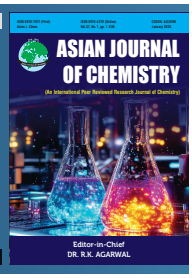
Asian Journal of Chemistry; Vol. 37, No. 7 (2025), 1761-1777

ASIAN JOURNAL OF CHEMISTRY

https://doi.org/10.14233/ajchem.2025.33728



Targeting SARS-CoV-2 Spike Protein with Regional Ethnomedicinal Plant-derived Phytocompounds: A Comprehensive Molecular Docking, ADMET and MD Simulation Study

AMBALIKA PHONGLO[1,2], DIKSHITA DOWERAH[3], SRUTISHREE SARMA[1], NAJIMA AHMED[1,4] and RAMESH CH. DEKA[1,*]

[1]CMML-Catalysis and Molecular Modelling Lab, Department of Chemical Sciences, Tezpur University, Napaam-784028, India
[2]Department of Chemistry, Anandaram Dhekial Phookan College, Haibargaon-782002, India
[3]Department of Chemistry, Arunachal University of Studies, Namsai-792103, India
[4]Center for Multidisciplinary Research, Tezpur University, Napaam-784028, India

*Corresponding author: E-mail: ramesh@tezu.ernet.in

Received: 24 March 2025; *Accepted*: 22 June 2025; *Published online*: 30 June 2025; AJC-22053

The achievement of efficient vaccine development for the SARS-CoV-2 virus marks a significant milestone, ultimately controlling the COVID-19 pandemic. However, the studies to find small molecules with potential SARS-CoV-2 inhibitory properties remained significant as there are not many treatment alternatives available. In this study, we studied biologically active phytocompounds present in regional ethnomedicinal plants like *Lindera neesiana*, *Litsea cubeba* and *Zanthoxylum armatum* DC using various computational techniques such as molecular docking, drug-likeness, ADMET study, MD simulation, *etc*. The SARS-CoV-2 spike protein's RBD (PDB ID: 6m0j) has been docked with 107 phytocompounds. Top 5 phytocompounds with the best docking scores were further screened for their drug-likeness and ADMET properties. Following that, 100 ns MD simulations including PCA analysis were performed, providing additional evidence of the stability of the studied protein-ligand complexes. Binding free energies were also evaluated using the MM-GBSA method, which revealed spathulenol to exhibit the most favourable energy signifying its strong binding affinity and stability within the RBD of the spike protein. This *in-silico* study identifies the specific phytocompounds present in *L. neesiana, L. cubeba* and *Z. armatum* DC as a potential inhibitor of SARS-CoV-2 spike protein.

Keywords: SARS-CoV-2, ADMET, Docking studies, MD simulation, PCA analysis, MM-GBSA.

INTRODUCTION

In order to tackle the COVID-19 pandemic, researchers across the world have ventured into designing and developing therapeutics against the SARS-CoV-2 virus [1]. Earlier reported zoonotic viral infections like SARS-CoV-1 (2003) and Middle East respiratory syndrome *i.e.* MERS (2012), though highly infectious and deadlier, these viruses were successfully confined preventing a worldwide pandemic [2-4]. In comparison to this, the recently emerged SARS-CoV-2 virus was extremely contagious and the news of the identification of new SARS-CoV-2 strains and cases continues to be reported now and then. Recent studies also suggest that the recently emerged SARS-CoV-2 strains have higher transmission rates as well as greater resistance against available vaccines. For instance, the omicron variant has been found to evade the antibody neutralization induced by Pfizer BNT162b2 mRNA vaccination [5-7]. Hence, to treat patients at the initial phase of infection and prevent the virus from further spreading, we still need to find novel small compounds that are safe to employ as oral antivirals [8]. Thus, the demand for drug compounds that are potent against SARS-CoV-2 still remains critical.

The structural and non-structural protein targets of SARS-CoV-2 that are crucial for the virus's survival have been long identified.1 SARS-CoV-2 has four structural proteins, namely, (i) spike (S), (ii) enveloped (E), (iii) membrane (M), and (iv) nucleocapsid (N) proteins [9]. Suppression of viral entry, assembly and replication have been the focus and among them, the spike protein is a primary drug target since it is important for host cell recognition and viral entry [8]. These spike proteins initiate the viral infection by binding to the host cell receptor angiotensin converting enzyme 2 (ACE2) [10-12]. The ACE2 receptor is present in lungs and the virus enters the lung cells by endocytosis [9,13,14]. This signifies that the compounds that can inhibit the interactions of S-protein/ACE2 would stop viral entry and thus prevent the subsequent infection [12].

The residues of SARS-CoV and SARS-CoV-2 spike proteins involved in making interaction with ACE2 at the binding site have undergone multiple mutations, which may in part explain the enhanced binding strength of SARS-CoV-2 with ACE2 [15]. However, Walls *et al.* [16] suggested that the feasibility of the viral membrane-fusion process is influenced by the conformational flexibility of the receptor binding domain (RBD) of SARS-CoV-2 S-protein and this would result in a stronger impact on the increased infectivity of SARS-CoV-2. Nevertheless, halting the initial viral binding to the ACE2 receptor still remains a valid mechanism for developing potential anti-SARS-CoV-2 therapeutics.

Furthermore, several spatial-temporal analyses suggest varied geographical conditions as one of the variables influencing the variations in the impacts of SARS-CoV-2 virus globally [17,18]. Similar spatial and spatial-temporal analyses also revealed that when about India, there were less infectivity as well as mortality rates in the Northeastern region and a few other states [19,20]. This could be due to different variables like the less dense population in the region, varying weather conditions or food habits and lifestyle patterns. Owing to its varied landscape, diversified ecology and favourable climatic conditions, the North-Eastern region has rich plant biodiversity making it distinct from other parts of the subcontinent [21,22]. The tribal ethnic communities living in this region not only consume these plants but also utilize them to treat different diseases. They have vast ethnomedicinal knowledge, the majority of which is yet to be reported to the scientific world [23].

Lindera neesiana (family: Lauraceae) [24], *Litsea cubeba* (family: Lauraceae) [25,26] and *Zanthoxylum armatum* DC (family: Rutaceae) [27,28] are enriched with various phytocompounds such as terpenoids, phenols, fatty acids, *etc*. In Northeast India, different parts of these plants are used for culinary purposes and have valuable medicinal uses. In many traditional medicines, *Litsea cubea*, locally known as mejankari and may chang (in China), is used for stomach ache and indigestion, as an anti-asthmatic, pain reliever, cold reliever, *etc.* [29]. Moreover, recent reports mention its antibacterial, antifungal, antioxidant, mosquito-repellent and anticancer effects [25,26]. Whereas, the fruits of *Lindera neesiana,* also known as Tejmui or Siltimur are chewed to treat tooth pain, diarrhea, gastric disorders and headaches [24]. While *Zanthoxylum armatum* DC, commonly known as Jabrang, is used to treat cough, bronchitis, rheumatism, carminative, anthelmintic, *etc.* [28,30]. Therefore, owing to its diverse pharmacological properties, we have studied the potential inhibitory properties of these three plants against SARS-CoV-2 spike protein using the *in silico* approach. To date, there is less report on the inhibition of SARS-CoV-2 spike protein by the bioactive compounds present in the chosen plants. We anticipate that the findings of this study would be helpful for more indepth research in the future and will also emphasize the geographical importance of the region.

EXPERIMENTAL

Ligand dataset preparation: In this computational study, a total of 107 phytocompounds from different parts of the *L. neesiana* [24], *L. cubeba* [25,26] and *Z. armatum* DC [27,28] plants were retrieved from reported literature and were used as ligand datasets. The standard 3D structures of ligands were obtained from PubChem database server (https://pubchem.ncbi.nlm.nih.gov/) in SDF format [31]. For docking calculation, the ligands were further prepared by converting them into PDBQT format using AutoDock Tools (ADT) [32].

Protein preparation: The three-dimensional coordinates of the target spike protein were retrieved from the Research Collaboratory for Structural Bioinformatics Protein Databank (RCSB PDB), https://www.rcsb.org/, having PDB ID 6m0j with a resolution of 2.45 Å [33,34]. The RBD domain of the S-protein was prepared by removing the bound ACE2 protein and all the water molecules from the protein structure. Furthermore, the target protein was processed by adding missing polar hydrogen atoms, followed by Gasteiger charges. Finally, after correcting all these aspects the structure file was converted to PDBQT format in AutoDock Tools (ADT) [32].

Grid generation and docking protocol: To obtain the best fit of the ligand molecules in the protein's binding site, molecular docking studies were performed using AutoDock4.2 software [35]. The grid box was generated by enclosing all the interface amino acid residues of the receptor binding domain (RBD) of the spike protein, important for binding to ACE2 receptor. The amino acid residues of the spike protein involved in interaction with ACE2 are Lys 417, Gly446, Tyr449, Tyr453, Leu455, Phe456, Ala475, Phe486, Asn487, Tyr489, Gln493, Gly496, Gln498, Thr500, Asn501, Gly502 and Tyr505 [34]. Accordingly, the grid centre was set to X = -36.649, Y = 30.63 and Z = 3.89 and dimensions of the grid box was set to 50 × 80 × 40 Å with grid spacing of 0.6 Å.

To obtain the best fitting conformation of the ligand compounds within the receptor's binding site, the ligands were kept flexible while treating the receptor as a rigid entity throughout the docking process. Using the auto grid tool, the grid maps of interaction energies between protein and various atom types present in the ligand were then pre-calculated [36]. Next, the Lamarckian Genetic Algorithm was used to evaluate the conformational states of the flexible ligand [37]. Following that docking calculations were carried out with default parameters [38]. The binding energy of the protein-ligand complex was calculated using the following equation which includes contributions from ΔG_{vdw} (dispersion/repulsion), ΔG_{hbond} (hydrogen bonding), ΔG_{elec} (electrostatic interaction), ΔG_{tor} (torsional constraints) and ΔG_{desolv} (desolvation effects) [39]:

$$\Delta G = \Delta G_{vdw} + \Delta G_{hbond} + \Delta G_{elec} + \Delta G_{tor} + \Delta G_{desolv} \quad (1)$$

The binding strength of the protein-ligand complex was predicted with the obtained binding energy values wherein the lowest binding energy (most negative) represents the highest binding affinity [39]. After docking, the molecular interactions between the docked protein-ligand complexes were analyzed by using the BIOVIA Discovery Studio visualizer [40].

Drug-likeness: Following docking results, drug-likeness properties were analyzed for the selected phytocompounds using the SwissADME server [41]. SwissADME server evaluates the physio-chemical properties of the compounds and

utilizes Lipinski's rule of 5 (RO5), Ghose's rule, Veber's rule, Egan's rule, *etc.* to predict their drug-likeness properties. As per Lipinski's RO5, a compound to exhibit drug-like characteristics should follow molecular weight, MW ≤ 500, number of hydrogen bond donors, HBD ≤ 5, number of hydrogen bond acceptor, HBA ≤ 10 and lipophilicity of the compound, $ClogP_{o/w} \leq 5$ or $MLogP_{o/w} \leq 4.15$ [42]. While Ghose rule states that an orally active compound should have a molecular weight in the range of 160-480 g/mol, $WlogP_{o/w}$ 0.4 to 5.6, total atomic count within 20 to 70 and molar refractivity between 40 to 130 [43]. These rules were further refined by applying Veber's and Egan's criteria, according to which a drug-like compound should satisfy rotatable bonds ≤ 10 and polar surface area (TPSA) ≤ 140 $Å^2$ (as per Veber's rule) [44] and $-1.0 \leq logP_{o/w} \leq 5.8$, TPSA ≤ 130 $Å^2$ (as per Egan's rule) [45].

ADMET study: ADMET is another important analysis of the drug discovery process which helps in examining the pharmacokinetic behaviour of the compounds. ADMET study was carried out for the selected compounds using the pkCSM server which evaluated various parameters associated with absorption (A), distribution (D), metabolism (M), excretion (E) and toxicity (T) of the concerned compounds [46]. Initially, we utilized the NCBI PubChem database to obtain the SMILES of the selected compounds, which were then used as an input file in the pkCSM server. Following that several parameters such as gastrointestinal (GI) absorption, solubility, skin permeability, blood–brain barrier (BBB) and central nervous system (CNS) permeability, CYP450 activity, Renal OCT2 binding, *etc.* were evaluated. In addition, oral rat acute toxicity and oral rat chronic toxicity levels, AMES and hepatoxicity, skin sensitization, hERG inhibition, *etc.* were analyzed to examine the toxicity of the phytocompounds.

Molecular dynamics (MD): To get more indepth information on the dynamics of the protein-ligand complex, a MD simulation was carried out for 100 ns using the AMBER18 program for the top 5 complexes that showed the highest binding affinity in docking calculations [47]. At first, quantum mechanical calculations of the selected ligands were performed at m06-2x/6-311G** level of theory in Gaussian09 software [48]. For simulation, ligand atoms were defined using GAFF force field, followed by preparation of set of ligand library and parameter files in the antechamber module of AMBER18 [49]. Similarly, using the leap module of Amber18, ff14SB force field was adopted for all the atoms in the receptor protein and the systems were neutralized with Cl^- ions [50]. The neutralized systems were then submerged in a TIP3P solvation octahedron box [51]. The energy minimization was carried out in two stages to remove any strains in the system. In both stages, each complex was optimized in 2500 steps of steepest descent followed by another 2500 steps of conjugate gradient minimization [52]. During the first stage of minimization, all atoms of the complex were weakly restrained to their initial coordinate whereas the second step of minimization was carried out without putting any restraints. The long-range electrostatic calculations were done using the Particle Mesh Ewald (PME) method and the non-bonded cut-off was set to 10 Å [53,54]. After minimization, the system was then heated gradually under NVT conditions for 100 ps from 0 K to 300 K [55]. Following that the systems were equilibrated under NPT conditions for 100 ps, maintaining a constant pressure of 1atm [56]. To constrain covalently bound hydrogen atoms, the SHAKE algorithm was applied [57]. Langevin thermostat was applied to maintain the temperature of the system at 300 K with a collision frequency of 1 ps^{-1}. Finally, the MD production run was carried out for 100 ns for each of the protein-ligand complexes in AMBER18 GPU implementation [58]. Analysis of MD trajectories was executed using the CPPTRAJ module implemented in AMBER 18 [59].

Principal component analysis (PCA) was performed to investigate the dynamical behaviour and conformational variations of the SARS-CoV-2 spike protein during the 100 ns simulation run. Several aspects of principal component analysis (PCA) such as eigenvectors, eigenvalues and their 2D projections concerning the first two principal components (PC1 *vs.* PC2 plot) and porcupine plots were extracted using the CPPTRAJ module of AMBER18. Free energy landscape (FEL), another important parameter of PCA to assess the conformational perturbations, was obtained by employing gmx sham module of the GROMACS software package [60]. The PCA results were plotted in XMgrace [61] and Origin (Origin Lab, USA) software and other dynamical motions were visualized in visual molecular dynamics (VMD) [62].

Binding free energy calculation: Following MD simulation, free energies of binding were calculated to validate the binding affinity of the studied phytocompounds towards the receptor binding domain (RBD) of the spike protein. In general, molecular mechanics generalized born surface area (MM-GBSA) and molecular mechanics poisson-boltzmann surface area (MM-PBSA) are two well-known methods to evaluate the binding free energy of protein-ligand complexes [63]. Hence, the MMPBSA.py module of AMBER18 was utilized to compute the binding free energies of phytocompounds from their MD simulation trajectories using the MM-GBSA method [64]. The binding free energy of the protein-ligand system is calculated following the equation:

$$\Delta G_{binding} = \Delta G_{complex} - \Delta G_{receptor} - \Delta G_{ligand} \quad (2)$$

here, the first term on the right-hand side represents the free energy of the bound protein-ligand adduct and the free energies of unbound protein and ligand are represented by the second and third terms, respectively [65]. Each free energy term on the right-hand side is estimated according to:

$$\Delta G = E_{MM} + G_{solvation} \quad (3)$$

where, E_{MM} takes contributions from several energy components like internal energy, van der Waals energy and electrostatic energy summed from molecular mechanics [66] *i.e.*

$$E_{MM} = \Delta G_{internal} + \Delta E_{vdw} + \Delta E_{electrostatic} \quad (4)$$

where $E_{internal}$ depicts strain energy due to bond stretching, angle bending and dihedral interactions. Whereas E_{vdw} and $E_{electrostatic}$ are the energies caused by non-bonded interactions depicted by Lennard-Jones and Coulomb equations, respectively [67].

Again, the free energy of solvation ($G_{solvation}$) is estimated as the sum of polar and non-polar contributions towards solvation energy [68].

RESULTS AND DISCUSSION

Molecular docking: To identify the compounds with potential viral entry inhibitory properties, the compounds reported in the three chosen plants *L. neesiana*, *L. cubeba* and *Z. armatum* were docked into the RBD of SARS-CoV-2 spike protein (PDB ID: 6M0J). In total, 107 compounds are listed based on binding affinity for the target protein (Table-1). Molecular docking revealed numerous phytocompounds to have appreciable binding affinity towards the active sites of the RBD domain of spike protein. However, for convenience, we shortlisted 10 docked protein-ligand complexes with the best binding energies for analysis using BIOVIA Discovery Studio. The phytochemical compounds showing the best binding affinity were α-caryophyllene, cadinene, spathuenol, germacrene A, allo-aromadendrene, exaltoin, camphor, copaene, β-caryophyllene and borneol. α-Caryophyllene was obtained as the best hit with a binding energy of -6.69 kcal/mol followed by cadinene and spathulenol with a binding energy of -6.63 kcal/mol and -6.47 kcal/mol respectively, as shown in Table-1.

TABLE-1
STUDIED PHYTOCOMPOUNDS AND THEIR PubChem ID, LIGAND SOURCE AND BINDING ENERGY AS OBTAINED FROM AutoDock4

S. No.	Ligand name	PubChem ID	Ligand source	Binding energy (kcal/mol)
1	α-Thujene	17868	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.80
2	Caryophyllene oxide	1742210	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-6.16
3	α-Pinene	6654	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-5.78
4	**α-Caryophyllene**	**23204**	***Litsea cubeba***	**-6.69**
5	Camphene	6616	*Lindera neesiana*; *Litsea cubeba*	-5.94
6	**β-Caryophyllene**	**5281515**	***Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC***	**-6.19**
7	β-Pinene	14896	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-5.87
8	*trans*-Carveol	94221	*Lindera neesiana*; *Zanthoxylum armatum DC*	-5.14
9	Myrcene	31253	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.42
10	*cis*-Carveol	330573	*Litsea cubeba*	-5.17
11	α-Fellandrene	7460	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.98
12	6-Methyl-5-hepten-2-one	9862	*Lindera neesiana*; *Litsea cubeba*	-3.90
13	β-Phellandrene	11142	*Litsea cubeba*; *Zanthoxylum armatum DC*	-5.05
14	*trans*-Linalool oxide	6432254	*Lindera neesiana*; *Zanthoxylum armatum DC*	-4.55
15	*cis*-Ocimene	5320250	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.44
16	*trans*-Ocimene	5281553	*Lindera neesiana*; *Litsea cubeba*	-4.53
17	*cis*-Sabinene hydrate	62367	*Lindera neesiana*; *Zanthoxylum armatum DC*	-5.02
18	Eucalyptol	2758	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-5.16
19	Palmitic acid	985	*Litsea cubeba*; *Zanthoxylum armatum DC*	-3.82
20	α-Terpinene	7462	*Litsea cubeba*; *Zanthoxylum armatum DC*	-4.86
21	P-Cymene	7463	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.58
22	γ-Terpinene	7461	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.45
23	o-Cymene	10703	*Litsea cubeba*; *Zanthoxylum armatum DC*	-4.92
24	**Camphor**	**2537**	***Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC***	**-6.28**
25	β-Cymene	10812	*Zanthoxylum armatum DC*	-5.03
26	(S)-Citronellol	7793	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-3.97
27	D-Limonene	440917	*Litsea cubeba*; *Zanthoxylum armatum DC*	-4.70
28	(R)-Citronellol	101977	*Litsea cubeba*	-4.07
29	Piperitone	6987	*Litsea cubeba*; *Zanthoxylum armatum DC*	-5.10
30	E-Citral	638011	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.28
31	β-Bisabolene	10104370	*Lindera neesiana*; *Litsea cubeba*	-5.78
32	Z-Citral	643779	*Lindera neesiana*; *Litsea cubeba*	-4.23
33	Citronellal	7794	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-3.97
34	Linalool	6549	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.26
35	Nerol	643820	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.46
36	Geraniol	637566	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.36
37	Terpinen-4-ol	11230	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.82
38	β-Elemene	6918391	*Lindera neesiana*; *Litsea cubeba*	-5.95
39	γ-Elemene	6432312	*Litsea cubeba*	-6.10
40	α-Terpineol	17100	*Lindera neesiana*; *Litsea cubeba*; *Zanthoxylum armatum DC*	-4.97
41	(Z)-β-Terpineol	8748	*Litsea cubeba*; *Zanthoxylum armatum DC*	-5.04
42	Myristicin	4276	*Lindera neesiana*	-4.56
43	Elemicin	10248	*Lindera neesiana*	-4.41
44	**Spathulenol**	**92231**	***Lindera neesiana***	**-6.47**
45	α-Campholenal	1252759	*Lindera neesiana*	-4.88

46	Verbenene	6427476	*Lindera neesiana*	-5.96
47	2,6-Dimethyl-5-heptanal	61016	*Lindera neesiana*	-4.02
48	Methyl eugenol	7127	*Lindera neesiana*	-4.33
49	Myrtenal	61130	*Lindera neesiana*	-5.37
50	Geranyl formate	5282109	*Lindera neesiana*	-4.19
51	S-(–)-Verbenone	29025	*Lindera neesiana*	-5.82
52	(E)-Limonene oxide	449290	*Litsea cubeba*	-4.91
53	(Z)-Limonene oxide	6452061	*Litsea cubeba*	-4.93
54	Isopulegol	170833	*Litsea cubeba*	-4.85
55	Isopulegone	34645	*Litsea cubeba*	-4.76
56	*cis*-Piperitol	85567	*Litsea cubeba*	-5.00
57	Isogeraniol	5362876	*Litsea cubeba*	-4.35
58	Geranic acid	5275520	*Litsea cubeba*	-3.99
59	**Copaene**	**19725**	***Litsea cubeba***	**-6.27**
60	Geranyl acetate	1549026	*Litsea cubeba*	-4.47
61	β-Farnesene	5281517	*Litsea cubeba*	-5.10
62	(+)-4-Carene	21674939	*Litsea cubeba*	-5.15
63	3-Carene	26049	*Litsea cubeba*	-4.87
64	**Borneol**	**64685**	***Litsea cubeba***	**-6.11**
65	2-Methyl-6-hepten-1-ol	565262	*Litsea cubeba*	-3.66
66	**Germacrene A**	**9548705**	***Litsea cubeba***	**-6.43**
67	**Cadinene**	**3032853**	***Litsea cubeba***	**-6.63**
68	Terpinolene	11463	*Litsea cubeba*	-4.80
69	Humulene epoxide II	10704181	*Litsea cubeba*	-5.94
70	Selina-6-en-4-ol	527220	*Litsea cubeba*	-6.08
71	α-Cadinol	10398656	*Litsea cubeba*	-6.60
72	Thujanol*	6432367	*Litsea cubeba*	-5.30
73	Bornyl acetate*	6448	*Litsea cubeba*	-5.54
74	Eugenol	3314	*Litsea cubeba*	-4.55
75	Germacrene D	6436582	*Litsea cubeba*	-5.95
76	Verbenol	61126	*Litsea cubeba*	-5.08
77	*n-trans*-Nerolidol	5284507	*Litsea cubeba*	-5.21
78	Ethyl myristate	31283	*Litsea cubeba*	-3.39
79	2,3-Dehydro-1,8-cineole	523035	*Litsea cubeba*	-5.73
80	**Allo-Aromadendrene**	**42608158**	***Zanthoxylum armatum DC***	**-6.41**
81	(E)-Nerolidol	5284507	*Zanthoxylum armatum DC*	-4.95
82	*trans*-Pipertiol	85568	*Zanthoxylum armatum DC*	-4.96
83	Isofenchol	91749788	*Zanthoxylum armatum DC*	-5.73
84	E-Methyl cinnamate	637520	*Zanthoxylum armatum DC*	-4.35
85	Z-Methyl cinnamate	6428458	*Zanthoxylum armatum DC*	-5.01
86	Cuminaldehyde	326	*Zanthoxylum armatum DC*	-4.55
87	Methyl palmitate	8181	*Zanthoxylum armatum DC*	-3.65
88	*cis*-9-Hexa-decenoic acid	445638	*Zanthoxylum armatum DC*	-3.18
89	*cis*-10-Octadecenoic acid	5282759	*Zanthoxylum armatum DC*	-3.31
90	Theaspirane	61953	*Zanthoxylum armatum DC*	-5.91
91	Sabinene	18818	*Zanthoxylum armatum DC*	-5.08
92	α-Asarone	636822	*Zanthoxylum armatum DC*	-4.75
93	β-Asarone	5281758	*Zanthoxylum armatum DC*	-4.77
94	Cuminol	325	*Zanthoxylum armatum DC*	-4.52
95	Citronellic acid	10402	*Zanthoxylum armatum DC*	-3.80
96	Artimisia alcohol	12308602	*Zanthoxylum armatum DC*	-4.78
97	**Exaltoin**	**543400**	***Zanthoxylum armatum DC***	**-6.29**
98	1-Methoxy-6,6-dimethyl-1-cyclohexene	580527	*Zanthoxylum armatum DC*	-4.64
99	*cis-p*-Menth-2-en-1-ol	13918681	*Zanthoxylum armatum DC*	-5.15
100	Menthoglycol Or PMD	19100	*Zanthoxylum armatum DC*	-4.90
101	α-Thujone	12304613	*Zanthoxylum armatum DC*	-5.51
102	α-Fenchol	439711	*Zanthoxylum armatum DC*	-5.81
103	Carvone	7439	*Zanthoxylum armatum DC*	-5.12
104	Tagetonol	522417	*Zanthoxylum armatum DC*	-3.95
105	Phellandral	89488	*Zanthoxylum armatum DC*	-4.93
106	2-Tridecanone	11622	*Zanthoxylum armatum DC*	-3.59
107	Undecan-2-one	8163	*Zanthoxylum armatum DC*	-3.80

Highlighted portion are the 10 best compounds that showed the highest binding affinity towards SARS-CoV-2 spike protein along with their binding energy, 2D structure, and ligand source

Molecular interactions of α-caryophyllene with the RBD of spike protein consisted of pi-sigma interactions with Phe497 and pi-alkyl interaction with Tyr449, Tyr453 and Tyr495 residues. Besides, it also makes van der Waals interaction with Ser494, Gly496, Arg403, Tyr505 and Asn501 (Fig. 1).

Cadinene made alkyl interaction with Arg403 and pi-alkyl contacts with Tyr449, Tyr495 and Tyr505 residues. Apart from these, Ser494, Gly496, Phe497, Tyr453 and Asn501 residues make van der Waals interactions with cadinene. Spathulenol is bound to the RBD of the spike protein by making direct hydrogen bonding with residues Ser494 and Gly496. Moreover, it also makes other hydrophobic interactions such as pi-alkyl with Phe497, Tyr495 and Tyr505 and alkyl with Arg403 residue. While residues Tyr453, Gln498 and Asn501 are involved in making non-bonded van der Waals interactions. The interactions of all the studied phytocompounds with the RBD of spike protein summarized with the number of interactions are provided in Table-2.

Germacrene A showed pi-alkyl interaction with Tyr449 and Tyr505 and other surrounding residues like Ser494, Tyr495, Gly496, Phe497, Arg403, Tyr453 and Asn501 participate in van der Waals interaction. In the case of allo-aromadenderene bound complex, the molecular interaction consisted of alkyl contacts with residue Arg403 and pi-alkyl interactions with residues Tyr453, Phe497, Tyr495 and Tyr505. Residues Ser494 Gly496 and Asn501 were involved in van der Waals interaction.

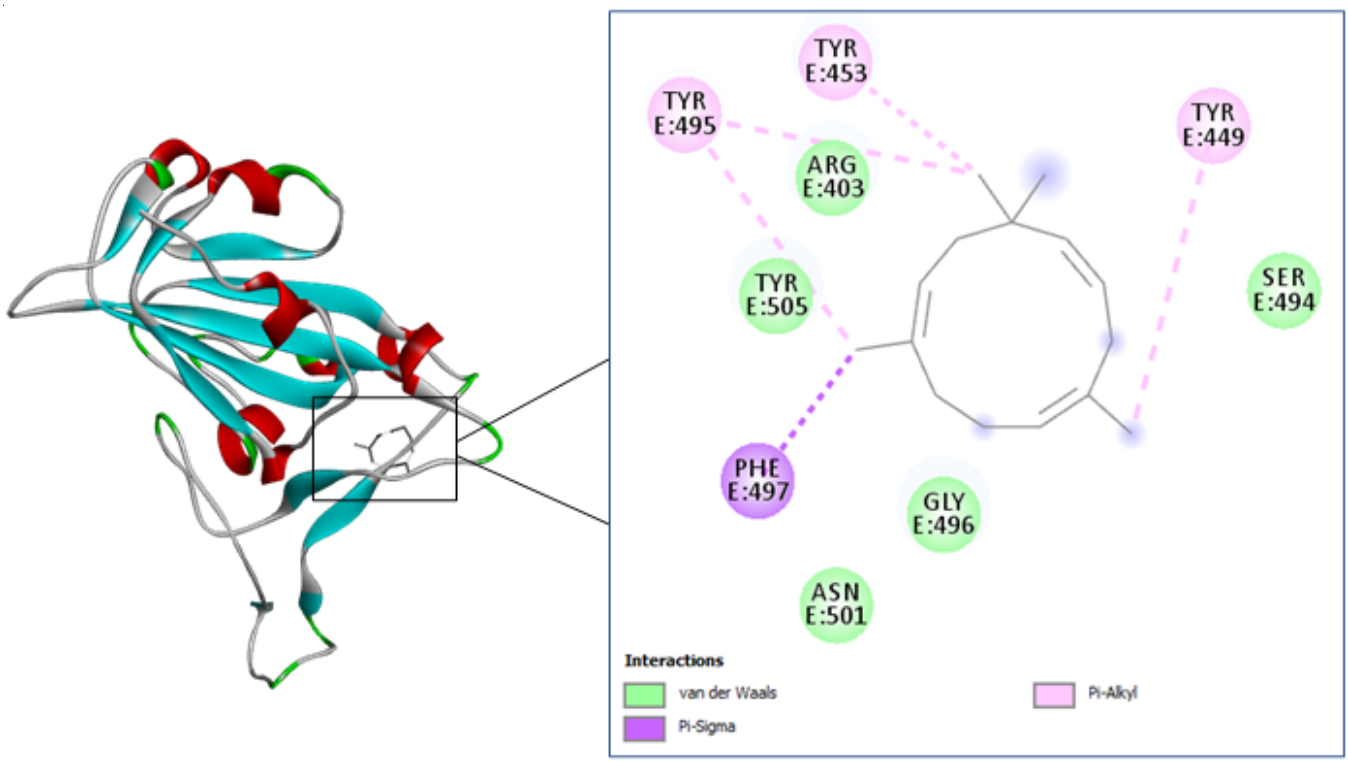


Fig. 1. α-Caryophyllene-spike protein complex and 2D image of their interaction

TABLE-2
INTERACTING RESIDUES OF SARS-CoV-2 SPIKE PROTEIN WITH THE PHYTOCOMPOUNDS SUMMARIZED WITH NUMBER OF INTERACTIONS

Ligand	H-bonding	Hydrophobic interaction (pi-sigma, pi-alkyl, alkyl, C-H bond)	van der Waals interactions
α-Caryophyllene	–	Tyr449(1), Tyr453(1), Tyr495(2), Phe497(1)	Ser494, Gly496, Arg403, Tyr505, Asn501
Cadinene	–	Arg403(1), Tyr449 (1), Tyr495(1), Tyr505(3)	Ser494, Gly496, Phe497, Tyr453, Asn501
Spathulenol	Ser494, Gly496	Arg403(1), Tyr495(2), Phe497(1), Tyr505(1)	Tyr453, Gln498, Asn501
Germacrene A	–	Tyr449 (1), Tyr505(2)	Ser494, Tyr495, Gly496, Phe497, Arg403, Tyr453, Asn501
Allo-Aromadendrene	–	Arg403(1), Tyr453(2), Tyr495(2), Phe497(1), Tyr505(1)	Ser494 Gly496, Asn501
Exaltoin	Gln493	–	Tyr449, Ser494, Tyr495, Gly496, Phe497, Gln498, Arg403, Tyr453, Asn501, Tyr505
Camphor	–	Tyr453(1), Arg457(1), Lys458(1), Pro491(1)	Phe456, Ile472, Gln474, Glu471, Ser469, Asp467, Arg454
Copaene	–	Tyr453(1), Tyr495(1), Tyr505(2)	Ser494, Gly496, Phe497, Gln498, Arg403, Tyr449, Asn501
β-Caryophyllene	–	Arg403(2), Tyr453(1), Tyr495(3), Phe497(2), Tyr505(1)	Ser494, Gly496, Tyr449, Asn501
Borneol	Glu471	Arg457(2), Lys458(1), Tyr473(2), Pro491(2)	Gln474, Asp467, Ile472, Ser469, Arg454, Phe456

Exaltoin interacted by making direct hydrogen bonding with Gln493 and non-bonded van der Waals interaction with Tyr449, Ser494, Tyr495, Gly496, Phe497, Gln498, Arg403, Tyr453, Asn501 and Tyr505 residues.

Camphor showed alkyl interaction with Lys458, Arg457 and Pro491 and pi-alkyl interaction with Tyr473. Other than these, it also showed van der Waals interaction with the surrounded residues such as Phe456, Ile472, Gln474, Glu471, Ser469, Asp467 and arg454. Copaene interacted by making pi-alkyl linkages with Tyr495, Tyr453 and Tyr505 residues. Furthermore, many surrounding residues Ser494, Gly496, Phe497, Gln498, Arg403, Tyr449 and Asn501 are involved in forming non-bonded van der Waals interaction. In case of β-caryophyllene bound complex, molecular interactions consist of alkyl interaction with residue Arg403 and pi-alkyl interaction with Tyr495, Phe497, Tyr453 and Tyr505 residues. While Ser494, Gly496, Tyr449 and Asn501 are involved in making non-bonded van der Waals contacts. Lastly, the caryophyllene oxide compound is bound to the RBD of spike protein by forming a C-H bond with residue Arg403 and pi-alkyl bonding with Tyr449 and Tyr505 residues. In addition to these, Tyr453, Ser494, Tyr495, Gly496, Phe497 and Asn501 residues were involved in making van der Waals interactions. Fig. 2 displays the 2D structures of interaction between all these phytochemical compounds and the RBD of the spike protein.

Till now, many studies have been reported regarding the structure of the SARS-CoV-2 spike protein in complex with the human ACE2 receptor. According to X-ray crystallography techniques, the contact residues of the RBD of the spike protein interacting with ACE2 are Lys 417, Gly446, Tyr449, Tyr453, Leu455, Phe456, Ala475, Phe486, Asn487, Tyr489, Gln493, Gly496, Gln498, Thr500, Asn501, Gly502 and Tyr505.74 After detailed analysis of the molecular interaction of the best binding phytochemical compounds from the chosen plants *L. neesiana*, *L. cubeba* and *Z. armatum*, we observed that most of the interactions are within the experimentally reported active site amino acid residues.

Drug likeness properties: After docking, the best 5 phytocompounds showing a higher binding affinity for RBD of SARS-CoV-2 spike protein were shortlisted to examine their drug-likeness properties. Drug-likeness is a concept in drug designing used for the pre-assessment of the physio-chemical properties of the studied compounds; as certain physio-chemical properties largely influence the oral availability of a bioactive compound. Here, Lipinski's rule of 5 (RO5) also known as Pfizer's rules, Ghose, Veber as well as Egan's rules were applied to evaluate the drug-likeness properties of the shortlisted ligands. However, it is not mandatory that a compound should obey all the rules to be a viable drug compound. As oral bioavailability of a drug compound does not affect the pharmacological potencies of drug [69]. Table-3 enlists the drug-likeness properties of the studied phytocompounds. Spathulenol was found to obey all the criteria of Lipinski's rule of 5 (RO5), implying that it has a molecular weight below 500 mg/mol, lipophilicity $ClogP_{o/w}$ less than 5 or $MLogP_{o/w}$ less than 4.15, number of H-bond donor and H-bond acceptor less than 5 and 10, respectively [42]. On the contrary, α-caryophyllene, cadinene, germacrene A and allo-aromadendrene violated one of Lipinski's rules as they exhibit $MLogP_{o/w} \leq 4.15$. However, their consensus $logP_{o/w}$ value was within the desired limits ($ClogP_{o/w} \leq 5$). Moreover, none of the studied compounds showed a violation of Ghose and Veber's rules of drug-likeness. Thus, they satisfy the ideal range of molecular weight (160-480 g/mol), $WlogP_{o/w}$ (0.4 to 5.6), molar refractivity (40 to 130) and total atomic count (20 to 70), as per Ghose's rule [43] and rotatable bonds (≤ 10), polar surface area (TPSA) (≤ 140 $Å^2$) as per Veber's rule [44]. They also meet the criteria of Egan's rule, which means they exhibit $logP_{o/w}$ within -1.0 to 5.8 and TPSA less than or equal to 130 $Å^2$ [45].

ADMET study: To understand the biological activity and pharmacokinetic behaviour of the compounds, ADMET property analysis is crucial in the process of drug discovery. Various absorption parameters like water solubility, Caco2 permeability, GI absorption *etc.* were evaluated for the five shortlisted compounds (Table-4). Their evaluated logS values were in the range of -6.188 to -3.866, signifying moderate to high solubility. All of the 5 compounds exhibit high Caco2 permeability (> 0.90), indicating their high oral absorption. A high gastro-

TABLE-3
PHYSICO-CHEMICAL PROPERTIES AND DRUG-LIKENESS OF THE FIVE STUDIED PHYTOCOMPOUNDS AS OBTAINED IN THE SwissADME SERVER

Physicochemical properties and drug-likeness	Compound name				
	α-Carphyllene	Cadinene	Spathulenol	Germacrene A	Allo-aromadendrene
Molecular weight (g/mol)	204.35	206.37	220.35	204.35	204.35
Number of heavy atoms	15	15	16	15	15
Number of rotatable bonds	0	1	0	1	0
Number of H-bond acceptor	0	0	1	0	0
Number of H-bond donor	0	0	1	0	0
Molar refractivity	70.42	69.52	68.34	70.68	67.14
TPSA ($Å^2$)	0.00	0.00	20.23	0.00	0.00
$MLogP_{o/w}$	4.53	5.56	3.67	4.53	5.65
Concensus $logP_{o/w}$	4.26	4.50	3.26	4.56	4.34
Lipinski; violation	Yes; 1	Yes; 1	Yes; 0	Yes; 1	Yes; 1
Ghose; violation	Yes; 0	Yes; 0	Yes; 0	Yes; 0	Yes; 0
Veber; violation	Yes; 0	Yes; 0	Yes; 0	Yes; 0	Yes; 0
Egan; violation	Yes; 0	Yes; 0	Yes; 0	Yes; 0	Yes; 0
Bioavailability score	0.55	0.55	0.55	0.55	0.55

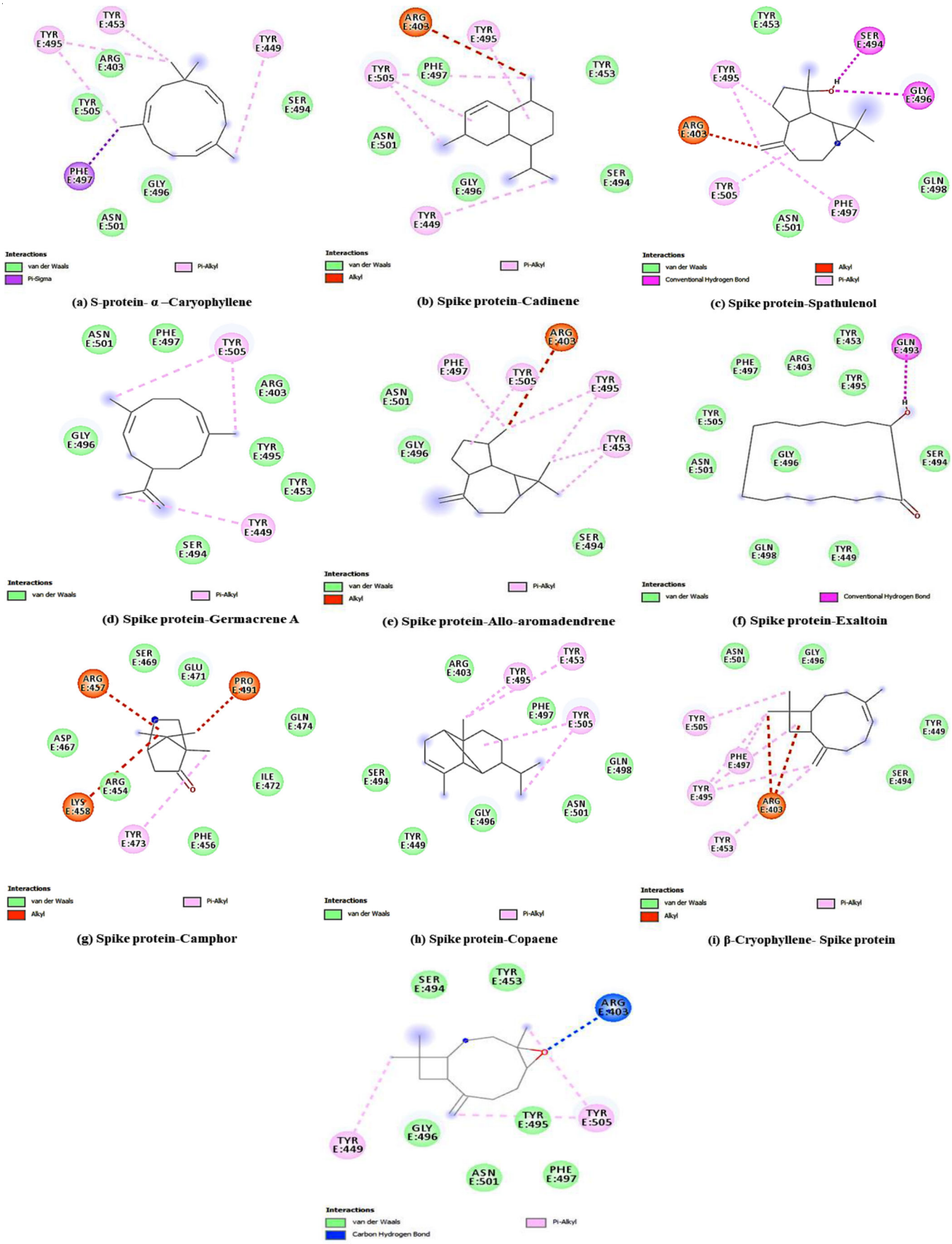


Fig. 2. 2D interaction mode between the highest binding compounds and RBD of SARS-CoV-2 spike protein

TABLE-4
ADME PROPERTIES AND TOXICITY PARAMETERS OF THE STUDIED PHYTOCOMPOUNDS AS EVALUATED BY pkCSM WEB SERVER

Property	Model name	α-Carphyllene	Cadinene	Spathulenol	Germacrene A	Allo-aromadendrene
Absorption	Water solubility (log mol/L)	-5.191	-6.188	-3.866	-5.412	-5.764
	Caco2 permeability (log Papp in 10^{-6} cm/s)	1.421	1.421	1.388	1.424	1.395
	Intestinal absorption (human) (% Absorbed)	94.682	96.179	93.235	94.928	95.302
	Skin permeability (log Kp)	-1.739	-1.552	-2.141	-1.587	-1.828
	P-Glycoprotein substrate	Yes	No	No	No	No
Distribution	BBB permeability (log BB)	0.663	0.806	0.6	0.695	0.822
	CNS permeability (log PS)	-2.555	-1.631	-2.447	-2.445	-1.769
Metabolism	CYP2D6 substrate	No	No	No	No	No
	CYP3A4 substrate	No	Yes	Yes	No	Yes
	CYP1A2 inhibitor	No	No	No	No	No
	CYP2C19 inhibitor	No	No	Yes	No	No
	CYP2C9 inhibitor	No	No	No	No	No
	CYP2D6 inhibitor	No	No	No	No	No
	CYP3A4 inhibitor	No	No	No	No	No
Excretion	Total clearance (log ml/min/kg)	1.282	1.099	0.895	1.443	0.926
	Renal OCT2 substrate	No	No	No	No	No
Toxicity	AMES toxicity	No	No	No	No	No
	Max. tolerated dose (human) (log mg/kg/day)	0.551	-0.028	0.077	0.535	-0.146
	hERG I inhibitor	No	No	No	No	No
	hERG II inhibitor	No	No	No	No	No
	Oral rat acute toxicity (LD_{50}) (mol/kg)	1.766	1.492	1.687	1.759	1.526
	Oral rat chronic toxicity (LOAEL) (log mg/kg bw/day)	1.336	1.456	1.39	1.347	1.332
	Hepatotoxicity	No	No	No	No	No
	Skin sensitization	Yes	Yes	No	No	Yes

intestinal (GI) absorption is indicated by GI absorption ≥ 30% and all of the studied compounds showed GI absorption value above 90% [70]. However, they possess low skin permeability as indicated by their logKp value (logKp > -2.5 indicates low skin permeability). Except for α-caryophyllene, none of them were substrates of P-gp (P-glycoprotein), which plays a major role in releasing toxins out of the cells [70].

When the distribution is concerned, they have log BB (blood-brain barrier) values above the ideal range (log BB > 0.3), which means that all of the compounds could cross the BBB. Compounds with log PS value greater than -2 are considered to have high CNS (central nervous system) permeability [70]. Except for cadinene and allo-aromadendrene, the remaining three phytocompounds have log PS values below -2. Thus, they are not able to readily penetrate the central nervous system.

Metabolic properties of drug compounds can be largely linked to their cytochrome P450 activity. Several pharmacological molecules are inactivated by cytochrome P450, whilst others are activated by it [71]. Thus, many unfavourable drug-drug interactions that could lead to drug toxicities and therapeutic failures could be prevented by recognizing whether the concerned drugs are CYP substrates or inhibitors [72]. It was observed that none of the phytocompounds were substrates of CYP2D6 enzyme. Whereas except for α-caryophyllene and germacrene A, the remaining three phytocompounds were CYP3A4 substrates. Also, as per metabolism data, all of the compounds were found to be non-inhibitors of CYP1A2, CYP2C9, CYP2D6, CYP3A4 and CYP2C19 enzymes, except for spathulenol which is a CYP2C19 inhibitor.

Excretion is a measure of drug clearance, which occurs mainly as a combination of renal clearance (excretion *via* kidneys) and hepatic clearance (metabolism in liver and biliary clearance). The evaluated total clearance values for the studied compounds were obtained in the range of 0.895 to 1.443, as given in Table-4. Moreover, none of the phytocompounds were substrates of renal OCT2 (organic cationic transporter 2), which plays a major role in the deposition and renal clearance of drugs. Renal OCT2 substrates are also likely to show unfavourable interactions when co-administered with OCT2 inhibitors.

When the toxicity factor is concerned, none of the studied phytocompounds exhibited hepatoxicity and AMES toxicity. They also did not show toxicity to the skin except for α-caryophyllene, cadinene and allo-aromadendrene, which could induce skin sensitization. Inhibition of hERG, which encodes potassium ion channels could lead to the development of cardiotoxicity [73]. All of the phytocompounds did not show inhibition of hERG I and hERG II, signifying no associated cardiotoxicity. In addition, the evaluated values of LD_{50}, LOAEL and maximum tolerated dose as predicted by the pkCSM server are also shown in Table-4. Therefore, from the analysis of drug-likeness and ADMET properties of the five shortlisted phytocompounds, it can be summarized that they exhibit admirable physio-chemical and pharmacokinetic properties.

Molecular dynamics simulation: Following molecular docking study, the apo-form of SARS-CoV-2 spike protein and the best five protein-ligand complexes *viz.* spike-α-caryophyllene, spike-cadinene, spike-spathulenol, spike-germacrene A and spike-allo-aromadendrene were subjected to molecular

dynamics (MD) simulation for 100 ns time scale. MD simulation study was performed to get deeper insights into the consequences of ligand binding on the dynamical properties like overall stability (RMSD), conformational fluctuations (RMSF) and compactness (Rg) of the protein-ligand complexes.

RMSD: RMSD analysis is performed to determine the structural stability of the protein backbone in the presence of ligands with respect to initial protein backbone conformation. During the course of the simulation, it computes the conformational perturbations, which occur in the protein backbone. Smaller deviation values suggest that the protein backbone exhibited steady behaviour, whereas larger RMSD value indicates the instability of the protein-ligand system. The RMSD plot obtained from the simulation trajectories for the Apo-spike protein and its complexes with α-caryophyllene, cadinene, spathulenol, germacrene A and allo-aromadendrene is shown in Fig. 3. The calculated average RMSD values of Apo-spike protein and the complexes of spike protein with all the selected 5 ligands fall within 2.06-2.65 Å (Table-5). This suggests that the spike protein did not undergo major deviation after the ligand inclusion and the behaviour of the spike protein backbone was quite similar to that of the Apo protein. The least deviation was observed in the case of cadinene docked complex with an average RMSD value of 2.13 Å and the generated RMSD profile was observed to overlap with Apo-spike protein in most of the time frames of the 100 ns simulation period. The complexes of α-caryophyllene and germacrene A also did not show any prominent deviation and both of them got equilibrated after around 70 ns with an average RMSD value of 2.15 Å.

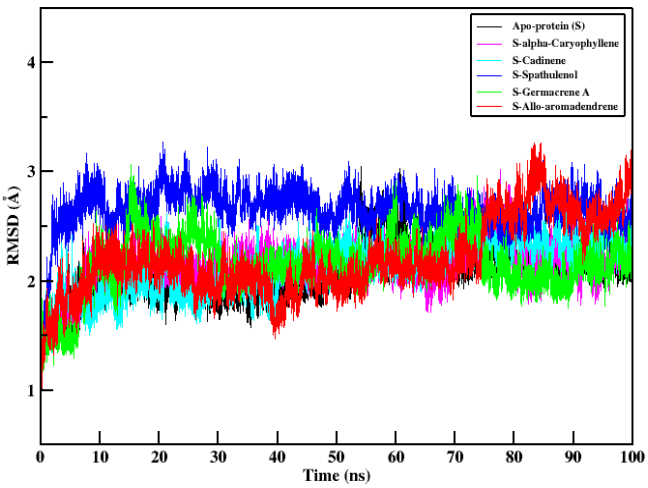


Fig. 3. RMSD plot of Apo form of SARS-CoV-2 spike protein and spike protein-ligand complexes for the 100 ns simulation time

In contrast, the allo-aromadendrene docked complex showed continual structural changes throughout the simulation timescale. The spike-spathulenol complex was observed with considerably higher deviations as compared to others with an average RMSD value of 2.65 Å. It showed an unusual rise in RMSD at the beginning up to 10 ns, however, after 10 ns it exhibited steady RMSD for the rest of the simulation period.

RMSD calculations were also carried out for the studied five phytocompounds and the obtained RMSD plot is shown in Fig. 4. The phytocompounds *e.g.* cadinene, spathulenol, germacrene A and allo-aromadendrene exhibited stable conformation and did not undergo any significant change in its orientation during the entire 100 ns simulation period. The calculated average RMSD values for cadinene, spathulenol, germacrene A and allo-aromadendrene were 0.69 Å, 0.20 Å, 0.26 Å, 0.69 Å and 0.29 Å, respectively. α-Caryophyllene was found to change its orientation at 10 ns and 65 ns, however, it got equilibrated after 70 ns, with an average RMSD value of 0.65 Å. Nonetheless, similar to other studied ligands like spathulenol, α-caryophyllene also remained positioned within the RBD of the SARS-CoV-2 spike protein.

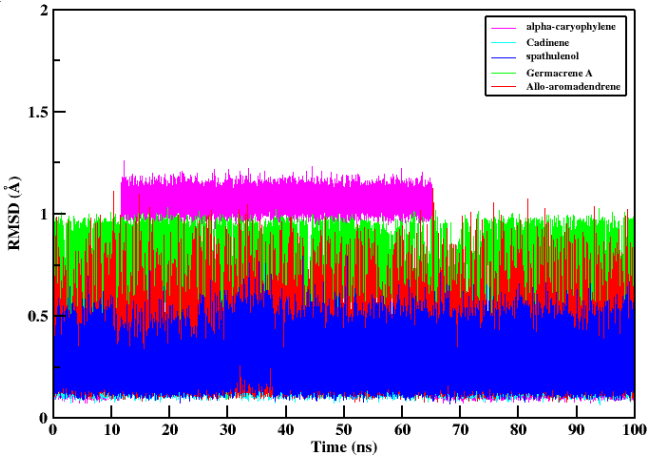


Fig. 4. RMSD plot of ligands relative to their initial structure for the 100 ns simulation time

Root mean square fluctuations (RMSF): Root mean square fluctuation (RMSF) computes amino acid residue-wise mobility of the protein backbone. A significant degree of mobility and instability in the protein receptor is indicated by the high fluctuations of amino acid residues. On the other hand, stable and rigid protein receptors are defined by low RMSF of amino acid residues. Fig. 5 shows the generated RMSF plot of all the studied protein-ligand complexes as well as of the Apo-

TABLE-5
CALCULATED VALUES OF AVERAGE RMSD, RADIUS OF GYRATION (Rg) AND SASA VALUES OF APO FORM OF SARS-CoV-2 SPIKE PROTEIN AND SPIKE PROTEIN-LIGAND COMPLEXES

System	RMSD (Å) (complex)	RMSD (Å) (ligand)	Radius of gyration (Å)	SASA (nm^2)
Apo-M^{pro}	2.06	–	18.39	98.43
M^{pro} - α-Caryophyllene	2.15	0.65	18.37	100.30
M^{pro} - Cadinene	2.13	0.20	18.32	98.52
M^{pro} - Spathulenol	2.65	0.26	18.33	100.82
M^{pro} - Germacrene A	2.15	0.69	18.35	100.02
M^{pro} - Allo-aromadendrene	2.20	0.29	18.36	99.55

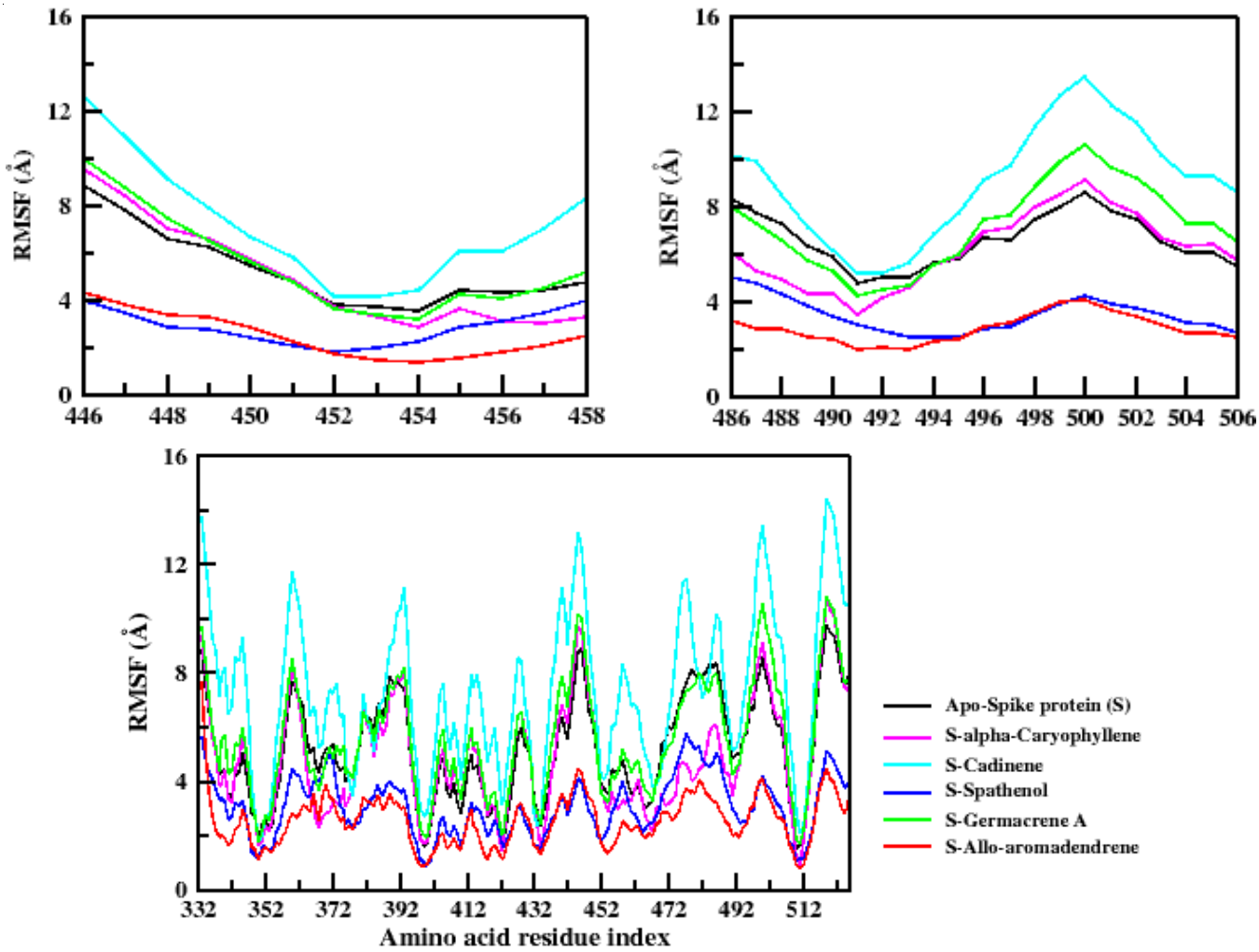


Fig. 5. RMSF plot of Apo form of SARS-CoV-2 spike protein and spike protein-ligand complexes for the 100 ns simulation time

form of the SARS-CoV-2 spike protein. As evident from the RMSF plot, spike-allo-aromadendrene displayed the lowest fluctuations of the involved amino acid residues. Similarly, the spike-spathulenol complex also maintained residue-wise fluctuations lower than the Apo-spike protein during the entire 100 ns simulation period. In case of both complexes, low fluctuations were observed in *aa* residues involved in the interaction of RBD with the ACE2 receptor. This signifies the stable and favourable incorporation of spathulenol and allo-aromadendrene ligands into the spike protein. The RMSF pattern of α-caryophyllene bound spike protein and Apo-spike protein were also found to be analogous, maintaining lower fluctuations in most of the residues of the receptor binding domain. The RMSF plot of spike-germacrene A complex was also more or less similar showing slightly more fluctuations in RBD region. On the contrary, the binding of cadinene to the RBD of spike protein had significantly elevated the fluctuations of all amino acid residues involved in comparison to Apo-spike protein, indicating unfavourable interaction of cadinene with the spike protein.

Radius of gyration (Rg): Furthermore, we have also evaluated the radius of gyration (Rg) value, which describes the compactness and folding of the receptor-ligand complexes during the molecular dynamic simulation period. While an unfolded or unstable form exhibits a higher Rg value, a compact form of the receptor is indicated by less variation in the gyration value. The Rg plot depicting the variation in Rg values over the 100 ns simulation period for the Apo-form of spike protein and its complexes with selected ligands is shown in Fig. 6. The average Rg values of Apo-spike protein, spike-α-caryophyllene, spike-cadinene, spike-spathulenol, spike-germacrene A and spike-allo-aromadendrene were obtained as 18.39, 18.37, 18.32, 18.33, 18.35 and 18.36 Å respectively. Rg plot of the spike-spathulenol complex was seemingly overlapping with that of the Apo-spike protein in most of the time frames of the 100 ns simulation period. Noticeably, Rg plots of spike protein complexes with the remaining four ligands *viz.* α-caryophyllene, cadinene, germacrene A and allo-aromadendrene also resembled those of Apo-spike protein, with substantially decreased Rg in the second half of simulation time scale. These findings suggest that the selected phytocompounds interact favourably, forming stable and rigid complexes with the RBD of the SARS-CoV-2 spike protein.

Solvent accessible surface area (SASA): SASA analysis is performed to evaluate the biomolecular surface area that is available to the surrounding molecules. SASA analysis helps

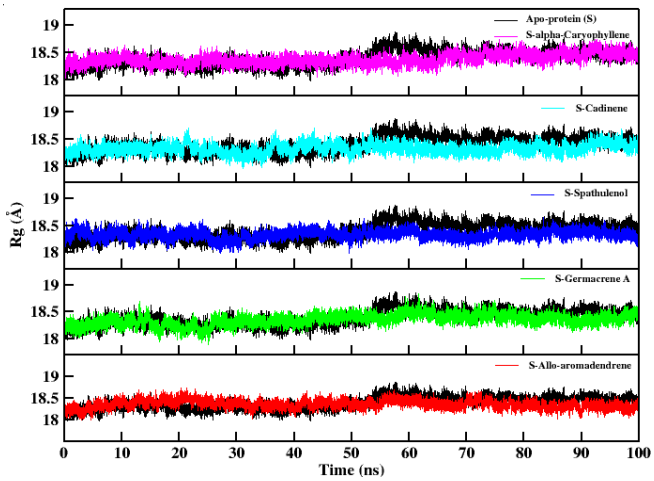


Fig. 6. Radius of gyration (Rg) plot of Apo form of SARS-CoV-2 spike protein and spike protein-ligand complexes for the 100 ns simulation time

understand the extent of protein volume expansion upon ligand inclusion. As seen in Fig. 7, SASA profiles of Apo-spike protein and all the studied spike-ligand complexes were more or less similar throughout the simulation. The evaluated average SASA values of Apo-spike protein, spike-α-caryophyllene, spike-cadinene, spike-spathulenol, spike-germacrene A and spike-allo-aromadendrene complexes were found to be 98.43, 100.30, 98.52, 100.82, 100.02 and 99.55 nm^2, respectively (Table-5). Thus, all the studied complexes showed seemingly little expansion in association with the ligand as compared to the Apo-spike protein. Thereby indicating a slight increase in the surface area accessible to solvent molecules. Nevertheless, their SASA profiles, particularly of spike-α-caryophyllene and

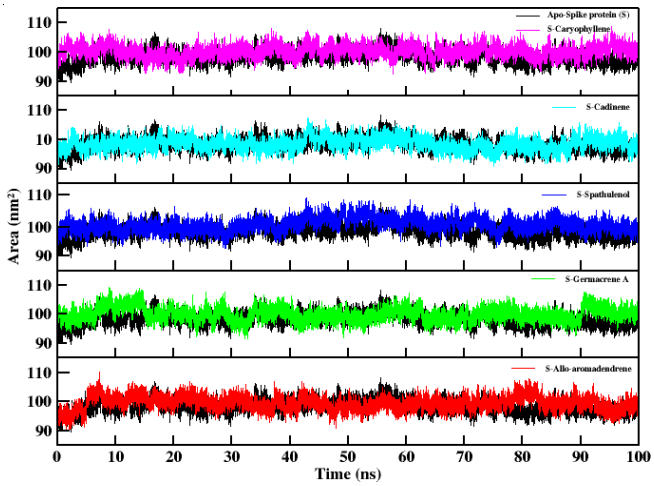


Fig. 7. Solvent accessible surface area (SASA) plot of Apo form of SARS-CoV-2 spike protein and spike protein-ligand complexes for the 100 ns simulation time

spike-spathulenol complexes were quite steady without any unusual change in SASA pattern. Thus, it can be summarized that these spike protein complexes on binding with phytocompounds have maintained decent stability in the solvated environment.

Hydrogen bond analysis: While analysing the stability of receptor-ligand complexes, it is important to consider the H-bond interactions, as these are one of the key transient interactions between the receptor and the ligand [70]. For all the 5 spike-ligand complexes, H-bonds distances were evaluated for 100 ns simulations to obtain the important H-bonds interactions with a decent percentage of occupancies. However, among the studied phytocompounds only spathulenol formed H-bonds with the spike protein residues. In case of the other four phytocompounds, *i.e.* α-caryophyllene, cadinene, germacrene A and allo-aromadendrene, no H-bonding interactions were observed between the ligands and the spike protein with appreciable occupancy. This may be attributed to the fact that these ligands do not contain any heteroatoms through which the ligands can form hydrogen bonds with amino acid residues of the spike protein. For spathulenol, H-bond interactions having a percentage of occupancy greater than 2% and the corresponding donor-acceptor bond distance are listed in Table-6. The highest H-bond occupancy was observed between Phe490 *aa* residue of spike protein as acceptor and oxygen (O1) atom of spathulenol as donor, displaying 13.92% occupancy. Phe490 residue also formed H-bond *via* N-atom as H-donor and oxygen (O1) atom of spathulenol as the H-acceptor atom, having 4.44% occupancy. In addition, oxygen (O1) atom of spathulenol also interacted with Thr495 and Ser494 residues, though the percentage of occupancy was extremely low (2.34% and 2.23% only). Thus, in terms of H-bond interactions, spathulenol is the most favourable phytocompound amongst the 5 studied ligands.

Principal component analysis (PCA): Since the spike-spathulenol system had favourable results in most of the analyzed parameters, therefore, to study the large-scale movements in the protein backbone, principal component analysis (PCA) was conducted for the Cα atom of spathulenol-bound spike protein and Apo-spike protein. For the studied systems, Fig. 8 shows the plot of the top 20 eigenvalues against their corresponding eigenvectors, as obtained from the diagonalization of the covariance matrix representing the atomic perturbations during 100 ns MD simulation. Fig. 8 shows that the major coordinated motions in the Cα atom of the protein backbone are predominantly represented by the first few eigenvalues. For Apo-protein and spike-spathulenol systems, the first five eigenvectors contribute 63.77% and 51.75% of the total movements, respectively over the 100 ns simulation period. In addition, since the first two principal components represented an extensive amount of motion variables in the trajectory, a

TABLE-6
HYDROGEN BOND OCCUPANCY OF THE LIGAND DURING 100 ns MD SIMULATION

Phytocompound	Donor	Acceptor	Occupancy (%)	Average distance (Å)
Spathulenol	LN3_195@O1	PHE_490@O	13.92	2.78
	PHE_490@N	LN3_195@O1	4.44	2.89
	LN3_195@O1	TYR_495@O	2.34	2.75
	SER_494@N	LN3_195@O1	2.23	2.89

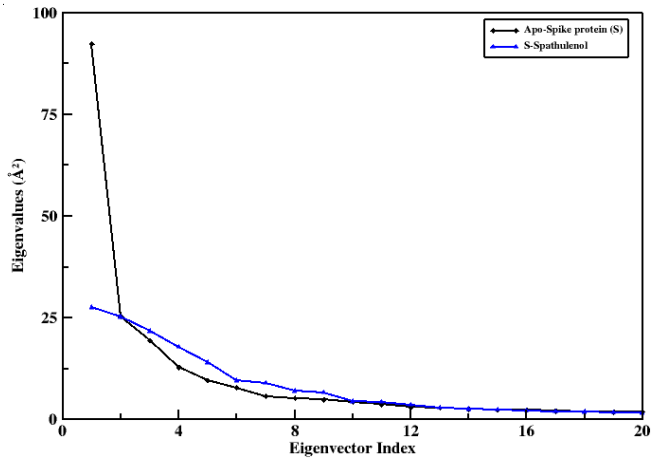


Fig. 8. 2D plot of eigenvalues *vs.* eigenvector indices obtained from the covariance matrix of Cα atomic motions of the studied protein-ligand systems for the 100 ns simulation time

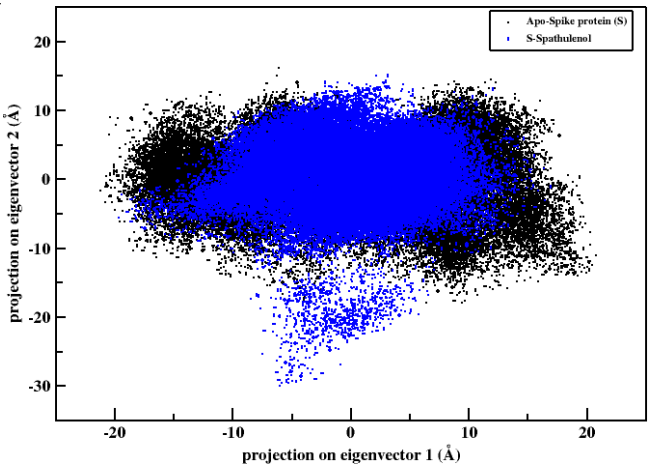


Fig. 9. 2D projection of eigenvector 1 *vs.* projection on eigenvector 2 (PC1 *vs.* PC2) of the studied protein-ligand systems for the 100 ns simulation time

2D projection of PC1 *vs.* PC2 was plotted (Fig. 9). In the PC1 *vs.* PC2 plot, less occupied space along the axes is interpreted to depict more stable motions than the dispersed ones. As per Fig. 9, in comparison to Apo-spike protein, spathulenol-bound spike protein is seen to project less conformational space along the PC1 axis, reflecting more restricted and compact Cα atomic motions in the protein backbone. These observations were also found to be aligned with RMSF findings.

Thereafter, Gibb's free energy landscape (FEL) was constructed using the PC1 and PC2 trajectory projections that depict the probability distribution of energy basins as variables (Fig. 10). Broadly, the appearance of low-energy basins in both the studied systems seems to be similar. However, in the Apo-spike protein, the presence of high-energy regions (red) separates the arrangement of low-energy basins (blue), making it appear to be dispersed. In contrast, the low-energy basins in spathulenol-bound spike protein were found to be clustered together, covering a larger surface area. Following that, for qualitative representation of Cα atomic motions of the protein backbone concerning the first eigenvector, porcupine plots were generated for the studied systems (Fig. 11). Porcupine plot interpretation can be done by comparing the length and direction of the arrows, which indicate the motion's magnitude and direction, respectively. The regions of the RBD (referred to as receptor binding motif or RBM) that directly interact with the ACE2 receptor on host cells are observed to be influenced noticeably by the incorporation of spathulenol to spike protein as compared to Apo-spike protein. The RBM region of the spike-spathulenol system was found to be substantially stable with decreased motion as demonstrated by the shorter magnitude of arrows compared to the long prominent arrows in the case of RBM of apo-spike protein (Fig. 11). Although enhanced motions were observed in the flexible terminal ends of the spike-spathulenol system, the central part of the RBD (core region) showed analogous motion behaviour preserving overall structure and stability.

MM-GBSA analysis: Calculations of MM-GBSA binding free energy were performed to validate the binding affinity and estimate the free energies of binding for five protein-ligand complexes. MM-GBSA calculation was carried out using the entire 50000 frames spanning over 100 ns of MD simulation

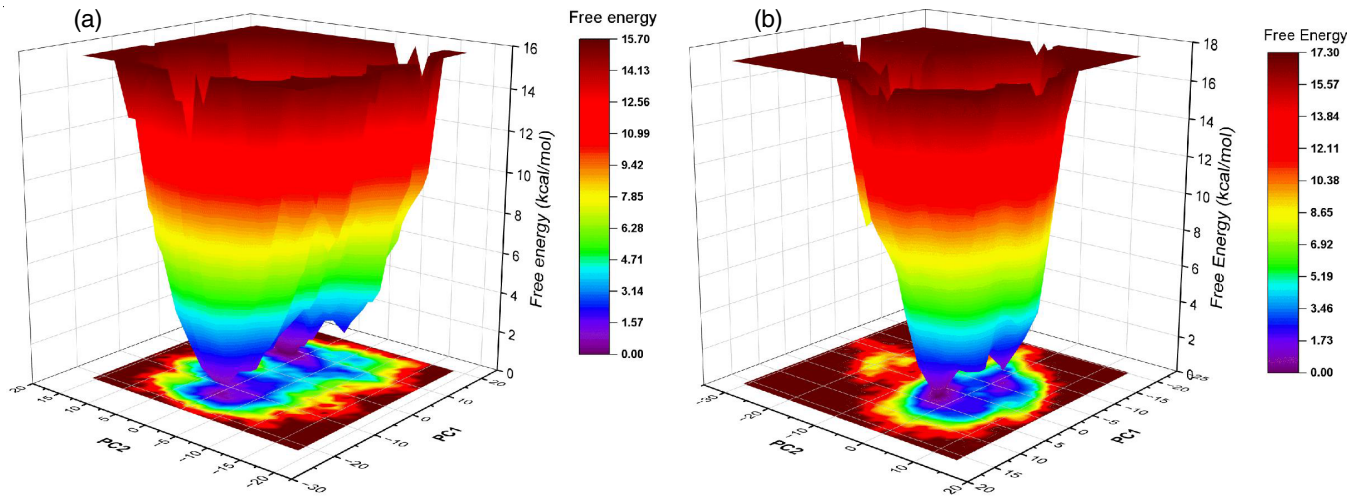


Fig. 10. Gibbs free energy landscape (FEL) generated using the PC1 and PC2 trajectory projections for (a) apo-spike protein and (b) spike-spathulenol system

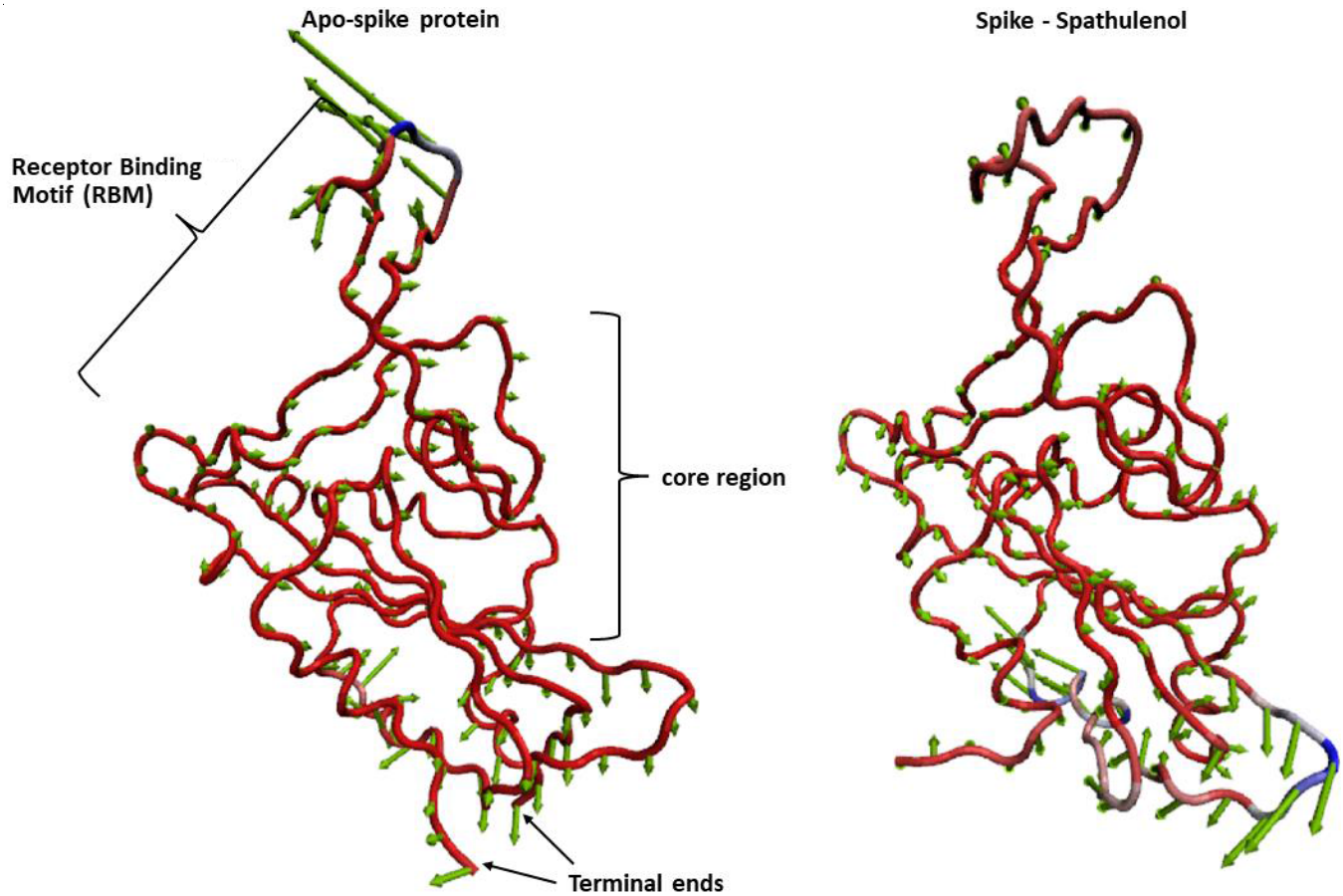


Fig. 11. Porcupine plots for the Cα atomic motions of protein backbone generated using PC1 for the studied systems

run. The free energies of binding for spike-α-caryophyllene, spike-cadinene, spike-spathulenol, spike-germacrene and spike-allo-aromadendrene were estimated to be -5.47, -8.78, -17.30, -2.42 and -10.45 kcal/mol, respectively (Table-7). These negative binding energy values indicate stable phytocompound binding to the RBD of the spike protein. Although, except for spathulenol, their binding affinity appears to be rather weak as evident in their small negative binding free energy values. Van der Waals interactions are the main factor favouring the obtained binding free energies of protein-ligand complexes, whereas electrostatic contributions were comparatively less. The most negative binding free energy in the spike-spathulenol complex could be correlated to the presence of hydrogen bonding interaction between spathulenol and spike protein residues. The 2D images of hydrogen bonding interaction between spathulenol and spike protein residues at various time intervals of 100 ns MD simulation run are given in Fig. 12. The combined contribution of more negative van der Waals energy and electrostatic energy in comparison to other protein-ligand complexes is also responsible for this observation. The different energy contributions towards MM-GBSA free energy of binding are represented in Fig. 13. This highly negative binding free energy of the spike-spathulenol system is also reflected in low conformational fluctuations of the receptor residues. Thus, based on docking, MD simulation and MM-GBSA calculation, it can be summarized that spathulenol is the most promising phytocompound to exhibit SARS-CoV-2 spike protein inhibitory property.

Conclusion

In this comprehensive study, three regional ethnomedicinal medicinal plants *Lindera neesiana*, *Litsea cubeba* and *Zanthoxylum armatum* were studied for their inhibitory properties towards key druggable SARS-CoV-2 target spike protein. The study employed various computational approaches like

TABLE-7
CALCULATED MM-GBSA FREE ENERGY OF BINDING OF SARS-CoV-2 SPIKE PROTEIN-LIGAND COMPLEXES OVER 100 ns MD SIMULATIONS

Compound	Calculated MM-GBSA value (kcal/mol)						
	ΔE_{vdw}	ΔE_{ele}	ΔE_{GB}	ΔE_{surf}	ΔG_{gas}	ΔG_{solv}	$\Delta G_{binding}$
α-Caryophyllene	-6.56	-0.27	4.58	-1.17	-8.88	3.40	-5.47
Cadinene	-10.66	-0.16	5.61	-1.71	-12.67	3.89	-8.78
Spathulenol	-14.23	-5.55	10.60	-2.13	-25.77	8.47	-17.30
Germacrene A	-3.84	-0.43	4.03	-0.83	-5.62	3.19	-2.42
Allo-aromadendrene	-13.21	-0.88	6.74	-1.94	-15.25	4.80	-10.45

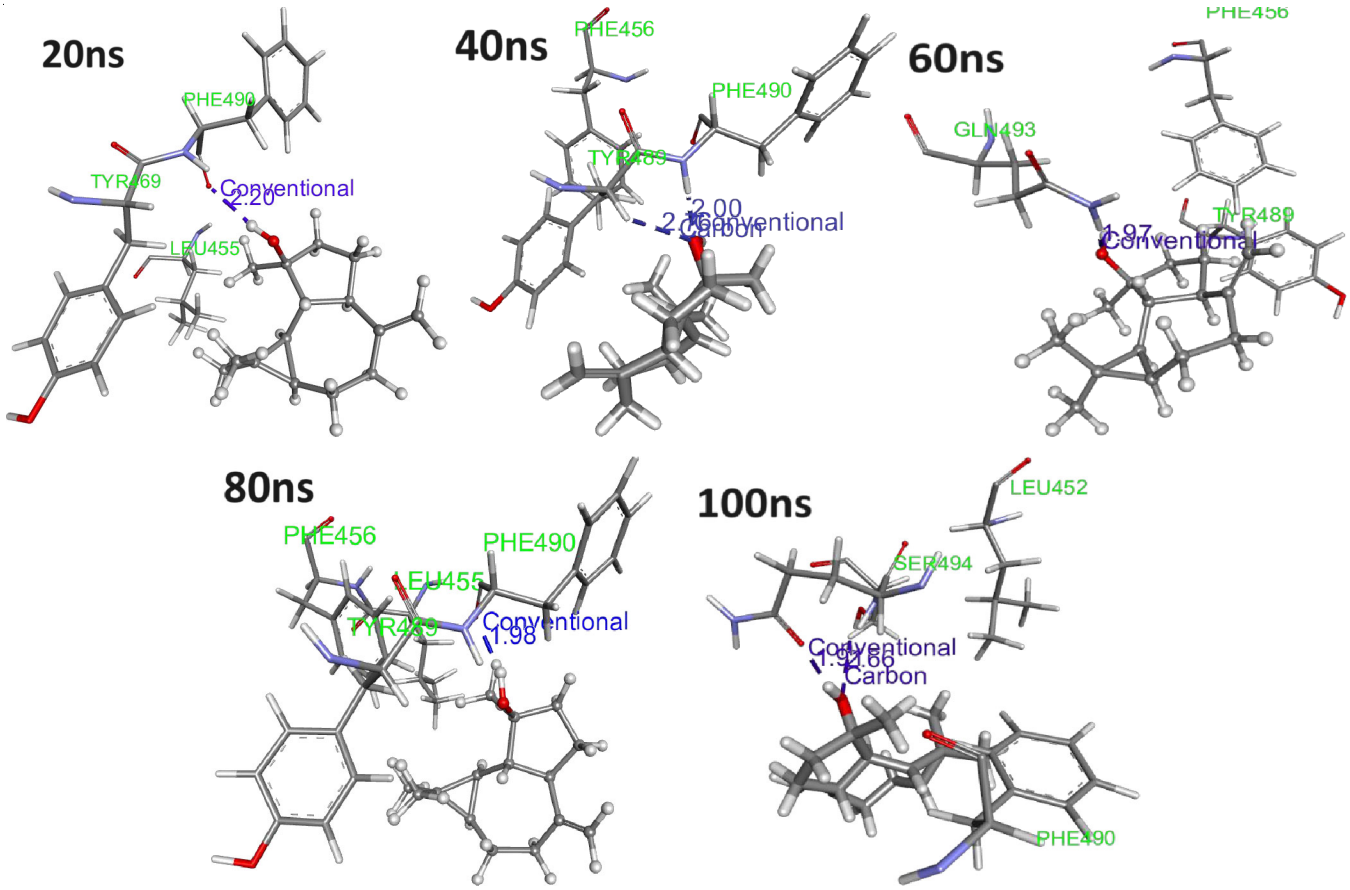


Fig. 12. H-bond interactions (both conventional H-bond and carbon-hydrogen bond) between SARS-CoV-2 spike protein and spathulenol compound at different time intervals of 100 ns MD simulation. The residues are shown in ball and stick representation while the ligand is shown in stick representation

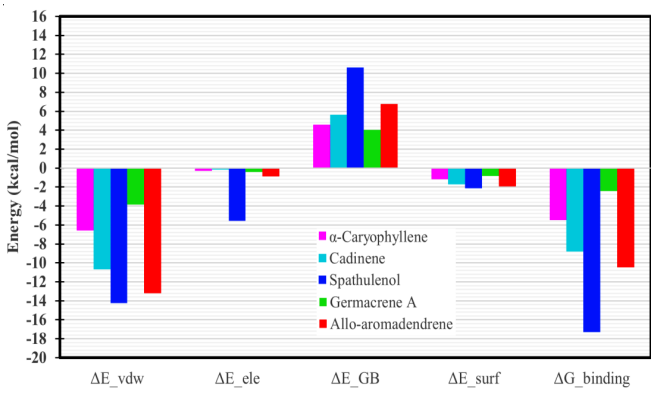


Fig. 13. MM-GBSA free energy of binding ($\Delta G_{binding}$) of best five spike protein-ligand complexes and its different energy contributions

molecular docking, drug-likeness, ADMET study, molecular dynamics simulation and MM-GBSA calculation. In total, we have docked 107 phytocompounds from the selected ethnomedicinal plants in search of new, potential SARS-CoV-2 spike protein inhibitors. Docking results suggested many phytochemical constituents of the chosen plants have good inhibition properties against spike protein. As per docking results, α-caryophyllene phytocompound showed the highest binding affinity towards RBD of spike protein with a binding energy value of -6.69 kcal/mol. The residues of the RBD of the target protein predominantly Arg403, Tyr449, Tyr453, Ser494, Tyr495 and Gly496, Phe497 and Tyr505 residues were observed to form hydrogen or hydrophobic interactions with the studied phytocompounds. Among 107 phytocompounds, 5 phytocompounds with the best binding energy values were shortlisted for further investigation. Drug-likeness and ADMET parameters were evaluated for the selected phytocompounds and overall, they displayed good physico-chemical and pharmacokinetic behaviour with no associated toxicities. The stabilities of the 5 spike-ligand complexes were further examined through 100 ns MD simulation. In particular, spathulenol phytocompound present in *L. neesiana* displayed the best desired result in all the studied parameters. The stable simulation trajectories, as observed in RMSD, RMSF, radius of gyration and SASA plots, provided evidence for the stable interaction of the spathulenol compound within the RBD of the spike protein. The MM-GBSA calculations further revealed spathulenol to have the most negative favourable binding energy, suggesting its strong affinity for RBD of the SARS-CoV-2 spike protein. Hence, the study outcomes support the promising inhibitory properties of the spathulenol compound of *Lindera neesiana* plant against the SARS-CoV-2 spike protein. It is anticipated that this computational investigation will offer avenues for further computational research, followed by experimental validation of its activity as a potential inhibitor of SARS-CoV-2 spike protein and will contribute to developing therapeutic drugs in the future for COVID-19 disease.

ACKNOWLEDGEMENTS

The authors are grateful to the Department of Chemical Sciences, Tezpur University for providing the computational facilities. One of the authors, S.S., also acknowledges the Department of Science and Technology, Government of India through the program named Innovation in Science Pursuit for Inspired Research (INSPIRE)-IF190945.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this article.

REFERENCES

1. A. Mishra, W.H. Khan and A.S. Rathore, *J. Chem. Inf. Model.*, **61**, 5708 (2021); https://doi.org/10.1021/acs.jcim.1c00994
2. A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng and T. Jiang, *Cell Host Microbe*, **27**, 325 (2020); https://doi.org/10.1016/j.chom.2020.02.001
3. S. Su, G. Wong, W. Shi, J. Liu, A.C. Lai, J. Zhou, W. Liu, Y. Bi and G.F. Gao, *Trends Microbiol.*, **24**, 490 (2016); https://doi.org/10.1016/j.tim.2016.03.003
4. P.A. Rota, M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Peñaranda, B. Bankamp, K. Maher, M. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D. Wang, D.D. Erdman, T.C.T. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Günther, A.D.M.E. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson and W.J. Bellini, *Science*, **300**, 1394 (2003); https://doi.org/10.1126/science.1085952
5. M. Rasmussen, F.T. Møller, V. Gunalan, S. Baig, M. Bennedbæk, L.E. Christiansen, A.S. Cohen, K. Ellegaard, A. Fomsgaard, K.T. Franck, N.B. Larsen, T.G. Larsen, R. Lassaunière, C. Polacek, A.G. Qvesel, R.N. Sieber, L.D. Rasmussen, M. Stegger, K. Spiess, M.-H.E. Tang, L.S. Vestergaard, T.E. Andersen, S.V. Hoegh, R.M. Pedersen, M.N. Skov, K. Steinke, T.V. Sydenham, M. Hoppe, L. Nielsen, T.G. Krause, H. Ullum and P. Jokelainen, *Euro Surveill.*, **28**, 2300460 (2023); https://doi.org/10.2807/1560-7917.ES.2023.28.36.2300460
6. L. Reeve, E. Tessier, A. Trindall, N.I.B. Abdul Aziz, N. Andrews, M. Futschik, J. Rayner, A. Didier'Serre, R. Hams, N. Groves, E. Gallagher, R. Graham, B. Kele, K. Hoschler, T. Fowler, E. Blandford, H. Mahgoub, J. Hoffmann, M. Ramsay, G. Dabrera, M. Chand, M. Zambon, A. Sharp, E. Heinsbroek and J. Lopez Bernal, *Euro Surveill.*, **28**, 2300489 (2023); https://doi.org/10.2807/1560-7917.ES.2023.28.39.2300489
7. A.S. Gonzalez-Reiche, H. Alshammary, S. Schaefer, G. Patel, J. Polanco, J.M. Carreño, A.A. Amoako, A. Rooker, C. Cognigni, D. Floda, A. van de Guchte, Z. Khalil, K. Farrugia, N. Assad, J. Zhang, B. Alburquerque, G. Kleiner, D. Andre, K.F. Beach, M.C. Bermúdez-González, G. Cai, N. Lyttle, L.C.F. Mulder, A. Oostenink, A.B.T. Salimbangon, G. Singh, M. van Kesteren, B. Monahan, J. Mauldin, M. Awawda, L.A. Sominsky, C. Gleason, K. Srivastava, R. Sebra, J.D. Ramirez, R. Banu, P. Shrestha, F. Krammer, A. Paniz-Mondolfi, E.M. Sordillo, V. Simon and H. van Bakel, *Nat. Commun.*, **14**, 3235 (2023); https://doi.org/10.1038/s41467-023-38867-x
8. J. Li, K.T. McKay, J.M. Remington and S.T. Schneebeli, *Sci. Rep.*, **11**, 16307 (2021); https://doi.org/10.1038/s41598-021-95826-6
9. S.S. Mishra, S. Ranjan, C.S. Sharma, H.P. Singh, S. Kalra and N. Kumar, *J. Biomol. Struct. Dyn.*, **39**, 4449 (2021); https://doi.org/10.1080/07391102.2020.1791957
10. Z.T. Muhseen, A.R. Hameed, H.M. Al-Hasani, M. Tahir ul Qamar and G. Li, *J. Mol. Liq.*, **320**, 114493 (2020); https://doi.org/10.1016/j.molliq.2020.114493
11. L. Wu, Q. Chen, K. Liu, J. Wang, P. Han, Y. Zhang, Y. Hu, Y. Meng, X. Pan, C. Qiao, S. Tian, P. Du, H. Song, W. Shi, J. Qi, H.-W. Wang, J. Yan, G.F. Gao and Q. Wang, *Cell Discov.*, **6**, 68 (2020); https://doi.org/10.1038/s41421-020-00210-9
12. J. Whisenant and K. Burgess, *ACS Med. Chem. Lett.*, **11**, 1076 (2020); https://doi.org/10.1021/acsmedchemlett.0c00233
13. M. Aminpour, M. Cannariato, A. Zucco, E. Di Gregorio, S. Israel, A. Perioli, D. Tucci, F. Rossi, S. Pionato, S. Marino, M.A. Deriu, K.K. Velpula and J.A. Tuszynski, *Biomedicines*, **9**, 1208 (2021); https://doi.org/10.3390/biomedicines9091208
14. R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo and Q. Zhou, *Science*, **367**, 1444 (2020); https://doi.org/10.1126/science.abb2762
15. B.J. Williams-Noonan, N. Todorova, K. Kulkarni, M.I. Aguilar and I. Yarovsky, *J. Phys. Chem. B*, **125**, 2533 (2021); https://doi.org/10.1021/acs.jpcb.0c11321
16. A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire and D. Veesler, *Cell*, **181**, 281 (2020); https://doi.org/10.1016/j.cell.2020.02.058
17. O.J. Matthew, A.O. Eludoyin and K.S. Oluwadiya, *Spat. Spatio-Temporal Epidemiol.*, **37**, 100417 (2021); https://doi.org/10.1016/j.sste.2021.100417
18. X.D. Yang, H.L. Li and Y.E. Cao, *Int. J. Environ. Res. Public Health*, **18**, 484 (2021); https://doi.org/10.3390/ijerph18020484
19. D.K. Ramasamy, *World J. Pharm. Life Sci.*, **6**, 303 (2020).
20. G.S. Bhunia, S. Roy and P.K. Shit, *Spatial Inform. Res.*, **29**, 661 (2021); https://doi.org/10.1007/s41324-020-00376-0
21. K. Kiewhuo, D. Gogoi, H.J. Mahanta, R.K. Rawal, D. Das and G.N. Sastry, *Comput. Biol. Chem.*, **100**, 107728 (2022); https://doi.org/10.1016/j.compbiolchem.2022.107728
22. A.A. Mao, T.M. Hynniewta and M. Sanjappa, *Indian J. Tradit. Knowled.*, **8**, 96 (2009).
23. R. Chakraborty, B. De, N. Devanna and S. Sen, *J. Nat. Prod. Plant Resour.*, **2**, 143 (2012).
24. S. Comai, S. Dall'Acqua, A. Grillo, I. Castagliuolo, K. Gurung and G. Innocenti, *Fitoterapia*, **81**, 11 (2010); https://doi.org/10.1016/j.fitote.2009.06.017
25. L. Si, Y. Chen, X. Han, Z. Zhan, S. Tian, Q. Cui and Y. Wang, *Molecules*, **17**, 7057 (2012); https://doi.org/10.3390/molecules17067057
26. H.P. Kayat, S.D. Gautam and R.N. Jha, *J. Pharmacogn. Phytochem.*, **5**, 58 (2016).
27. T.P. Singh and O.M. Singh, *Pharmacogn. J.*, **11**, (2011).
28. M. Tiwary, S.N. Naik, D.K. Tewary, P.K. Mittal and S. Yadav, *J. Vector Borne Dis.*, **44**, 198 (2007).
29. M. Kamle, D.K. Mahato, K.E. Lee, V.K. Bajpai, P.R. Gajurel, K.S. Gu and P. Kumar, *Plants*, **8**, 150 (2019); https://doi.org/10.3390/plants8060150
30. V. Nath, P. Buragohain and H. Sharma, *Curr. Trends Pharm. Res.*, **8**, 158 (2021).
31. S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, B.A. Shoemaker, Q. Li, P.A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang and E.E. Bolton, *Nucleic Acids Res.*, **51(D1)**, D1373 (2023); https://doi.org/10.1093/nar/gkac956
32. G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell and A.J. Olson, *J. Comput. Chem.*, **30**, 2785 (2009); https://doi.org/10.1002/jcc.21256
33. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig I.N. Shindyalov and P.E. Bourne, *Nucleic Acids Res.*, **28**, 235 (2000); https://doi.org/10.1093/nar/28.1.235
34. J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang and X. Wang, *Nature*, **581**, 215 (2020); https://doi.org/10.1038/s41586-020-2180-5
35. G.M. Morris, R. Huey and A.J. Olson, *Curr. Protoc. Bioinformatics*, **24**, 8 (2008); https://doi.org/10.1002/0471250953.bi0814s24
36. D.S. Goodsell and A.J. Olson, *Proteins*, **8**, 195 (1990); https://doi.org/10.1002/prot.340080302
37. S. Gupta, A.K. Singh, P.P. Kushwaha, K.S. Prajapati, M. Shuaib, S. Senapati and S. Kumar, *J. Biomol. Struct. Dyn.*, **39**, 4334 (2021); https://doi.org/10.1080/07391102.2020.1776157

38. S. Jena, P. Munusami, B. Mm and K. Chanda, *Virusdisease*, **32**, 65 (2021); https://doi.org/10.1007/s13337-021-00666-7
39. S. Murugesan, S. Kottekad, I. Crasta, S. Sreevathsan, D. Usharani, M.K. Perumal and S.N. Mudliar, *Comput. Biol. Med.*, **136**, 104683 (2021); https://doi.org/10.1016/j.compbiomed.2021.104683
40. D.S. Biovia, Discovery Studio Modeling Environment. Dassault Syst. Release, San Diego, version 4 (2020).
41. A. Daina, O. Michielin and V. Zoete, *Sci. Rep.*, **7**, 42717 (2017); https://doi.org/10.1038/srep42717
42. C.A. Lipinski, F. Lombardo, B.W. Dominy and P.J. Feeney, *Adv. Drug Deliv. Rev.*, **64**, 4 (2012); https://doi.org/10.1016/j.addr.2012.09.019
43. A.K. Ghose, V.N. Viswanadhan and J.J. Wendoloski, *J. Comb. Chem.*, **1**, 55 (1999); https://doi.org/10.1021/cc9800071
44. D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward and K.D. Kopple, *J. Med. Chem.*, **45**, 2615 (2002); https://doi.org/10.1021/jm020017n
45. W.J. Egan, K.M. Merz and J.J. Baldwin, *J. Med. Chem.*, **43**, 3867 (2000); https://doi.org/10.1021/jm000292e
46. pKCSM; http://biosig.unimelb.edu.au/pkcsm/
47. D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman, AMBER 2018, University of California, San Francisco (2018).
48. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Aricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven Jr., J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J. Cioslowski, J.V. Ortiz and D.J. Fox, Gaussian 09, Revision A.1, Gaussian Inc., Wallingford CT (2009).
49. J.W. Ponder and D.A. Case, *Adv. Protein Chem.*, **66**, 27 (2003); https://doi.org/10.1016/S0065-3233(03)66002-X
50. J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, **11**, 3696 (2015); https://doi.org/10.1021/acs.jctc.5b00255
51. D.J. Price and C.L. Brooks III, *J. Chem. Phys.*, **121**, 10096 (2004); https://doi.org/10.1063/1.1808117
52. S. Ahmad, H.W. Abbasi, S. Shahid, S. Gul and S.W. Abbasi, *J. Biomol. Struct. Dyn.*, **39**, 4225 (2021); https://doi.org/10.1080/07391102.2020.1775129
53. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, **98**, 10089 (1993); https://doi.org/10.1063/1.464397
54. U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee and L.G. Pedersen, *J. Chem. Phys.*, **103**, 8577 (1995); https://doi.org/10.1063/1.470117
55. H.C. Andersen, *J. Chem. Phys.*, **72**, 2384 (1980); https://doi.org/10.1063/1.439486
56. K. Kodchakorn, Y. Poovorawan, K. Suwannakarn and P. Kongtawelert, *J. Mol. Graph. Model.*, **101**, 107717 (2020); https://doi.org/10.1016/j.jmgm.2020.107717
57. J.P. Ryckaert, G. Ciccotti and H.J. Berendsen, *J. Comput. Phys.*, **23**, 327 (1977); https://doi.org/10.1016/0021-9991(77)90098-5
58. A.W. Gotz, M.J. Williamson, D. Xu, D. Poole, S. Le Grand and R.C. Walker, *J. Chem. Theory Comput.*, **8**, 1542 (2012); https://doi.org/10.1021/ct200909j
59. D.R. Roe and T.E. Cheatham III, *J. Chem. Theory Comput.*, **9**, 3084 (2013); https://doi.org/10.1021/ct400341p
60. GROMACS, https://doi.org/10.5281/zenodo.10721181
61. P.J. Turner, XMGRACE, Version 5.1. 19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR (2005).
62. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, **14**, 33 (1996); https://doi.org/10.1016/0263-7855(96)00018-5
63. R. Ghosh, A. Chakraborty, A. Biswas and S. Chowdhuri, *J. Biomol. Struct. Dyn.*, **40**, 2053 (2022); https://doi.org/10.1080/07391102.2020.1835728
64. B.R. Miller III, T.D. McGee Jr., J.M. Swails, N. Homeyer, H. Gohlke and A.E. Roitberg, *J. Chem. Theory Comput.*, **8**, 3314 (2012); https://doi.org/10.1021/ct300418h
65. P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case and T.E. Cheatham, *Acc. Chem. Res.*, **33**, 889 (2000); https://doi.org/10.1021/ar000033j
66. J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman and D.A. Case, *J. Comput. Chem.*, **25**, 1157 (2004); https://doi.org/10.1002/jcc.20035
67. I. Massova and P.A. Kollman, *Perspect. Drug Discov. Des.*, **18**, 113 (2000); https://doi.org/10.1023/A:1008763014207
68. G.R. Bickerton, G.V. Paolini, J. Besnard, S. Muresan and A.L. Hopkins, *Nat. Chem.*, **4**, 90 (2012); https://doi.org/10.1038/nchem.1243
69. D.E. Pires, T.L. Blundell and D.B. Ascher, *J. Med. Chem.*, **58**, 4066 (2015); https://doi.org/10.1021/acs.jmedchem.5b00104
70. T. Lynch and A.M.Y. Price, *Am. Fam. Physician*, **76**, 391 (2007).
71. C.C. Ogu and J.L. Maxa, *Baylor Univ. Med. Center Proc.*, **13**, 421 (2000); https://doi.org/10.1080/08998280.2000.11927719
72. S.M. Lamothe, J. Guo, W. Li, T. Yang and S. Zhang, *J. Biol. Chem.*, **291**, 20387 (2016); https://doi.org/10.1074/jbc.M116.743138