreased to 50% in spike glycoprotein.

As a result of the RBCI analysis performed to detect the most effective phytochemicals against the all protein

Table 6. Drug-likeness properties of docked monoterpene hydrocarbons against spike glycoprotein of 2019-nCoV, TMPRSS2, CatB and CatL.

No	Compound	Number of rotatable bonds	TPSA[1]	Consensus Log P	Log *S* (ESOL[2])	Drug-likeness (Lipinski's rule-of-five)
1	Carvone	1	17.07 Å^2	2.44	-2.41	Yes (0 Violation)
2	Camphene	0	0.00 Å^2	3.43	-3.34	Yes (1 violation; MLOGP[3]> 4.15)
3	Sabinene	1	0.00 Å^2	3.25	-2.57	Yes (1 violation; MLOGP > 4.15)
4	β-Pinene	0	0.00 Å^2	3.42	-3.31	Yes (1 violation; MLOGP > 4.15)
5	Myrcene	4	0.00 Å^2	3.43	-3.05	Yes (0 Violation)
6	α-Terpineol	1	20.23 Å^2	2.58	-2.87	Yes (0 Violation)
7	Thymol	1	20.23 Å^2	2.80	-3.19	Yes (0 Violation)
8	Carvacrol	1	20.23 Å^2	2.82	-3.31	Yes (0 Violation)
9	α-Thujene	1	0.00 Å^2	3.15	-2.41	Yes (1 violation; MLOGP > 4.15)
10	β-Phellandrene	1	0.00 Å^2	3.07	-2.79	Yes (0 Violation)
11	α-Terpinene	1	0.00 Å^2	3.30	-3.30	Yes (0 Violation)
12	*p*-Cymene	1	0.00 Å^2	3.50	-3.63	Yes (1 violation; MLOGP > 4.15)
13	Limonene	1	0.00 Å^2	3.37	-3.50	Yes (0 Violation)
14	Linalool	4	20.23 Å^2	2.66	-2.40	Yes (0 Violation)
15	Borneol	0	20.23 Å^2	2.38	-2.51	Yes (0 Violation)

[1]TPSA: Topological polar surface area, [2] ESOL: Estimated aqueous solubility, [3] MLOGP: Moriguchi Log P.

Table 7. ADMET profiles of monoterpene hydrocarbons.

No	Compound	BBB[1] permeation	P-gp substrate[2]	CYP[3] inhibition	AMES toxicity	Hepatotoxicity	LD_{50} in rat (mol/kg)
1	Carvone	Yes	No	No	No	No	1.707
2	Camphene	Yes	No	Yes (CYP2C9)	No	No	1.554
3	Sabinene	Yes	No	No	No	No	1.549
4	β-Pinene	Yes	No	Yes (CYP2C9)	No	No	1.673
5	Myrcene	Yes	No	No	No	No	1.643
6	α-Terpineol	Yes	No	No	No	No	1.923
7	Thymol	Yes	No	Yes (CYP1A2)	No	Yes	2.074
8	Carvacrol	Yes	No	Yes (CYP1A2)	No	Yes	2.074
9	α-Thujene	Yes	No	No	No	No	1.589
10	β-Phellandrene	Yes	No	No	No	No	1.709
11	α-Terpinene	Yes	No	No	No	No	1.766
12	*p*-Cymene	Yes	No	Yes (CYP2D6)	No	No	1.827
13	Limonene	Yes	No	Yes (CYP2C9)	No	No	1.880
14	Linalool	Yes	No	No	No	No	1.704
15	Borneol	Yes	No	No	No	No	1.707

[1] BBB: Blood-brain barrier, [2] P-gp: P-glycoprotein substrate, [2] CYP: Cytochrome P.

targets, carvone was clearly superior to others. It was found to exhibit drug-like property according to the Lipinski's rule-of-five. As it is understood from the ADMET profile, carvone can effectively pass through BBB. P-gp is an important cell membrane component responsible for the removal of toxic molecules from the cell. This protein,

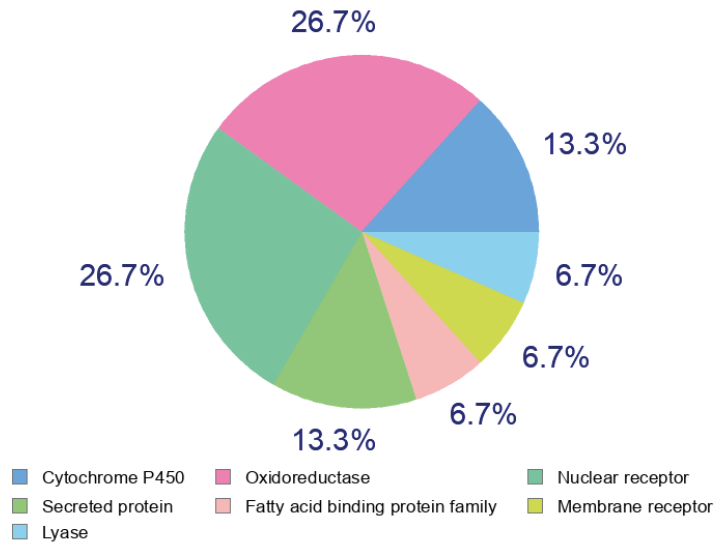

Figure 10. Target prediction of carvone.

which is found in many organisms, is thought to be one of the cell's defense mechanisms against harmful substances (Nguyen et al., 2020). As can be seen from Table 7, where ADMET profiles of monoterpenoids were given, carvone is not a substrate for P-gp. This finding shows that carvone is not perceived by the cell as a toxic molecule. In addition, the results of AMES toxicity and hepatotoxicity analysis support this finding. Carvone has also been found not to inhibit CYPs involved in the oxidation of xenobiotics and some endogenous harmful compounds (Rettie and Jones, 2005; Spector and Kim, 2015). This finding means that with carvone intake, metabolic activities based on oxidation can continue without disruption. LD_{50} dose of carvone in rats was also found to be better than those of sabinene, camphene, α-thujene, myrcene, β-pinene, and linalool. In Figure 10, the intracellular targets of carvone are included. According to this pie chart, the vast majority of intracellular targets of carvone are oxidoreductases and nuclear receptors. Oxidoreductases are enzymes that transfer the electron from the donor molecule to the acceptor. Usually they use NADP or NAD+ as the cofactors. In many organisms, transmembrane oxidoreductases are known to form the electron transfer chain (Rac and Fulgosi, 2020). Nuclear receptors are responsible for delivering the signal carried by some hormones (e.g., thyroid or some steroidal hormones) to the nucleus of the cell. Thus, these proteins regulate the expression of certain specific genes along with some other proteins (Evans, 1988). Although it is known that carvone targets these proteins, it is thought that detailed analyzes should be made to understand whether carvone has an inhibitory and/or activator effect on the activities of these proteins.

As a result of the literature research, the molecular interaction between spike glycoprotein of 2019-nCoV and carvone, camphene, α-terpineol, thymol, carvacrol, camphor, β-phellandrene, menthol, linalool, and borneol was analyzed and the binding energies of these molecules were found to be in the range of –3.4/–3.7 kcal (Smith and Smith, 2020). The data in question were found to be weaker than the data presented in the current study. However, no reports of the interaction of other monoterpenoids with the spike protein have been found. Additionally, no study was found documenting the interaction between neither carvone nor other monoterpenoids and TMPRSS2, CatB, and CatL. Therefore, the data presented here were considered to be the first records for the literature.

References

Andersen K, Rambaut A, Lipkin W, Holmes E, Garry R (2020). The proximal origin of SARS-CoV-2. Nature Medicine 26: 450-452.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J et al. (2009). BLAST+: architecture and applications. BMC Bioinformatics 10: 421.

Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H et al. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. The Lancet 395: 514-523.

Colovos C, Yeates TO (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Science 2: 1511-1519.

Corman V, Lienau J, Witzenrath M (2019). Coronaviruses as the cause of respiratory infections. Der Internist 60: 1136-1145.

Daina A, Michielin O, Zoete V (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports 7: 42717.

Daina A, Michielin O, Zoete V (2019). SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Research 47: W357-W364.

De Wit E, van Doremalen N, Falzarano D, Munster VJ (2016). SARS and MERS: recent insights into emerging coronaviruses. Nature Reviews Microbiology 14: 523-534.

Evans RM (1988). The steroid and thyroid hormone receptor superfamily. Science 240: 889-895.

Fehr AR, Channappanavar R, Perlman S (2017). Middle East respiratory syndrome: emergence of a pathogenic human coronavirus. Annual Review of Medicine 68: 387-399.

Glowacka I, Bertram S, Müller MA, Allen P, Soilleux E et al. (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. Journal of Virology 85: 4122-4134.

Greenspan PD, Clark KL, Tommasi RA, Cowen SD, McQuire LW et al. (2001). Identification of dipeptidyl nitriles as potent and selective inhibitors of cathepsin B through structure-based drug design. Journal of Medicinal Chemistry 44: 4524-4534.

Guan Y, Zheng B, He Y, Liu X, Zhuang Z et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 302: 276-278.

Guex N, Peitsch MC, Schwede T (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30: S162-173.

Hardegger LA, Kuhn B, Spinnler B, Anselm L, Ecabert R et al. (2011). Halogen bonding at the active sites of human cathepsin L and MEK1 kinase: efficient interactions in different environments. ChemMedChem 6 (11): 2048-2054.

Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 181 (2): 271–280.e8.

Huang C, Wang Y, Li X, Ren L, Zhao J et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet 395 (10223): 497-506.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993). PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography 26: 283-291.

Lau SK, Woo PC, Li KS, Huang Y, Tsoi H-W et al. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proceedings of the National Academy of Sciences 102 (39): 14040-14045.

Li F, Li W, Farzan M, Harrison SC (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309 (5742): 1864-1868.

Li W, Moore MJ, Vasilieva N, Sui J, Wong SK et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426: 450-454.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews 23 (1-3): 3-25.

Matsuyama S, Nagata N, Shirato K, Kawase M, Takeda M et al. (2010). Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. Journal of Virology 84: 12658-12664.

Meneguzzo F, Ciriminna R, Zabini F, Pagliaro M (2020). Accelerated production of hesperidin-rich citrus pectin from waste citrus peel for prevention and therapy of COVID-19. Preprints. doi: 10.20944/preprints202003.0386.v1

Morris GM, Lim-Wilby M (2008). Molecular docking. Methods in Molecular Biology 443: 365-382.

Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E (2020). A novel coronavirus emerging in China—key questions for impact assessment. New England Journal of Medicine 382: 692-694.

Nasab RR, Hassanzadeh F, Khodarahmi GA, Rostami M, Mirzaei M et al. (2017). Docking study, synthesis and antimicrobial evaluation of some novel 4-anilinoquinazoline derivatives. Research in Pharmaceutical Sciences 12 (5): 425-433.

Nguyen PH, Sigdel KP, Schaefer KG, Mensah GAKet al. (2020). The effects of anthracycline drugs on the conformational distribution of mouse P-glycoprotein explains their transport rate differences. Biochemical Pharmacology 174: 113813 doi: 10.1016/j.bcp.2020.113813

Pedretti A, Villa L, Vistoli G (2004). VEGA--an open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. Journal of Computer-Aided Molecular Design 18: 167-173.

Pires DE, Blundell TL, Ascher DB (2015). pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. Journal of Medicinal Chemistry 58: 4066-4072.

Prakash D, Suri S, Upadhyay G, Singh BN (2007). Total phenol, antioxidant and free radical scavenging activities of some medicinal plants. International Journal of Food Sciences and Nutrition 58: 18-28.

Rac A, Fulgosi H (2020). Qualitative and quantitative dataset of TROL protein interaction with C3 and C4 ferredoxin: NADP(+) oxidoreductases. Data in Brief 28: 105038.

Ramírez D, Caballero J (2018). Is it reliable to take the molecular docking top scoring position as the best solution without considering available structural data? Molecules 23: 1038.

Remmert M, Biegert A, Hauser A, Soding J (2011). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nature Methods 9: 173-175.

Rettie AE, Jones JP (2005). Clinical and toxicological relevance of CYP2C9: drug-drug interactions and pharmacogenetics. Annual Reviews of Pharmacology and Toxicology 45: 477-494.

Ricci CG, Netz PA (2009). Docking studies on DNA-ligand interactions: building and application of a protocol to identify the binding mode. Journal of Chemical Information and Modeling 49: 1925-1935.

Sampangi-Ramaiah MH, Vishwakarma R, Shaanker RU (2020). Molecular docking analysis of selected natural products from plants for inhibition of SARS-CoV-2 main protease. Current Science 118 (7): 1087-1092.

Sanner MF (1999). Python: a programming language for software integration and development. Journal of Molecular Graphics and Modelling 17: 57-61.

Sharma S (1996). Two group discriminant analysis. Applied multivariate techniques. New York, NY, USA: John Wiley & Sons, Inc, pp. 237-286.

Shulla A, Heald-Sargent T, Subramanya G, Zhao J, Perlman S et al. (2011). A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. Journal of Virology 85: 873-882.

Singh BN, Singh B, Singh R, Prakash D, Dhakarey R et al. (2009). Oxidative DNA damage protective activity, antioxidant and anti-quorum sensing potentials of Moringa oleifera. Food and Chemical Toxicology 47: 1109-1116.

Singh BN, Singh B, Singh R, Prakash D, Sarma B et al. (2009). Antioxidant and anti-quorum sensing activities of green pod of Acacia nilotica L. Food and Chemical Toxicology 47 (4): 778-786.

Singh BN, Singh B, Singh R, Prakash D, Singh D et al. (2009). Polyphenolics from various extracts/fractions of red onion (Allium cepa) peel with potent antioxidant and antimutagenic activities. Food and Chemical Toxicology 47 (6): 1161-1167.

Singh HB, Singh BN, Singh SP, Nautiyal CS (2010). Solid-state cultivation of Trichoderma harzianum NBRI-1055 for modulating natural antioxidants in soybean seed matrix. Bioresource Technology 101 (16): 6444-6453.

Smith M, Smith JC (2020). Repurposing therapeutics for covid-19: Supercomputer-based docking to the sars-cov-2 viral spike protein and viral spike protein-human ace2 interface. ChemRxiv. doi: 10.26434/chemrxiv.11871402

Spector AA, Kim H-Y (2015). Cytochrome P450 epoxygenase pathway of polyunsaturated fatty acid metabolism. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids 1851 (4): 356-365.

Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J et al. (2020). QMEANDisCo—distance constraints applied on model quality estimation. Bioinformatics 36: 1765-1771.

Tallei TE, Tumilaar SG, Niode NJ, Fatimawali F, Kepel BJ et al. (2020). Potential of plant bioactive compounds as SARS-CoV-2 main protease (Mpro) and spike (S) glycoprotein inhibitors: a molecular docking study. Preprints. doi: 10.20944/preprints202004.0102.v2

Thuy BTP, My TTA, Hai NTT, Hieu LT, Hoa TT et al. (2020). Investigation into SARS-CoV-2 resistance of compounds in garlic essential oil. ACS Omega 5 (14): 8312-8320.

Wang C, Horby PW, Hayden FG, Gao GF (2020). A novel coronavirus outbreak of global health concern. The Lancet 395: 470-473.

Wilson S, Greer B, Hooper J, Zijlstra A, Walker B et al. (2005). The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. Biochemical Journal 388: 967-972.

Xiang Z (2006). Advances in homology protein structure modeling. Current Protein and Peptide Science 7 (3): 217-227.

Zar J (1996). Biostatistical analysis. London, UK: Prentice–Hall International Inc.

Zhu N, Zhang D, Wang W, Li X, Yang B et al. (2020). China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine 382: 727-733.