



Cite this: *RSC Adv.*, 2020, 10, 31385

# Curcumin to inhibit binding of spike glycoprotein to ACE2 receptors: computational modelling, simulations, and ADMET studies to explore curcuminoids against novel SARS-CoV-2 targets

Dhivya Shanmugarajan,<sup>a</sup> Prabitha P.,<sup>a</sup> B. R. Prashantha Kumar <sup>\*a</sup> and B. Suresh<sup>b</sup>

The recent emergence of the novel coronavirus (SARS-CoV-2) has raised global concern as it is declared a pandemic by the WHO. However, to date, there is no current regimen to mitigate the molecular pathogenesis of SARS-CoV-2 virus. Curcuminoids, bioactive ingredients present in *Curcuma longa* (turmeric), are known to exhibit diverse pharmacological properties. To the best of our understanding to date, SARS-CoV-2 uses angiotensin-converting enzyme 2 (ACE2) for the host cellular entry. This is mediated via proteins of SARS-CoV-2, especially the spike glycoprotein receptor binding domain. Accordingly, our primary objective is to thwart virus replication and binding to the host system, leading us to probe curcuminoids efficiency towards key surface drug target proteins using the computational biology paradigm approach. Specifically, fourteen natural curcuminoids were studied for their possibility of inhibiting SARS-CoV-2. We studied their *in silico* properties towards SARS-CoV-2 target proteins by homology modelling, ADME, drug-likeness, toxicity predictions, docking molecular dynamics simulations and MM-PBSA free energy estimation. Among the curcuminoids docked to the receptor binding domain of SARS-CoV-2 spike glycoprotein, the keto and enol forms of curcumin form strong hydrogen bond interactions with ACE2 binding residues Q493, T501, Y505, Y489 and Q498. Molecular dynamics simulations, free energy binding and interaction energy validated the interaction and stability of the docked keto and enol forms of curcumin.

Received 8th April 2020

Accepted 8th August 2020

DOI: 10.1039/d0ra03167d

rsc.li/rsc-advances

## 1. Introduction

The world has recently witnessed an outbreak of a potentially lethal coronavirus, nCovid-19, an urgent public health issue with an increase in morbidity and mortality. Coronaviruses belonging to the family Coronaviridae significantly threaten human health, and other species. Recent findings claim that the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is evolutionarily closely related to bat coronavirus and SARS-CoV with a homology of  $\geq 95\%$  and  $\geq 70\%$ , respectively.<sup>1–3</sup> SARS-CoV-2 is a rapidly transmissible virus that tends to change its genetic material to survive in various environmental conditions. Structurally, this virus contains a nucleocapsid (N), where its genome is packed inside a helical capsid. A membrane protein (M) and small envelope protein play a major role in the virus assembly, and the spike glycoprotein or S protein plays a key role in host entry by harboring certain crucial amino acid residues of human angiotensin-converting enzyme 2 (*hACE2*). The

spike glycoprotein forms a large protrusion on the surface of the virus, exhibiting a crown-like appearance, and hence the name coronaviruses,<sup>4–8</sup> as shown in Fig. 1.

Similar to SARS-CoV, SARS-CoV-2 also facilitates its entry into human angiotensin-converting enzyme 2 (ACE-2) through the receptor-binding domain (RBD) of the spike glycoprotein. Angiotensin-converting enzyme 2 (ACE-2), is an enzyme located

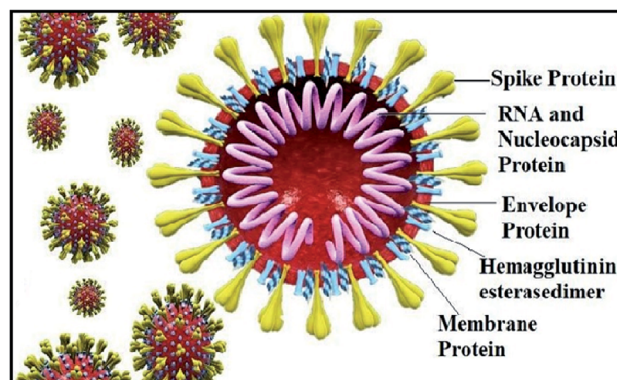


Fig. 1 Structure of coronavirus showing its structural proteins.

<sup>a</sup>Department of Pharmaceutical Chemistry, JSS College of Pharmacy, JSS Academy of Higher Education & Research, Mysuru 570 015, India. E-mail: brprashanthkumar@jssuni.edu.in; Fax: +91-821-2548359; Tel: +91-821-2548353

<sup>b</sup>JSS Academy of Higher Education & Research, Mysuru 570 015, India



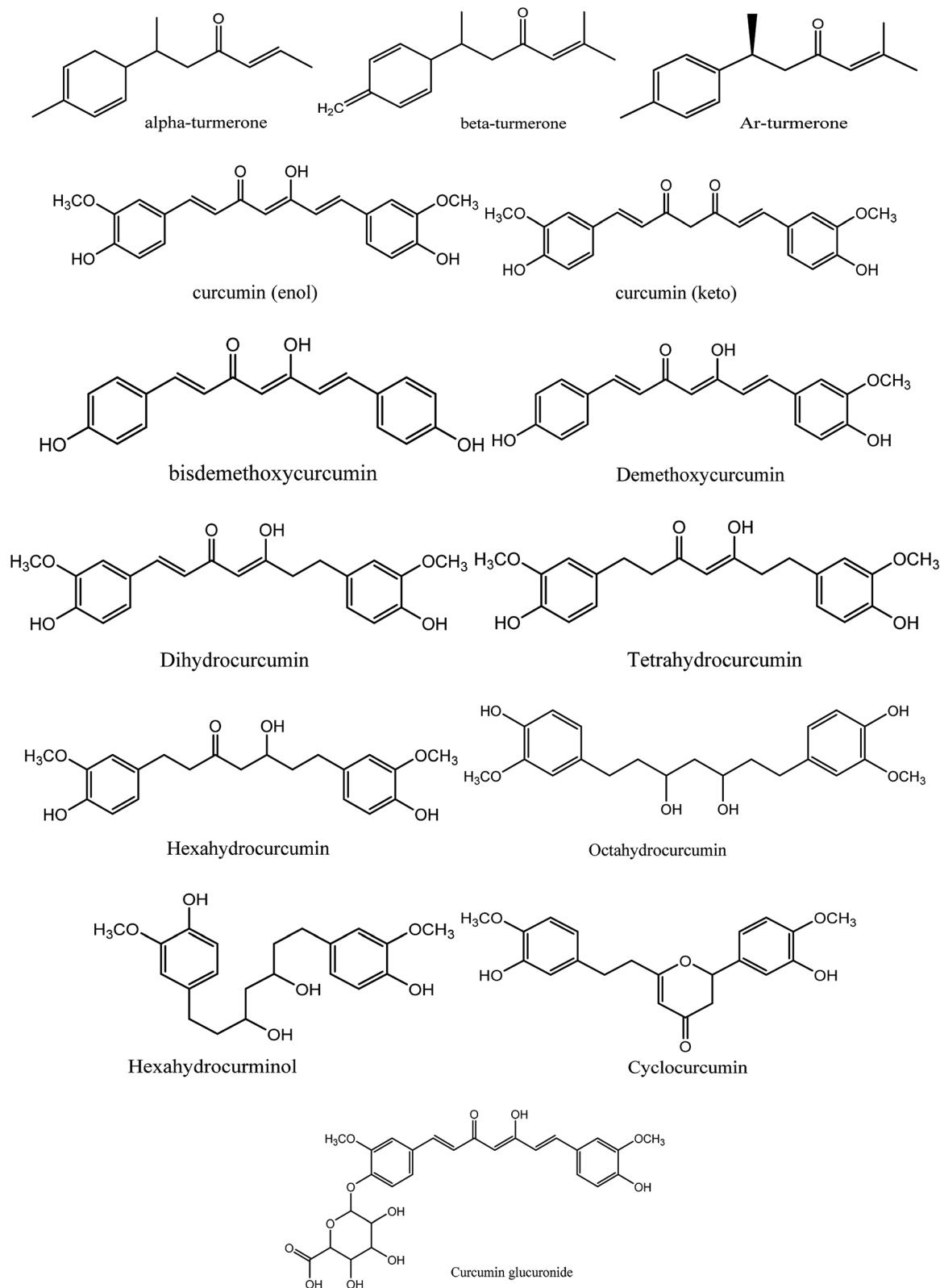


Fig. 2 Chemical structures of 14 natural curcumin derivatives used in this study.

on various parts of the body including alveolar epithelial cells of the lung, intestinal absorptive cells or enterocytes of small intestine, venous endothelial cells of the kidney, endothelial

cells of the heart and renal tubular epithelial cells.<sup>9</sup> Moreover, the binding affinity of SARS-CoV-2 on hACE2 is 10–20-fold higher than SARS-CoV-2002.<sup>10</sup> Hence, targeting the entry of the



SARS-CoV-2 spike glycoprotein RBD is considered a new therapeutic intervention.<sup>11</sup> However, in the current pandemic situation, traditional drug discovery or vaccine development is a daunting and time-consuming task, which can be offset by computational-aided drug design.

Moreover, instead of designing new lead compounds, drug repurposing or the use of natural products can be an alternative for the development of new antiviral compounds, which will enhance the speed of research in this area. In the current study, we focused on the natural plant *Curcuma longa*, commonly known as turmeric, a perennial herbaceous rhizomatous plant belonging to the ginger family Zingiberaceae, which is widely used in India.<sup>12</sup> *Curcuma longa* chemical constituents are widely used for treating various ailments and possess a wide variety of therapeutic properties including antiviral,<sup>13</sup> analgesic,<sup>14</sup> antimicrobial,<sup>15</sup> antiproliferative,<sup>16</sup> and anti-inflammatory<sup>17</sup> activity (Fig. 2).

Among the various chemical compounds, curcumin has gained importance among researchers because its compounds have been exhibited activity against viruses such as the human immunodeficiency virus (HIV), dengue virus, herpes simplex virus (HSV), hepatitis virus, influenza A virus (IAV), and Ebola virus.<sup>18–21</sup> Recently, reports on computational drug design demonstrated the therapeutic potential of curcumin as a dual inhibitory agent acting on S-protein and ACE-2.<sup>22</sup> Thus, the

main rationale of this study was to search for potential curcumin derivatives against these two drug targets, *i.e.* the spike glycoprotein RBD region and envelope protein. Besides, the pharmacokinetics, pharmacodynamics and drug-likeness of all the curcumin derivatives were studied. Finally, a molecular dynamics simulation study was performed to understand the stability of the protein–ligand complex with time, and a CHARMM interaction analysis and free energy binding using MM-PBSA demonstrated the stability of curcumin complexed with the spike glycoprotein.

## 2. Results and discussion

### 2.1 Homology modelling

Homology models are the backbone of structural biology, where a structure without NMR or X-ray crystallography data can be modelled using various machine learning algorithms. Moreover, it is a reliable and widely used technique for predicting unknown protein structures. The experimental protein structures of the novel coronavirus SARS-CoV-2 belonging to the genus Betacoronavirus are limited in structural databases. Therefore, the template modelling concept was implemented. Sequences were scanned in the TMpred tool to identify the TM-alpha helices among the three query sequences, and only the envelope protein showed transmembrane helices (Fig. 3A).

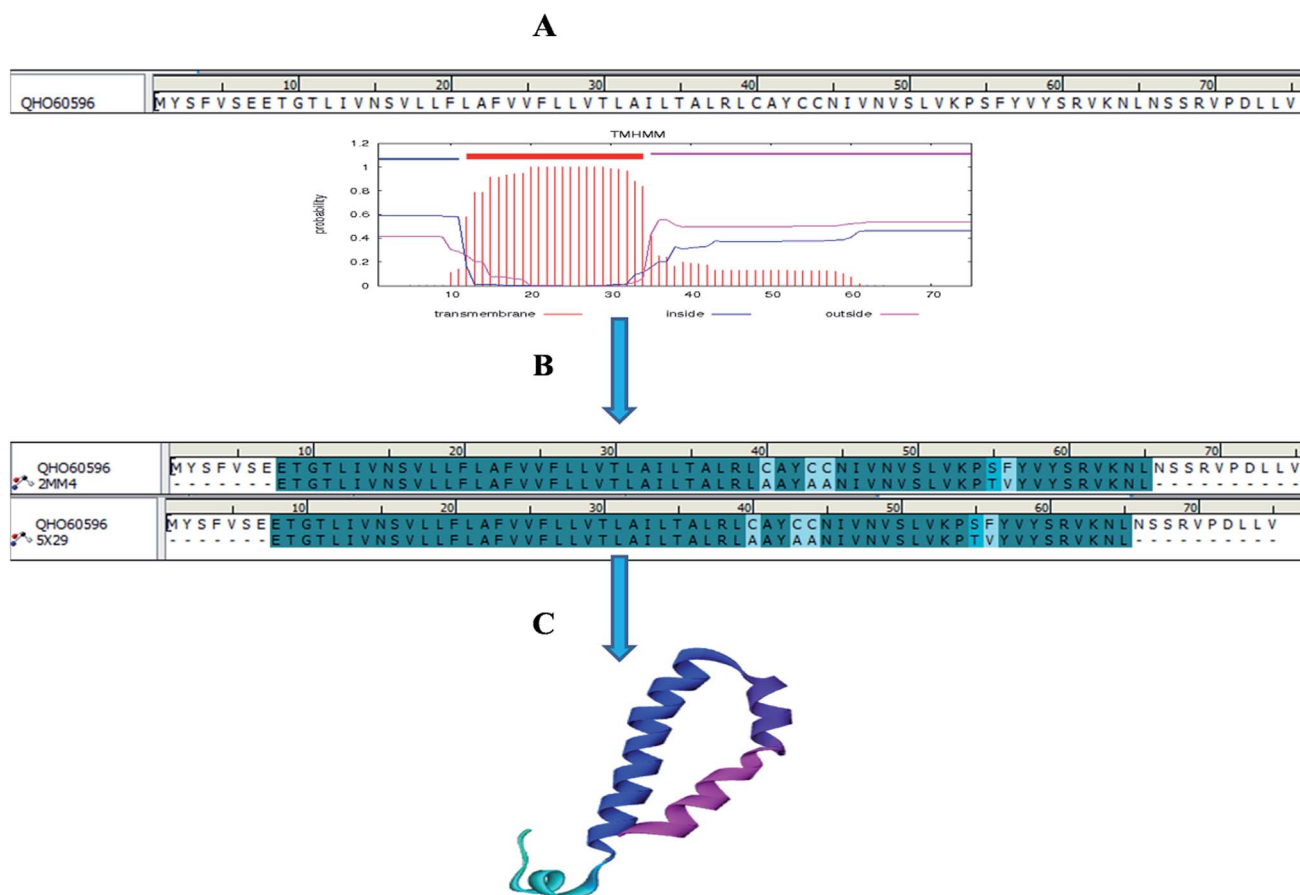


Fig. 3 Envelope protein of SARS-CoV-2. (A) Transmembrane prediction, (B) query to template alignment, and (C) homology modelled structure.





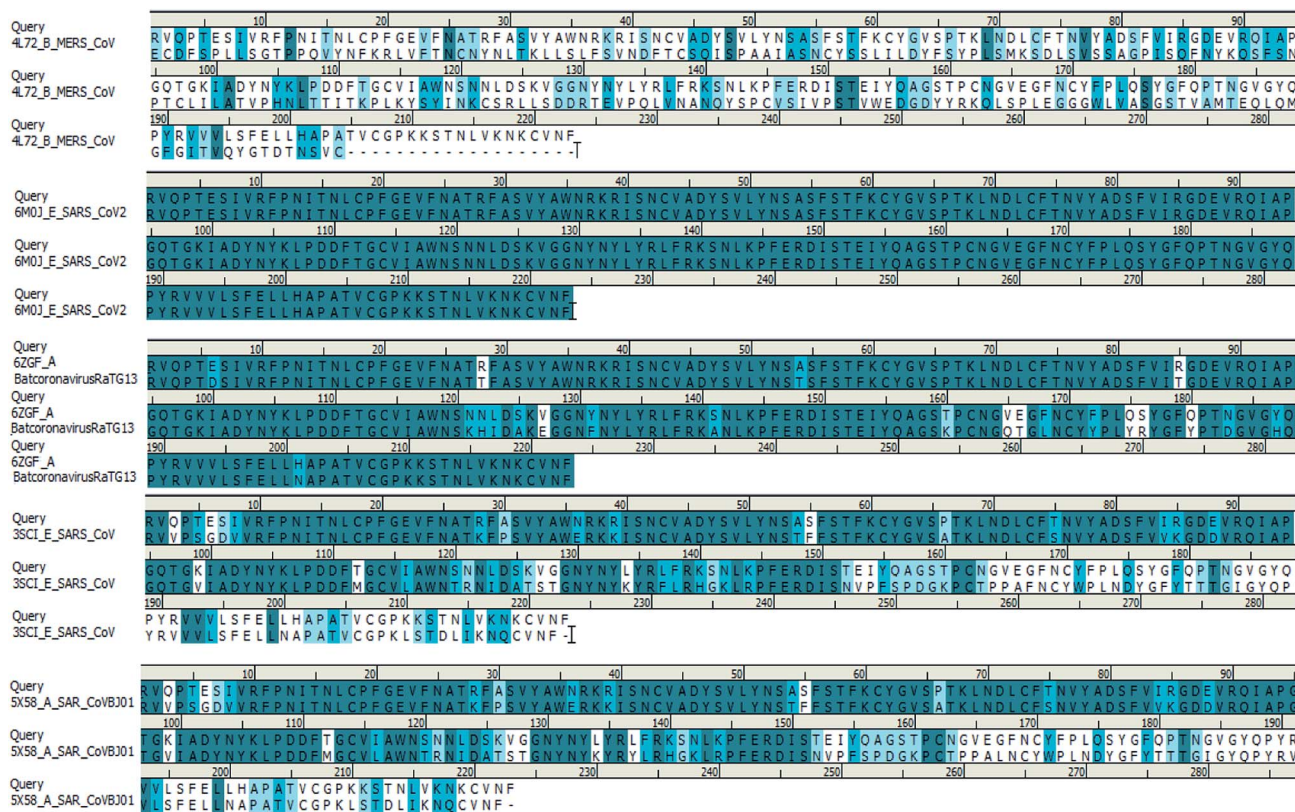


Fig. 4 Spike glycoprotein\_RBD query sequence of SARS-CoV-2 alignment with different template structures of MERS-CoV, batcoronavirusRaTG13, SARS-CoV and SARS-CoVBj01.

We searched for templates for each query sequence using BLAST search against PDB databases. According to the BLAST search of the spike glycoprotein\_RBD (333–526) region, it shows 100% template coverage with the RBD region of SARS-CoV-2 (PDB ID 6M0J\_E) with a resolution of 2.45 Å. Also, the homology of bat coronavirus RaTG13 (PDB ID 6ZGF\_A), SARS-CoV protein (PDB ID 3SCI\_E), SARS-CoV BJ01 (PDB ID 5X58\_A) and SARS-CoV-2 is about 90.13%, 73.54%, and 73.09% identity with query coverage of 100%, respectively. In contrast, the homology of the MERS S-protein (PDB ID 4L72\_B) and SARS-CoV-2 S-protein RBD is about 24% identity and the query coverage rate is 44%. The BLAST search results of the query sequence to the PDB structure alignment is depicted in Fig. 4.

Subsequently, the BLAST search of the envelope protein (QHS34548) of SARS-CoV-2 showed the homology of SARS-CoV PDB ID's 2MM4\_A and 5X29\_A. SARS-CoV envelope protein 5X29\_A shows maximum query coverage of 82% with 88.71% identity and 2MM4\_A of SARS-CoV sequence alignment with a query of only about 77% with 91.38% identity. However, the aligning query sequence with individual templates of SARS-CoV using the MODELLER algorithm showed 83% sequence similarity and 81.5% sequence identity (Fig. 3B). Hence, template 5X29\_A was used for homology modelling of the SARS-CoV-2 envelope protein. However, no template match was found with MERS-CoV. On the other hand, the membrane protein sequence for template structure identification showed no significant similarity results with the default scoring matrix of

BLASTp. Consequently, other scoring matrices were introduced to check the structural similarity of the sequence, excluding an iterative process that showed no significant similarity to the query. Hence, the modelled envelope protein (Fig. 3C) and available X-ray crystallography structure 6M0J\_E spike glycoprotein\_RBD of SARS-CoV-2 were used for further studies.

## 2.2 Structure validation process

The modelled envelope protein structure from the SARS-CoV structural parameters was studied for validation purposes. The best-modeled structure using MODELLER was initially chosen based on the lower discrete optimized protein energy (DOPE)<sup>23</sup> of  $-4812.83 \text{ kcal mol}^{-1}$ , and the higher total PDF (probability density function) of 165.148 shows 92.98 ERRAT quality factor. The Procheck tool evaluates the residue-by-residue stereochemical quality, structure geometry and distribution of Phi and Psi angle amino acids in favoured, allowed and generously allowed region of a modelled protein.<sup>24</sup> The Ramachandran plot (RP) showed 91.8% amino acids in the core region or most favoured regions and 8.2% residues were in the additional allowed regions with no outliers (Fig. 5A). Most of the amino acids in the favoured region signify that the modelled structure of the envelope protein is reliable, and can be equally compared with NMR structure quality. QMEANBrane is a unique structure quality assessing SWISS-MODEL tool explicitly used for modelled transmembrane protein investigation.<sup>25</sup> The envelope protein probed in SWISS-



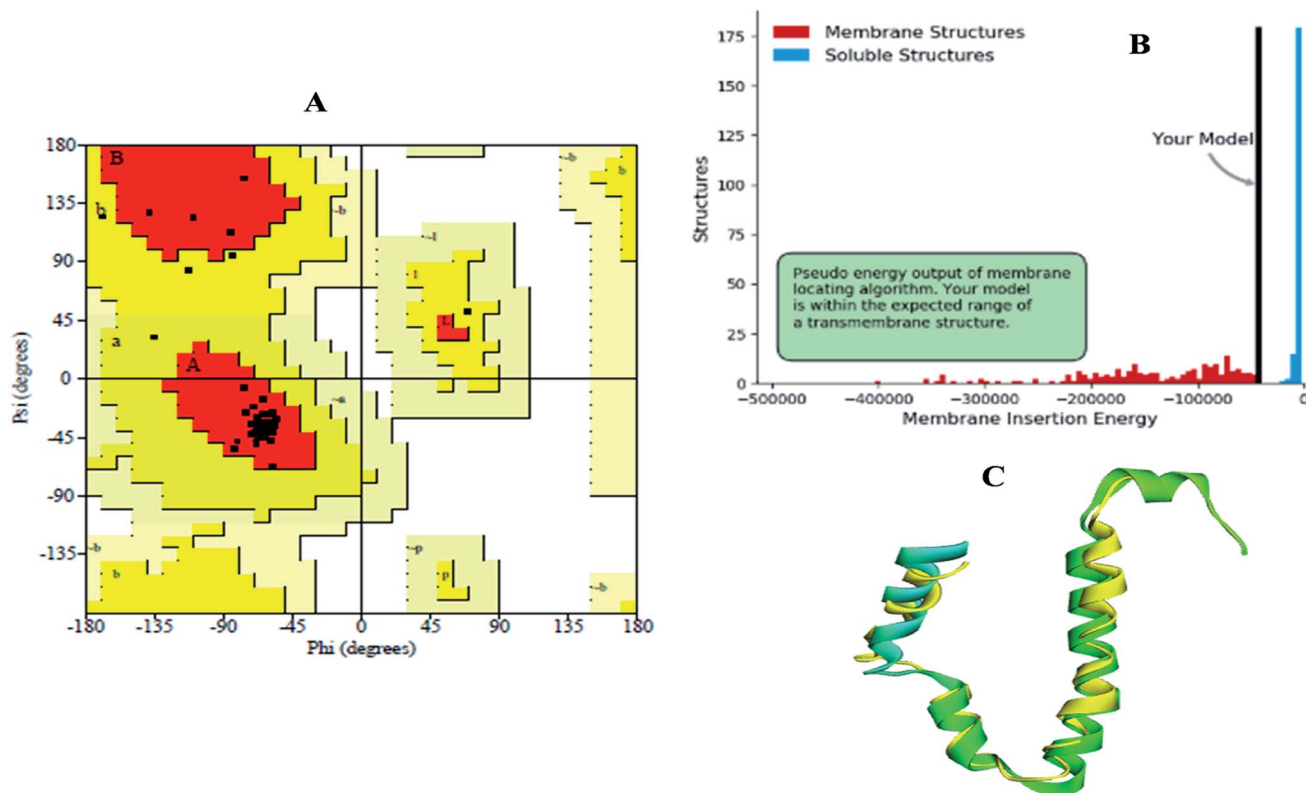


Fig. 5 (A) Ramachandran plot, (B) transmembrane structure validation, and (C) superimposed model of modelled structure and template structure.

MODEL was within the transmembrane insertion energy, and also satisfied statistical potentials in the naturally occurring oligomeric state (Fig. 5B). The template (5X29) and modelled

structure superimposition are shown in Fig. 5C, where the overall structure root mean square deviation is  $<2.5$  Å. The quality of the structure was further confirmed using the RAMPAGE tool. In

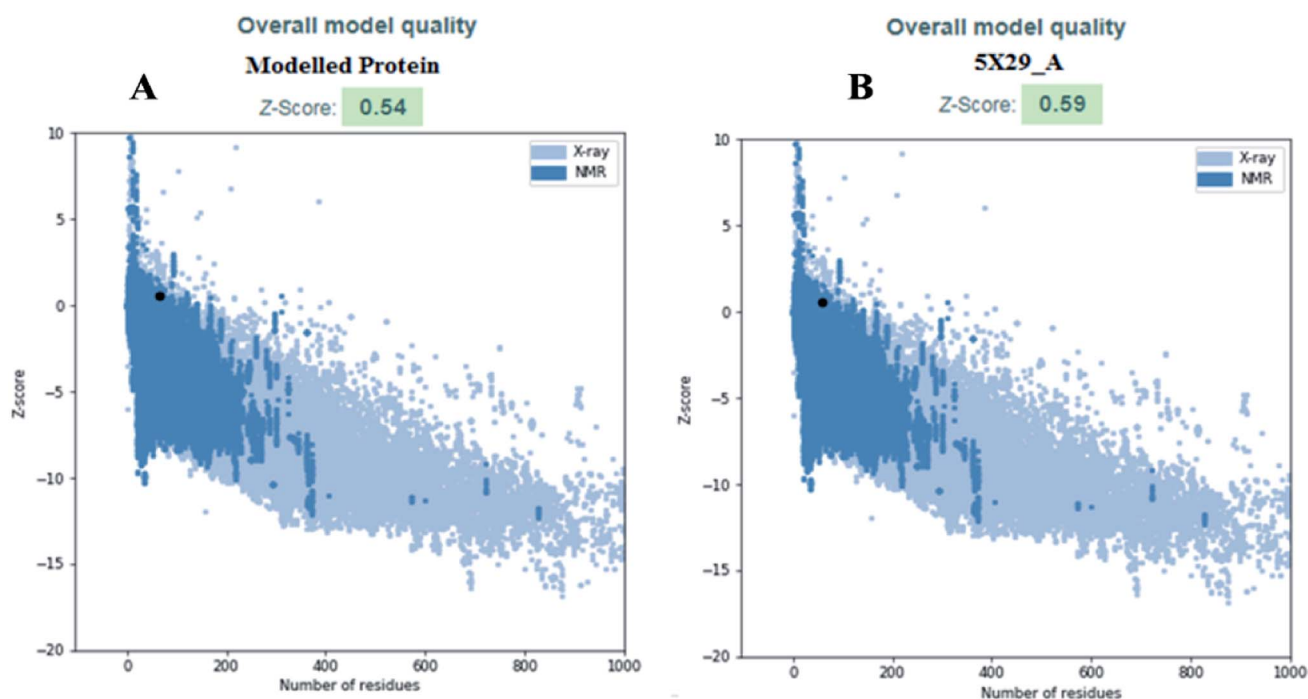


Fig. 6 (A) Modelled envelope protein and (B) template structure used for modelling envelope protein.

addition, ProSA structure analysis<sup>26</sup> was performed for the modelled protein and template structure, where the Z-scores of 0.54 and 0.59, respectively, located in the NMR structure region further confirm the modelled protein structure quality (Fig. 6).

### 2.3 Computational pharmacoinformatic study on curcumin derivatives

ADMET, TOPKAT and drug-likeness assessment for lead molecules can provide insight into their quantitative structure–property relationship. Thus, the models used in the *in silico*

pharmacokinetic study were obtained through a quantitative structure–activity relationship (QSAR). Furthermore, they reduce the research time, biological waste and cost. ADMET was initially used to screen a large library of compounds. Descriptors such as the blood–brain barrier (BBB) penetration, hepatotoxicity and CYP2D6 enzyme are high priority models followed by Ames mutagenicity and carcinogenicity. All the curcuminoid compounds were tested using the ADMET model. The pharmacokinetic analysis results are shown in (Fig. 7). According to the ADMET plot  $\text{Alog } p_{98}$  vs. PSA, it can be

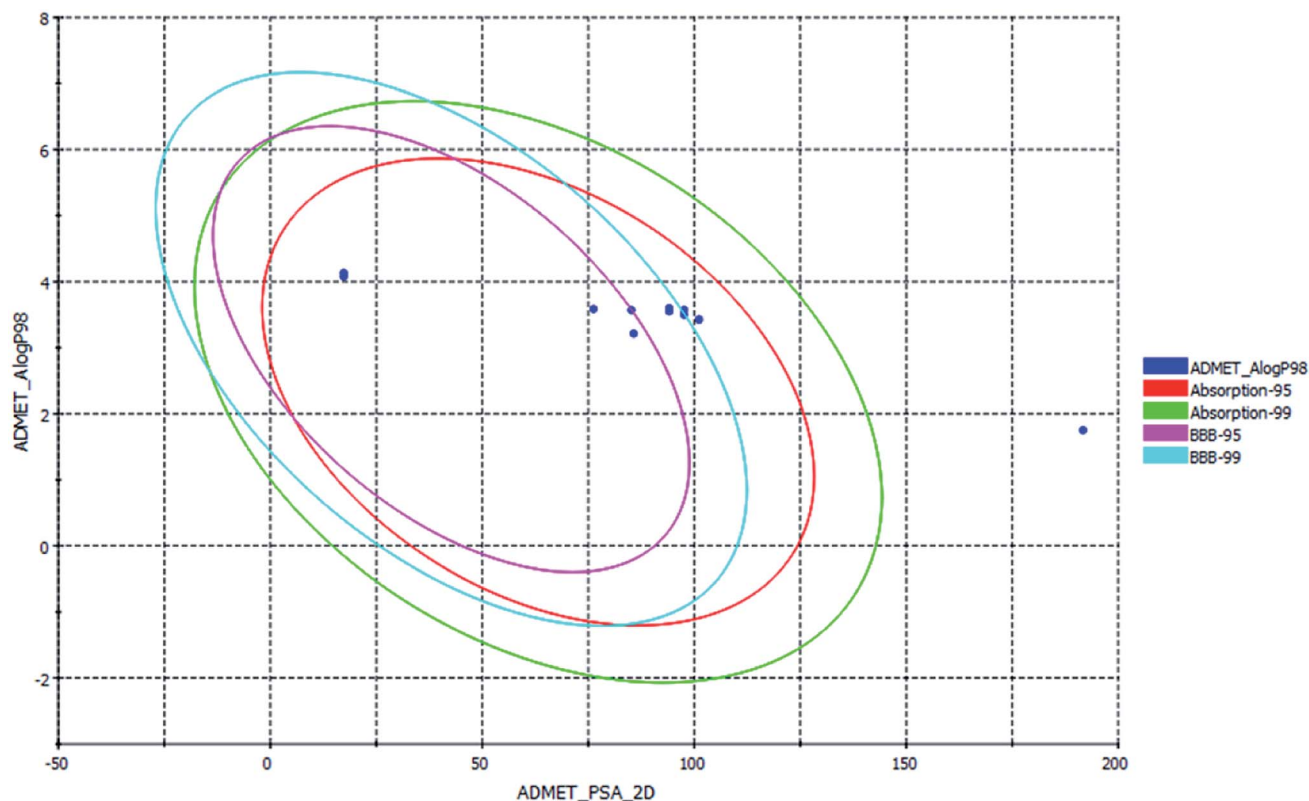


Fig. 7 ADMET plot for curcuminoids, indicating their ideal properties.

Table 1 *In silico* ADMET properties of curcuminoids

Compound name	Solubility	BBB	CPY2D6	Hepatotoxic	HIA	Alog $p_{98}$	PSA
Keto-curcumin	3	3	NI	NT	0	3.554	94.092
Dihydrocurcumin	3	3	NI	NT	0	3.577	94.092
Octahydrocurcumin	3	4	NI	NT	0	3.427	101.12
Tetrahydrocurcumin	3	3	NI	NT	0	3.6	94.092
Alpha-turmerone	2	0	NI	NT	0	4.079	17.3
Beta-turmerone	2	0	NI	NT	0	4.133	17.3
Bisdemethoxycurcumin	3	2	NI	NT	0	3.587	76.232
Hexahydrocurcumin	3	3	NI	NT	0	3.5	97.607
Demethoxycurcumin	3	2	NI	NT	0	3.57	85.162
Cyclocurcumin	3	2	NI	NT	0	3.213	85.722
Curcumin glucuronide	3	4	NI	NT	3	1.752	191.7
Enol-curcumin	3	4	NI	NT	0	3.573	97.607
Ar-turmerone	2	0	NI	NT	0	4.335	17.3
Hexahydroxycurcuminol	3	4	NI	NT	0	3.427	101.12





Table 2 TOPKAT analysis of curcuminoids

Compound name	NTP carcinogen <sup>a</sup>	Ames mutagen <sup>a</sup>	Rat TD50 <sup>b</sup>	Rat oral LD50 <sup>c</sup>	Rat inhalation LC50 <sup>d</sup>	Fathead minnow LC50 <sup>d</sup>	<i>Daphnia</i> EC50 <sup>e</sup>
Keto-curcumin	NC	NM	57.8227	2.81353	1200.80	0.000386	1.61229
Hexahydroxycurcuminol	NC	NM	1.53493	14.2048	1538.66	0.005549	0.481126
Dihydrocurcumin	NC	NM	63.4126	8.19582	1048.43	0.000318	1.05586
Octahydrocurcumin	NC	NM	1.53493	14.2048	1538.66	0.005549	0.481126
Tetrahydrocurcumin	NC	NM	25.0721	11.8342	1126.18	0.000875	0.600233
Alpha-turmerone	NC	NM	72.5357	1.09783	23297.40	0.000578	8.87751
Beta-turmerone	NC	NM	55.48	1.38263	26 058	0.000136	6.67372
Bisdemethoxycurcumin	NC	NM	168.944	1.00939	657.833	0.000988	3.7944
Hexahydrocurcumin	NC	NM	1.13932	8.65315	1963.47	0.004877	1.09273
Demethoxycurcumin	NC	NM	58.7615	2.39167	1645.26	0.000454	3.67163
Cyclocurcumin	NC	NM	24.1476	1.76033	3601.59	0.001092	1.03814
Curcumin glucuronide	NC	NM	1.42569	8.725	74.3232	0.004108	0.687055
Ar-turmerone	NC	NM	100.468	1.43641	16932.10	0.001595	26.6017
Enol-curcumin	NC	NM	53.0906	3.39621	927.783	0.000536	0.391523

<sup>a</sup> Non-Carcinogen (NC), Non-Mutagen (NM). <sup>b</sup> TD50 (mg per kg body weight per day). <sup>c</sup> LD50 (g per kg body weight), mg m<sup>-3</sup> h<sup>-1</sup>. <sup>d</sup> LC50 (g L<sup>-1</sup>). <sup>e</sup> EC50 (mg L<sup>-1</sup>).

observed that compound 71315012 (curcumin glucuronide) is outside of the ellipses due to its low human intestinal absorption (HIA) and undefined BBB penetration, whereas some molecules such as 5318039 (hexahydrocurcumin), 11068834 (octahydrocurcumin) and hexahydrocurcumin are near to 99% confidence ellipses of the BBB, and the remaining compounds are inside 95% and 99% confidence ellipses of HIA. Also, the curcuminoids compounds exhibit good membrane permeability, except curcumin glucuronide, which has a polar surface area<sup>27</sup> (PSA) of <140 Å<sup>2</sup>. Overall, the curcuminoids compounds are non-toxic to hepatic cells, non-inhibitors of a metabolic enzyme (CPYD26), and exhibit very high to medium penetration across the BBB except a few molecules (Table 1).

Similarly, the *in silico* pharmacodynamic models of the curcumin derivatives are free from carcinogens and mutagens. Quantitative-structure toxicity relationship (QSTR)-based toxicity parameters such as rat (TD50), rat\_oral (LD50),

rat\_inhalation, fathead minnow (LC50) and *Daphnia* (EC50) are summarized in Table 2. With only one RO5 violation allowed, Alog *P* of ≤5, molecular weight of ≤500, hydrogen bond acceptors (HBA) of ≤10 and hydrogen bond donor (HBD) of ≤5 are also known as Lipinski's rule of 5 or Pfizer's rule of five, which is used to evaluate the drug-likeness property of lead molecules.<sup>28</sup> Compounds that conform to these RO5 rules may be pharmacologically and biologically active, in addition to Veber's rule of rotatable bonds (RB) of <10 and PSA of <140 Å<sup>2</sup> (ref. 29) for orally bioactive compounds. Accordingly, all the compounds except curcumin glucuronide follow both Lipinski's and Veber's rule. Overall, 13 compounds possess oral bioavailability by obeying the drug-likeness rule tabulated in (Table 3).

## 2.4 Biological significance of receptor–ligand docking

Molecular docking is a key tool to understand the mode of binding of a compound to the active site of the target proteins.

Table 3 Drug-likeness property of curcuminoids

Compound name	PubChem CID	Lipinski's rule of 5				Veber's rule	
		HBD	HBA	MW	Alog <i>P</i>	RB	PSA
Keto-curcumin	969516	6	2	368.38	3.554	8	93.06
Dihydrocurcumin	10429233	6	2	370.396	3.577	9	93.06
Octahydrocurcumin	11068834	6	4	376.443	3.427	10	99.38
Tetrahydrocurcumin	124072	6	2	372.412	3.6	10	93.06
Alpha-turmerone	14632996	1	0	218.335	4.079	4	17.07
Beta-turmerone	196216	1	0	218.335	4.133	4	17.07
Bisdemethoxycurcumin	5315472	4	2	308.328	3.587	6	74.6
Hexahydrocurcumin	5318039	6	3	374.428	3.5	10	96.22
Demethoxycurcumin	5469424	5	2	338.354	3.57	7	83.83
Cyclocurcumin	69879809	6	2	368.38	3.213	5	85.22
Curcumin glucuronide	71315012	12	5	544.504	1.752	11	189.28
Enol-curcumin	381330244	6	3	368.38	3.573	7	96.22
Ar-turmerone	558221	1	0	216.319	4.335	4	17.07
Hexahydroxycurcuminol	Sketched	6	4	376.443	3.427	10	99.38



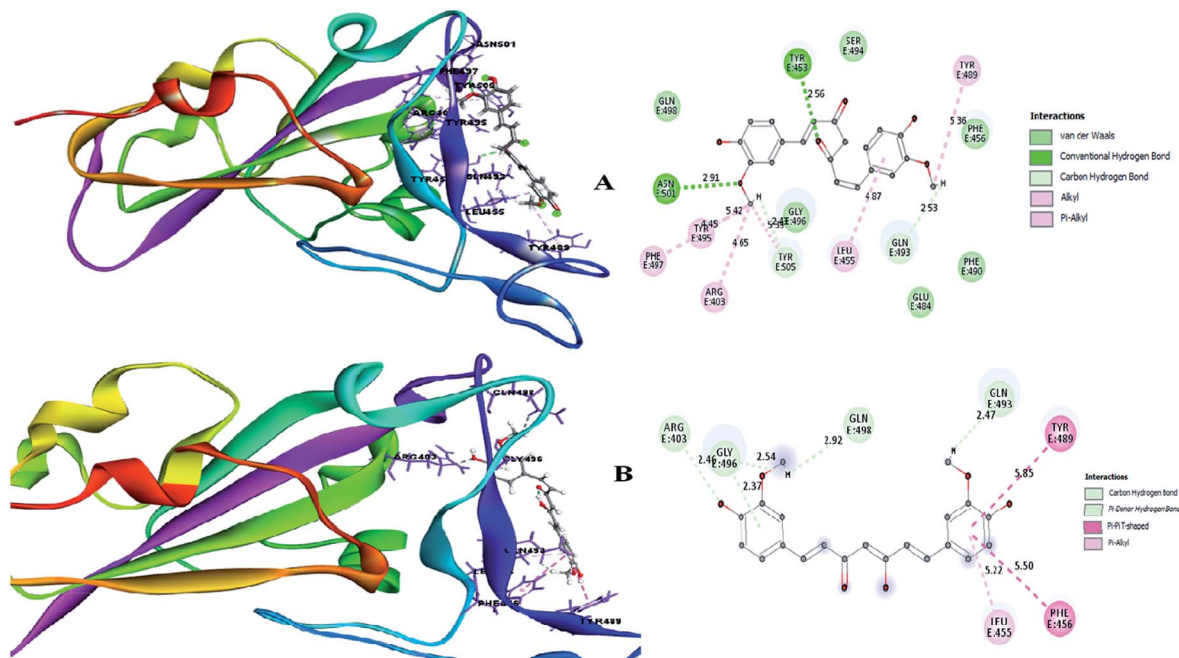


Fig. 8 (A) Receptor–ligand interaction of S-glycoprotein (RBD) with the keto form of curcumin and (B) receptor–ligand interaction of S-glycoprotein (RBD) with the enol form of curcumin.

Two cell surface proteins are the main drug targets of novel SARS-CoV-2. A team of researchers working on the genus Alpha/Betacoronavirus has reported that each virus species uses different entry points to invade the human system, where the DPP4 and APN receptors are used by MERS-CoV<sup>30</sup> and HCoV-229E, respectively.<sup>31</sup> Similarly, the spike glycoprotein or surface glycoprotein or S-glycoprotein of SARS-CoV-2 and SARS-CoV uses the angiotensin-converting enzyme 2 (ACE2) receptor, a type-I transmembrane metallopeptidase, for cellular

entry.<sup>32,33</sup> It was recently reported that SARS-CoV-2 harbors the F486, N487, Y489, Q493, Q498, T500, N501, G502 and Y505 amino acid residues present in the receptor-binding motif (RBM) of the receptor-binding domain (RBD) of the S-glycoprotein to interact with the ACE2 receptor in humans.<sup>1</sup>

Therefore, the present study aimed to find the binding interaction of curcuminoid derivatives to key target residues of SARS-CoV-2 through docking studies using the CDocker algorithm, a grid-based CHARMM simulation tool. Among the

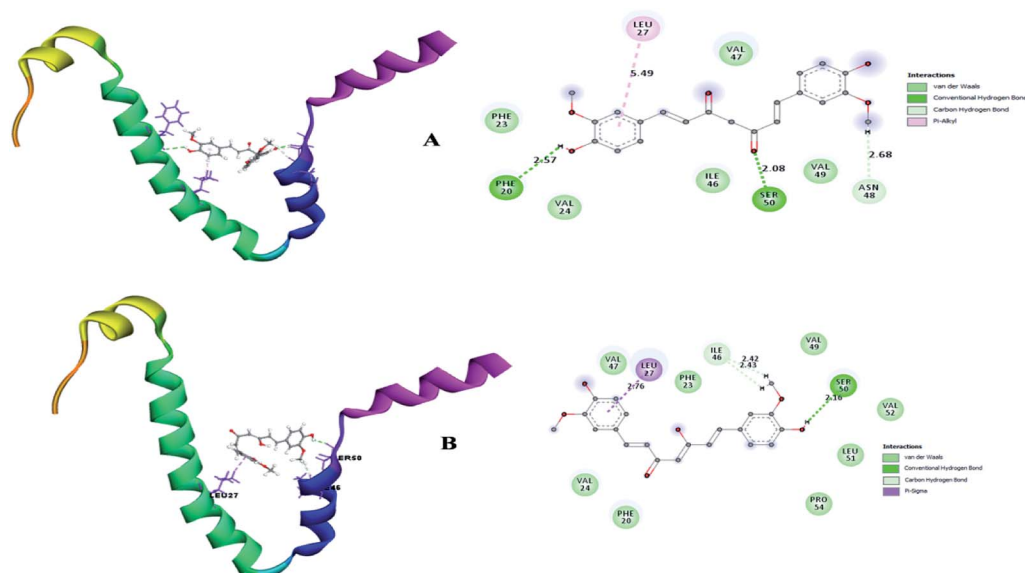


Fig. 9 (A) Receptor–ligand interaction of S envelope protein with the keto form of curcumin and (B) receptor–ligand interaction of the envelope protein with the enol form of curcumin.





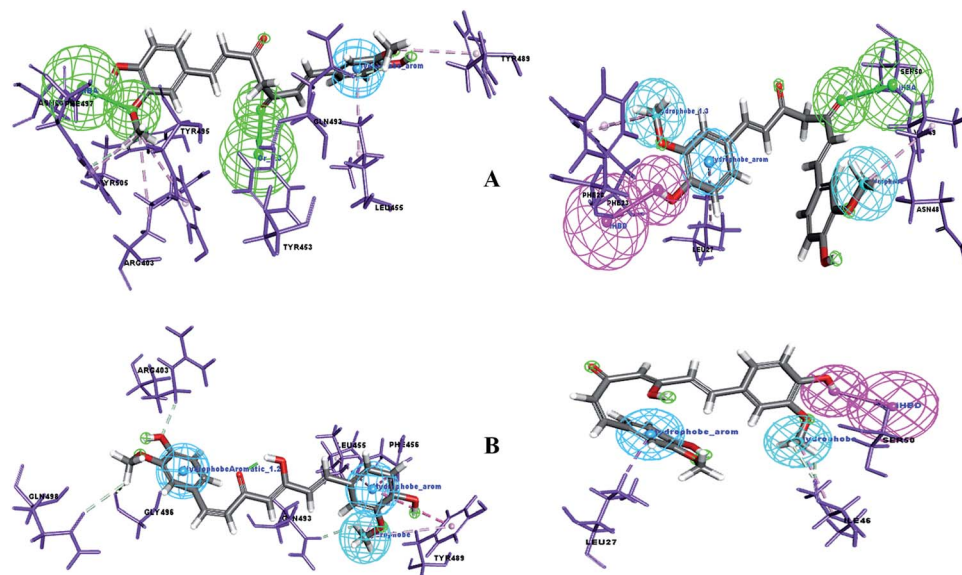


Fig. 10 (A) Pharmacophore interaction of drug targets with the keto form of curcumin and (B) pharmacophore interaction of drug targets with the enol form of curcumin. Pharmacophore features green: Hydrogen Bond Acceptor (HBA), pink: Hydrogen Bond Donor (HBD), dark blue: hydrophobe aromatic, and cyan: hydrophobe.

14 compounds docked to the site of RBM, only two forms of curcumin, namely, its keto and enol forms, interact to form strong hydrogen bond interaction with the key residues Q493, N501, Y505, Y489, and Q498 with the CDocker docking score of  $-20.753 \text{ kcal mol}^{-1}$  and  $-16.8067 \text{ kcal mol}^{-1}$  (Fig. 8A and B), respectively. Thus, the receptor–ligand interaction study indicated that both forms of the curcumin pharmacophore have the ability to interact with the spike glycoprotein, anchoring the residues of ACE2. Also, blocking these residues possibly does not facilitate its interaction with the human ACE2 receptor, and thereby viral infection can be controlled.

A small envelope protein (E-protein) is another drug target protein that plays three functional roles in SARS-CoV-2, including viral assembly,<sup>34</sup> pathogenesis of the virus<sup>35</sup> and release of virions.<sup>36</sup> Targeting the E-protein functions aborts the viral assembly process, and consequently the formation of immature virions.<sup>8</sup> On contrary, the exact function of the envelope protein is still enigmatic because various studies have shown that even without E-protein, the virus utilizes accessory proteins to form its core structure, but the virus efficiency is reduced by a hundred-to a thousand-fold during morphogenesis.<sup>37</sup> During docking, it was observed that curcumin keto ( $-19.174 \text{ kcal mol}^{-1}$ ) and enol

( $-14.115 \text{ kcal mol}^{-1}$ ) interact with the TM alpha helix of the E-protein to form hydrogen and hydrophobic interactions (Fig. 9A and B), respectively. Thus, curcumin may be a candidate compound for treating SARS-CoV-2.

## 2.5 Receptor–ligand interaction pharmacophore generation for SAR analysis of curcumin derivatives

A pharmacophore is a key tool to derive a structure–activity relationship (SAR) to distinguish the active functional groups of chemical compounds that are responsible for their biological activity. Curcumin is already known for its pleiotropism of pharmacological activity.<sup>38</sup> Previously, the remarkable effect of dietary phenolic curcumin (1,3-dicarbonyl group) was reported for HIV drug targets, including integrase, Tat-mediated trans-activation of the HIV-LTR and protease.<sup>39</sup> In the current study, SAR was probed for two forms of curcumin, *i.e.* its keto form (1,3-dicarbonyl group) and enol form (1,3-keto–enol). To identify the probable functional moieties of curcuminoids, the receptor–ligand interaction complex was used to enumerate the pharmacophore. The pharmacophore interaction features generated for keto and enol curcumin are depicted in Fig. 10. It can clearly be observed that the keto form of curcumin shares

Table 4 Dynamics and binding energy parameters of the receptor–ligand complex

Energy parameters in ( $\text{kcal mol}^{-1}$ )	S-glycoprotein		Envelope protein	
	Keto	Enol	Keto	Enol
Complex energy	−7378.69	−7350.24	1105.8	1129.009
Complex entropy	−29.9815	−29.9808	−28.1418	−28.1414
Binding energy	−56.57	−50.42	−51.71	−48.66
Potential energy	−10902.4	−10931.5	−2578.60	−2658.74
Total energy	−8630.45	−8679.34	−1741.24	−1725.18
Kinetic energy	2271.98	2252.18	837.431	833.556



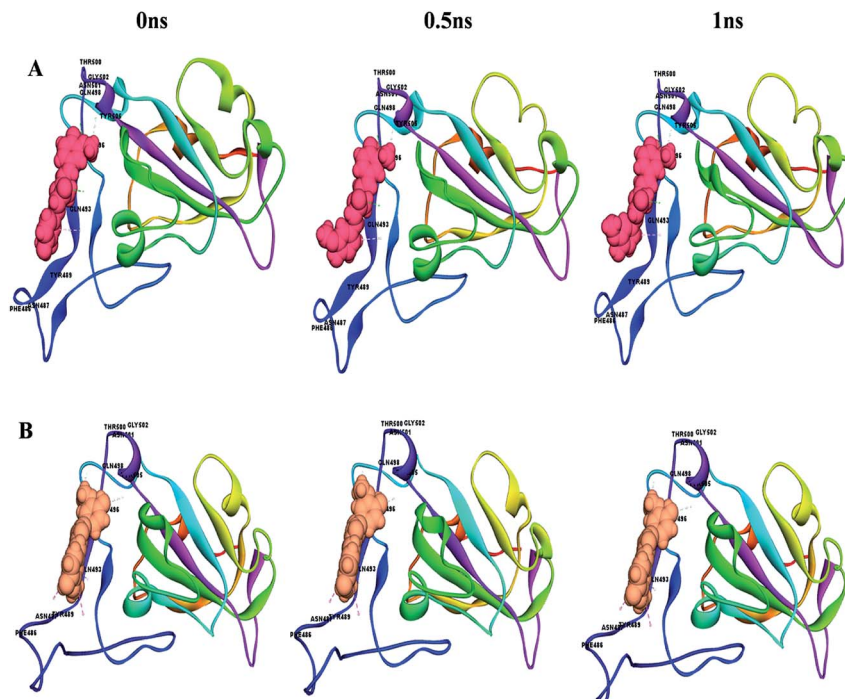


Fig. 11 Conformational changes with time for (A) S-glycoprotein with the keto form of curcumin and (B) S-glycoprotein with the enol form of curcumin.

a common HBA feature ( $=O$ ) with two drug target proteins and no acceptor feature was mapped for enol curcumin, which contains  $=O$  and  $OH$  chemical groups (1,3-keto-enol moiety).

Uniquely, the keto form of curcumin containing the  $=O$  chemical group showed a tendency to bind with the active site amino acid in both drug targets. In contrast, in the enol form,

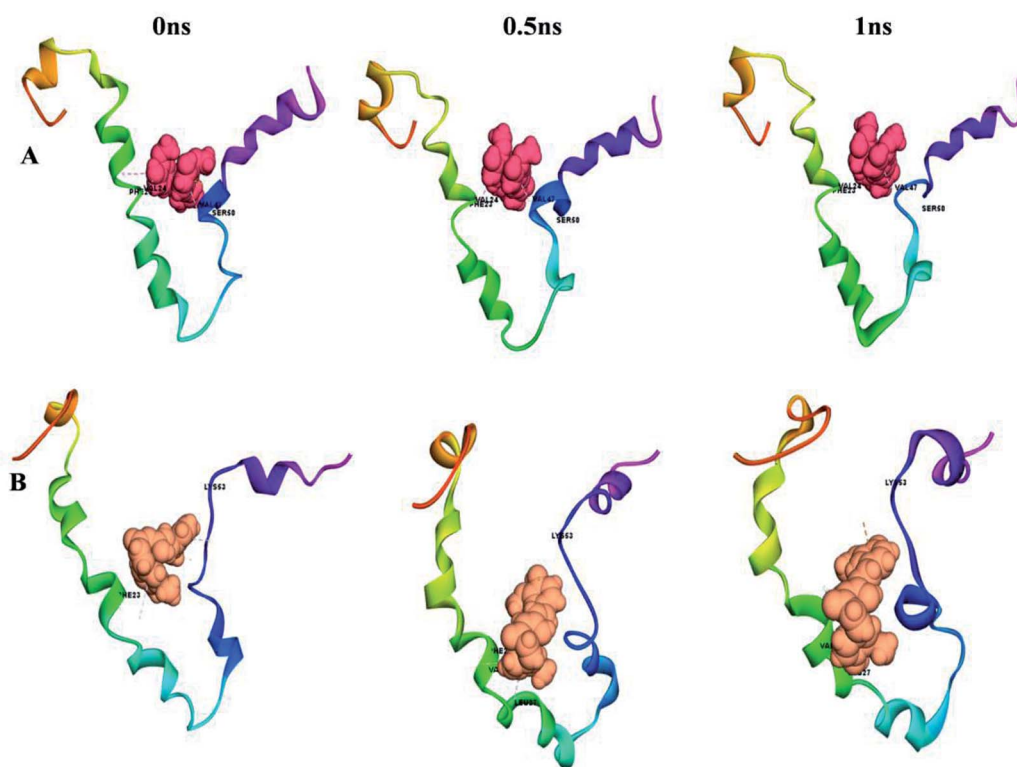


Fig. 12 Conformational changes with time for (A) envelope protein with the keto form of curcumin and (B) envelope protein with the enol form of curcumin.



the =O group is replaced with OH, which showed no binding interaction with the receptor sites. This reveals that the keto groups attached to the curcumin structure can be considered the important and crucial functional pharmacophore. Similarly, the enol form shares common pharmacophore hydrophobe aromatics and hydrophobes with the drug targets. Furthermore, to confirm this, the significance of substructure elimination of the  $-OCH_3$  moiety was studied. Accordingly, the removal of the  $OCH_3$  moiety in the curcumins showed no docking or interaction pharmacophore generation, which implies that together with the main core of curcumin, the  $OCH_3$  functional group is also required for binding interaction with the drug targets of SARS-CoV-2.

## 2.6 Binding energy and stability of the complex

Understanding the affinity of a compound to its target protein/receptors is the main objective of the structure-based drug design and drug discovery process. Specifically, the binding of a lead molecule to a receptor or signalling protein may alter its biological activity, and thus molecular structure recognition is considered a fundamental component in the virtual screening of drugs.

This can easily be achieved by means of calculating the binding energy of a docked receptor–ligand complex. The binding energy of the S-glycoprotein and envelope protein with the keto and enol forms of curcumin was computed using the equation ( $\text{energy}_{\text{binding}} = \text{energy}_{\text{complex}} - (\text{energy}_{\text{ligand}} + \text{energy}_{\text{receptor}})$ ),<sup>40</sup> where the negative energy of the binding complex shows the strength of the protein–ligand interaction (Table 4). A molecular dynamics study was carried out to study the stability of the receptor–ligand interaction. At the end of the study, binding site conformational changes were observed for the SARS-CoV-2 drug target proteins. The S-glycoprotein and envelope protein bound with the keto form of curcumin showed less secondary structural conformational changes compared to that with the enol form (Fig. 11 and 12), respectively. Further, significant flexibility and departure from the initial structure from molecular dynamics were estimated using the root mean square deviation (RMSD).

According to Fig. 13A, it can be observed that the RMSD deviation of the S-glycoprotein\_keto deviations is within 2.1 Å, also the stability of the complex was maintained throughout the molecular dynamics study. In contrast, the S-glycoprotein\_enol complex gradually changed to maintain its structural stability,

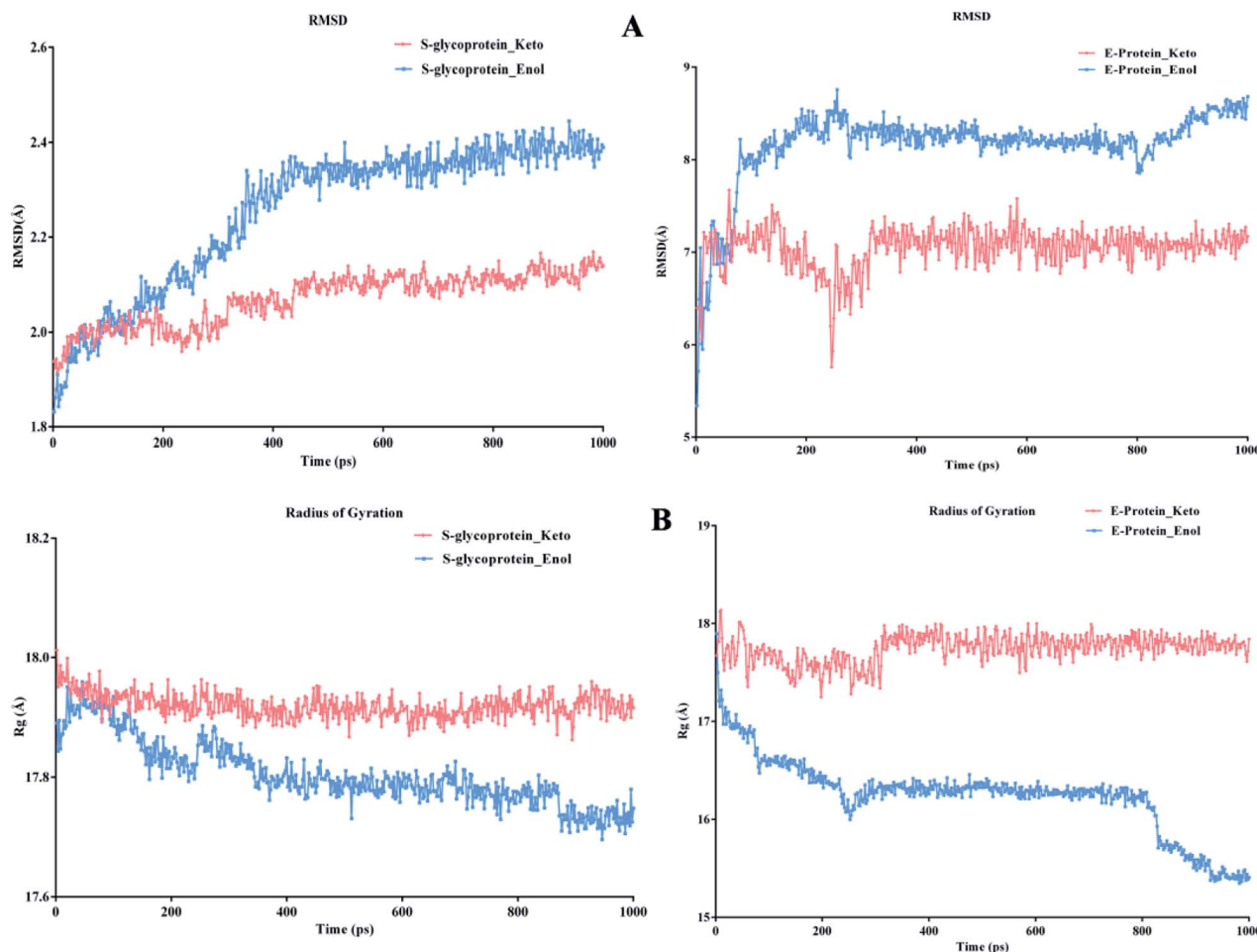


Fig. 13 (A) RMSD graph for S-glycoprotein and envelope protein protein–ligand complex and (B)  $R_g$  graph for S-glycoprotein and envelope protein protein–ligand complex.



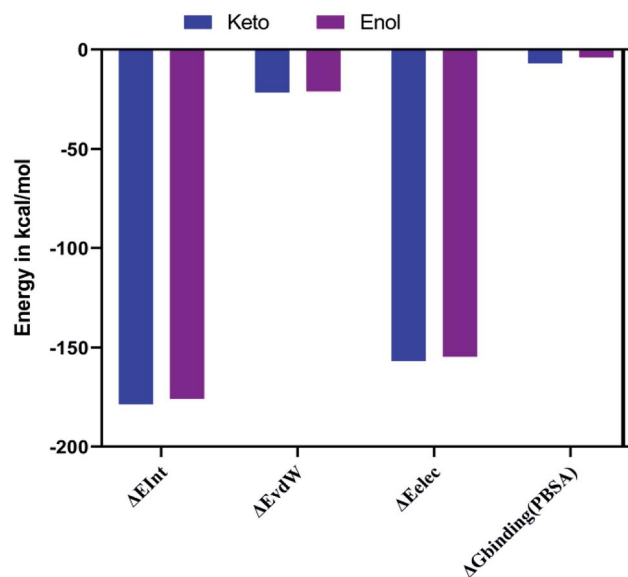


Fig. 14 Graph showing the interaction ( $\Delta E_{int}$ ), van der Waals ( $\Delta E_{vdW}$ ), electrostatic ( $\Delta E_{elec}$ ), and free energy binding ( $\Delta G_{binding}$ ) for both the keto-curcumin-S-protein RBD and enol-curcumin-S-protein RBD.

but the deviation did not exceed 2.5 Å. Subsequently, the conformational fluctuation of the envelope protein with the keto and enol curcumin complexes showed a sharp peak at the beginning of the molecular dynamics. Conversely, after 0.2 ns, the conformational changes of both curcumin complexes were less up to the end of the molecular dynamics study. Another time-dependent analysis is the radius of gyration ( $R_g$ ), which is used to measure the compactness of proteins. Fig. 13B illustrates that the S-glycoprotein complexes with the keto and enol form showed negligible changes in protein structural folding, with  $R_g$  values in the range of 17.6 to 18. In contrast, the  $R_g$  value

of the envelope protein with the enol form of curcumin changed over time, but at the end of the study, the structural compactness of the protein was retained, although the envelope protein complex with the keto form of curcumin showed less fluctuation and was stable throughout the molecular dynamics study. Thus, overall, the complex stability is relatively higher for the keto-bound complex than the enol complex.

## 2.7 Free and interaction energy parameter analysis for receptor–ligand complex

The combination approach of molecular mechanics energies with Poisson–Boltzmann surface area (MM/PBSA) is used to estimate the free energy of the binding of small ligands to biological macromolecules. The binding free energies for the two keto-curcumin-S-protein RBD and enol-curcumin-S-protein RBD complexes were estimated using the MM-PBSA method. The calculated binding free energy was  $-4.067$  kcal mol $^{-1}$  for enol-curcumin. Dongling *et al.*<sup>41</sup> described that the favorable interaction of a complex can be measured in terms of electrostatic interaction. The analysis of the free energy components showed that the binding free energy of  $-7.006$  kcal mol $^{-1}$  and more favorable electrostatic interaction (Fig. 14) for the keto-curcumin-S-protein RBD complex are the main reason for the higher affinity of keto-curcumin compared to enol-curcumin. On the other hand, the van der Waals interaction is more favorable for both the keto-curcumin and enol-curcumin conformations by  $-21.671$  kcal mol $^{-1}$  and  $-21.168$  kcal mol $^{-1}$ , respectively. Finally, the interaction energy of the complex was calculated using the CHARMM energy, which exhibited a 0.98-fold change difference for keto to enol curcumin. Comparatively, all the energy values are more favorable for both forms of curcumin with slight fold changes in energy values.

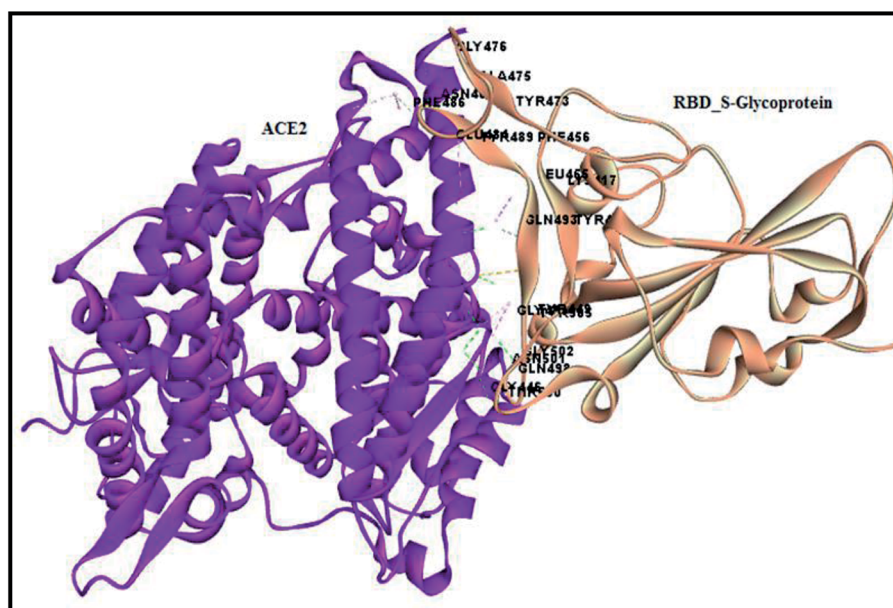


Fig. 15 Analyzed interface residues of ACE2 RBD region of S-glycoprotein in SARS-COV-2.





### 3. Conclusion

SARS-CoV-2 has resulted in a devastating pandemic with global concern; however, present therapies in virology fail to prevent its effects. Currently, there is exigency in identifying novel leads with anti-viral properties to impede viral pathogenesis in the host system. Thus, two important curcuminoids of turmeric, *i.e.*, its curcumin keto and enol forms, were demonstrated to be complementary to bind with the S-glycoprotein and envelope protein of SARS-CoV-2. However, the keto form of curcumin is more favourable for both these drug targets considering its docking score, binding energy and molecular dynamics simulation. Thus, this study indicates that surface proteins are key drug target proteins of SARS-CoV-2, and probably curcumin blocks essential biologically active drug target residues, thereby attenuating the viral infection. Hence, this computational biology approach identifies curcumin as a drug candidate for further investigation in treating SARS-CoV-2. However, this was an initial study to identify the active pharmacophore of the compound and its binding site structural complementary for two drug targets. Thus, in the future, we aim to perform large-scale molecular dynamics, and *in vitro* and *in vivo* experiments to confirm the efficacy of curcumin against SARS-CoV-2.

### 4. Materials and methods

#### 4.1 Homology modelling

The coronavirus protein sequences reported in the Indian State of Kerala such as the envelope protein (QHS34548), membrane protein (QIA98586) and spike glycoprotein\_RBD region (QIA98583) structures not reported in the PDB databases were determined using the template-based modelling technique. All the sequences were downloaded from the National Genomics Data Center (<https://bigd.big.ac.cn/ncov/>). Two approaches of a basic local alignment search tool for protein (BLAST\_p) were performed, where one was a direct hit on a server (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and the other technique was through the BLAST\_NCBI search protocol in Biovia DS2019. The template structures were selected by considering maximum identity and query coverage with less positive and *E*-values, where template hits with SARS-CoV/MERS-CoV species were given preference. Query sequences were aligned with the template structure followed by build homology modelling using MODELLER 9.17v9.<sup>42</sup>

#### 4.2 Modelled structure validation

The homology model protein was subjected to quality assessment in various structural validation servers and tools. In Discovery Studio, the best model structure was selected based on three parameters, including superimposing the best model structures with the PDB template structure to calculate the root mean square deviation using the align and superimpose protein protocol. Statistically, a lower DOPE score and high PDF total energy represent the best quality of model structures with stable conformations. Besides, external web servers such as ERRAT (<https://servicesn.mbi.ucla.edu/ERRAT/>), Procheck (<https://servicesn.mbi.ucla.edu/PROCHECK/>), and Rampage (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) were used to assess the structure quality. Some structures that showed more outliers in the Ramachandran plot were further optimized by the energy minimization technique provided with an RMS gradient of 0.1 kcal mol<sup>-1</sup> Å<sup>-2</sup> to prevent bad steric contact of atoms.<sup>43</sup> Besides, a deficiency in side chain and loop amino acids was processed using side and loop refinement by CHARMM simulation and force field.<sup>44</sup> The best-validated structures were used for further docking analysis.

servicesn.mbi.ucla.edu/PROCHECK/), and Rampage (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) were used to assess the structure quality. Some structures that showed more outliers in the Ramachandran plot were further optimized by the energy minimization technique provided with an RMS gradient of 0.1 kcal mol<sup>-1</sup> Å<sup>-2</sup> to prevent bad steric contact of atoms.<sup>43</sup> Besides, a deficiency in side chain and loop amino acids was processed using side and loop refinement by CHARMM simulation and force field.<sup>44</sup> The best-validated structures were used for further docking analysis.

#### 4.3 Chemical structure preparation

The natural form of 14 curcumin derivative compounds were retrieved from the PubChem compound and substance database (<https://pubchem.ncbi.nlm.nih.gov/>). All compounds were prepared using the ligand protocol to generate various 3D conformations, isomers, and tautomers of the compounds to remove duplicity and to fix bad valences.

#### 4.4 ADMET, drug-likeness and toxicity predictions

The curcumin derivative was submitted to the ADMET and TOPKAT tools of small molecule protocol for the *in silico* pharmacokinetics and pharmacodynamic studies. The pK<sub>a</sub> study included parameters such as human intestinal absorption (HIA), aqueous solubility, blood-brain-barrier penetration (BBB), cytochrome CYP2D6 inhibition, plasma protein binding (PPB) and hepatotoxicity. In the pharmacodynamic study, we included animal models such as NTP rodent carcinogenicity, rat\_oral LD50, rat\_TD50, fathead minnow LC50, *Daphnia* EC50, rat inhalational LC50 and Ames mutagenicity. All these models were developed and validated based on a quantitative-structure toxicity relationship (QSTR).

#### 4.5 Molecular docking and binding energy

Molecular docking is a lock and key process to identify compounds with complementary structures to drug target proteins. The CDOCKER algorithm is a grid-based and molecular dynamic simulation-implemented docking protocol, which was utilized for docking compounds to the binding site of the protein. The spike glycoprotein and envelope proteins are two novel drug targets for the dreadful covid-19 virus, which are the binding sites identified using a receptor-based cavity tool in Discovery Studio 2019. However, the spike glycoprotein (S-glycoprotein) RBD region binding to ACE2 receptor interactions were analyzed using the analyze protein interface tool in macromolecules to spot hydrogen, hydrophobic and other interactions, which was considered as the binding site for compound docking (Fig. 15). Both proteins were placed in an equal grid spacing in right angles in the 3D direction of the input site sphere coordinates with the radius of -38.36X, 31.32Y, 3.33Z, 20.1 Å and 37.37X, 33.46Y, 17.81Z, 26.873 Å for the RBD region of the S-glycoprotein and envelope protein, respectively. The docked complexes were further analyzed and validated *via* the negative CDOCKER energy and receptor-ligand interactions. Further, to understand and quantify the strength of the interactions between a ligand and protein, the



binding energies between all the docked poses of the ligand with receptor were studied by utilizing the binding energy protocol with CHARMM. This is a crucial process in drug discovery and lead optimization. The best pose with the lowest binding energy was considered for the molecular dynamics simulation.

#### 4.6 Molecular dynamics and simulation

The top compounds with the best pose in receptor–ligand interaction for the two different drug target proteins after docking validation were selected for molecular dynamics simulation. The whole system was subjected to CHARMM force field to satisfy bonded and non-bonded interactions, followed by standard dynamics cascade in a five-step simulation protocol. Initially, two 500 steps of minimization were performed using the steepest descent and conjugate gradient.<sup>45</sup> The minimized complexes were gradually driven from 50 K to a final target temperature of 300 K, followed by equilibration simulations. Finally, production was performed for 1000 ps for four complexes. The leapfrog dynamics integrator and shake constraint were introduced throughout the molecular dynamics simulation to study the bonded and non-bonded interaction. The molecular dynamics trajectory was determined for complex stability and time-dependent analysis (RMSD and radius of gyration ( $R_g$ )) using the Biovia Discovery Studio 2019 analyze trajectory protocol (Dassault Systèmes).<sup>46</sup>

#### 4.7 CHARMM interaction energy and molecular mechanics-Poisson–Boltzmann surface area (MM-PBSA) analysis

This study was focused on the impact of curcuminoids on the stability of the S-protein. Thus, the complex structure of keto-curcumin-S-protein RBD and enol-curcumin-S-protein RBD was used as a starting point for calculating the binding free energies. The 1000 ps molecular dynamics simulation was carried out using CHARMM. In CHARMM, molecular dynamics simulations are performed using a classical mechanics approach, in which Newton's equations of motion are integrated for all atoms in the system.<sup>47</sup> Thus, for each MD-simulated complex, we calculated the free binding energy ( $\Delta G_{\text{binding}}$ ) values for the 500 conformations of the MD trajectory. During the simulation, one conformation or snapshot was saved every 2 ps up to 1000 ps, and the final  $\Delta G_{\text{binding}}$  is the average of 500 conformations of the receptor–ligand complex. In the MM-PBSA method, the free energy of the protein–ligand binding ( $\Delta G_{\text{binding}}$ ) is obtained from the difference between the free energies of the protein–ligand complex ( $G_{\text{complex}}$ ) and the unbound receptor/protein ( $G_{\text{protein}}$ ) and ligand ( $G_{\text{ligand}}$ ) as follows:

$$\Delta G_{\text{binding}} = \Delta G_{\text{complex}} - [\Delta G_{\text{protein}} + \Delta G_{\text{ligand}}]$$

The output conformations of the molecular dynamics simulation were further sampled to study the interaction energy between sets of atoms across all conformations using CHARMM. The interface interacting residues of the receptor

and ligand were selected as two atom sets together with a cut-off distance of 12–10 Å to calculate the non-bonded interactions as follows:

$$\Delta E_{\text{Int}} = \Delta E_{\text{vdW}} + \Delta E_{\text{ele}}$$

where,  $\Delta E_{\text{Int}}$  = interaction energy of ligand–protein,  $\Delta E_{\text{vdW}}$  = van der Waals energy of ligand–protein, and  $\Delta E_{\text{elec}}$  = electrostatic energy of ligand–protein.

## Funding

This research project has not received any funding from any agencies.

## Conflicts of interest

The authors confirm that this article content has no conflicts of interest.

## Acknowledgements

We would like to thank the management of JSSAHER, Mysuru for providing facilities to carry out the research.

## References

- 1 M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen and M. A. Müller, *Cell*, 2020, **181**(2), 271–280.
- 2 C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu and Z. Cheng, *Lancet*, 2020, **395**(10223), 497–506.
- 3 N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei and T. Yu, *Lancet*, 2020, **395**(10223), 507–513.
- 4 Y. Liu, A. A. Gayle, A. Wilder-Smith and J. Rocklöv, *J. Travel Med.*, 2020, **27**, 1–7.
- 5 F. Li, *Annu. Rev. Virol.*, 2016, **3**, 237–261.
- 6 S. A. Kopecky-Bromberg, L. Martínez-Sobrido, M. Frieman, R. A. Baric and P. Palese, *J. Virol.*, 2007, **81**(2), 548–557.
- 7 W. Spaan, D. Cavanagh and M. C. Horzinek, *J. Gen. Virol.*, 1988, **69**(12), 2939–2952.
- 8 D. Schoeman and B. C. Fielding, *Virology*, 2019, **16**(1), 69.
- 9 Z. Florian and M. E. Rebeaud, *Front. Cardiovasc. Med.*, 2020, **7**, 71.
- 10 D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C. L. Hsieh, O. Abiona, *et al.*, *Science*, 2020, **6483**(367), 1260–1263.
- 11 H. Zhang, J. M. Penninger, Y. Li, *et al.*, *Intensive Care Med.*, 2020, **46**, 586–590.
- 12 S. Prasad and B. B. Aggarwal, *Herbal medicine: biomolecular and clinical aspects*, 2011, 13.
- 13 K. I. Priyadarsini, *Molecules*, 2014, **19**(12), 20091–20112.
- 14 C. C. Colpitts, L. Schang, H. Rachmawati, A. Frentzen, S. Pfaender, P. Behrendt and P. Meuleman, *Gut*, 2014, **63**(7), 1137–1149.



- 15 U. J. Eke-Okoro, R. B. Raffa, J. V. Pergolizzi Jr, F. Breve, R. Taylor Jr and NEMA Research Group, *J. Clin. Pharm. Ther.*, 2018, **43**(4), 460–466.
- 16 B. N. Waghela, A. Sharma, S. Dhumale, S. M. Pandey and C. Pathak, *PLoS One*, 2015, **10**(2), 1–8.
- 17 I. Chattopadhyay, K. Biswas, U. Bandyopadhyay and R. K. Banerjee, *Curr. Res.*, 2004, **87**, 44–53.
- 18 D. Praditya, L. Kirchhoff, J. Brüning, H. Rachmawati, J. Steinmann and E. Steinmann, *Front. Microbiol.*, 2019, **10**, 912–919.
- 19 A. Balasubramanian, R. Pilankatta, T. Teramoto, A. M. Sajith, E. N. Wulia, A. Kulkarni and R. Padmanabhan, *Antiviral Res.*, 2019, **162**, 71–78.
- 20 M. K. Kim, J. C. Park and Y. Chong, *Nat. Prod. Commun.*, 2012, **7**(1), 57–59.
- 21 B. C. Mounce, T. Cesaro, L. Carrau, T. Vallet and M. Vignuzzi, *Antiviral Res.*, 2017, **142**, 148–157.
- 22 J. Dandapat, A. B. Jena, N. Kanungo, V. Nayak and G. Chainy, *Insights from Computational study and implication for intervention*, 2020.
- 23 M. Y. Shen and A. Sali, *Protein Sci.*, 2006, **15**(11), 2507–2524.
- 24 M. W. MacArthur, R. A. Laskowski and J. M. Thornton, *Curr. Opin. Struct. Biol.*, 1994, **4**(5), 731–737.
- 25 G. Studer, M. Biasini and T. Schwede, *Bioinformatics*, 2014, **30**(17), i505–i511.
- 26 M. Wiederstein and M. J. Sippl, *Nucleic Acids Res.*, 2007, **35**, W407–W410.
- 27 R. Barret, Importance and Evaluation of the Polar Surface Area (PSA and TPSA), *Therapeutic Chemistry, Fundamentals*, 2018, pp. 89–95.
- 28 C. A. Lipinski, *Drug Discovery Today: Technol.*, 2004, **1**(4), 337–341.
- 29 D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, *J. Med. Chem.*, 2002, **45**(12), 2615–2623.
- 30 V. S. Raj, H. Mou, S. L. Smits, D. H. Dekkers, M. A. Müller, R. Dijkman and V. Thiel, *Nature*, 2013, **495**(7440), 251–254.
- 31 C. L. Yeager, R. A. Ashmun, R. K. Williams, C. B. Cardellicchio, L. H. Shapiro, A. T. Look and K. V. Holmes, *Nature*, 1992, **357**(6377), 420–422.
- 32 X. Y. Ge, J. L. Li, X. L. Yang, A. A. Chmura, G. Zhu, J. H. Epstein, Y. J. Zhang and Y. J. Zheng, *Nature*, 2013, **503**(7477), 535–538.
- 33 H. Hofmann, M. Geier, A. Marzi, M. Krumbiegel, M. Peipp, G. H. Fey and S. Pöhlmann, *Biochem. Biophys. Res. Commun.*, 2004, **319**(4), 1216–1221.
- 34 B. G. Hogue and C. E. Machamer, *Importance of Conserved Cysteine Residues in the Coronavirus Envelope Protein*, American Society of Microbiology, 2008, pp. 179–200.
- 35 J. M. Jimenez-Guardeno, J. L. Nieto-Torres, M. L. DeDiego, J. A. Regla-Nava, R. Fernandez-Delgado, C. Castaño-Rodriguez and L. Enjuanes, *PLoS Pathog.*, 2014, **10**(8), 1–9.
- 36 T. R. Ruch and C. E. Machamer, *J. Virol.*, 2011, **85**(2), 675–685.
- 37 M. L. DeDiego, E. Álvarez, F. Almazán, M. T. Rejas, E. Lamirande, A. Roberts and L. Enjuanes, *J. Virol.*, 2007, **81**(4), 1701–1713.
- 38 H. S. Zhang, Z. Ruan and W. W. Sang, *J. Cell. Physiol.*, 2011, **226**, 3385–3391.
- 39 A. Minassi, G. Sánchez-Duffhues, J. A. Collado, E. Muñoz and G. Appendino, *J. Nat. Prod.*, 2013, **76**(6), 1105–1112.
- 40 J. Tirado-Rives and W. L. Jorgensen, *J. Med. Chem.*, 2006, **49**(20), 5880–5884.
- 41 Z. Dongling, Y. Lei, J. Hanyong, G. Shanshan and H. Weiwei, *Int. J. Mol. Sci.*, 2014, **15**, 17284–17303.
- 42 B. Webb and A. Sali, *Protein structure modeling with MODELLER*, Humana Press, New York, NY, 2014, pp. 1–15.
- 43 V. Z. Spassov, P. K. Flook and L. Yan, *Protein Eng., Des. Sel.*, 2008, **21**(2), 91–100.
- 44 B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr, L. Nilsson, R. J. Petrella, B. Roux and A. Caflisch, *J. Comput. Chem.*, 2009, **30**(10), 1545–1614.
- 45 S. Dhivya, C. Suresh Kumar, V. Bommuraj, R. Janarthnam, M. Chandran, T. Usha and S. K. Middha, *J. Biomol. Struct. Dyn.*, 2018, **36**(9), 2361–2372.
- 46 S. Dhivya, N. Premjanu, M. Ganesh, J. A. Lakshmi, M. Sushil Kumar and C. Sureshkumar, in *Lipid Bilayers: Properties, Behavior and Interactions*, Nova Science Publisher, New York, 2019, Ch. 6, pp. 203–228.
- 47 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford, 1987.

