Identification of Potential Mpro Inhibitors for the Treatment of COVID-19 by Targeted Covalent Inhibition: An In Silico Approach

Dushyant V. Patel, Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, India
Divya M. Teli, Department of Pharmaceutical Chemistry, L. M. College of Pharmacy, India
Ashish M. Kanhed, Shobhaben Pratapbhai Patel School of Pharmacy and Technology Management, SVKM’s NMIMS University, India
Nirav R. Patel, Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, India
Bhavik S. Shah, Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, India
Amisha K. Vora, Shobhaben Pratapbhai Patel School of Pharmacy and Technology Management, SVKM’s NMIMS University, India
Mahesh T. Chhabria, Department of Pharmaceutical Chemistry, L. M. College of Pharmacy, India
Mange Ram Yadav, Parul University, India

ABSTRACT

A novel coronavirus is the causative agent identified for the current COVID-19 outbreak. Globally, more than 43 million people have been infected by this virus. The total number of deaths has surpassed 1.6 million across 210 countries due to the current pandemic. Till date, there is no specific therapeutic agent available for its treatment. Mpro, a non-structural protein cleaves viral polyproteins into other non-structural proteins. Inhibition of Mpro could prevent the virus replication projecting it as a potential candidate for anti-COVID-19 drug development. The authors report herein 10 top-ranked curcumin derivatives as non-peptide covalent-binding Mpro inhibitors using systematic virtual screening approach. Detailed ligand-receptor interaction analysis conferred that the α,β-unsaturated carbonyl moiety of curcumin functions as a warhead to yield a Michael adduct with Cys145 of the catalytic dyad of Mpro. Collectively, these results have offered new high affinity molecules for the development of potential drugs for the treatment of COVID-19.

KEYWORDS

Coronavirus, Covalent Docking, COVID-19, Curcumin, Mpro, Severe Acute Respiratory Syndrome

INTRODUCTION

Three major outbreaks of acute respiratory syndrome induced by coronaviruses (CoVs) have been witnessed in the last two decades. The first outbreak was Severe Acute Respiratory Syndrome (SARS) in 2003 with Guangdong, China as epicenter (Peiris et al., 2004) followed by Middle East Respiratory Syndrome (MERS) in 2012 in Saudi Arabia (Zaki et al., 2012) and now the novel coronavirus disease (COVID-19), first reported in Wuhan, China in late 2019(Wang et al., 2020). World Health Organization (WHO) declared COVID-19 outbreak as a global pandemic on 11th
March 2020 (Cucinotta & Vanelli, 2020). The spread of the disease around the world and the number of infected patients and mortality of COVID-19 are increasing exponentially day by day. As per the WHO report (28th October, 2020), COVID-19 is affecting more than 210 countries and territories with over 43,766,712 confirmed cases and over 1,663,459 total deaths around the world (World Health Organization, 2020). Currently, the global fatality rate is around 3.80% (calculated as deaths per confirmed cases). A large number of people are being identified as COVID positive every day in the USA followed by India, Brazil, Russia, South Africa and other countries.

CoVs are a group of enveloped, positive-sense, single-stranded RNA viruses belonging to the *Corona viridae* family. They induce respiratory, neurological and gastrointestinal complications of varying severity in human hosts. Novel coronavirus (2019-nCoV) also known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) belonging to the category of β-coronavirus is the causative agent of COVID-19 (Chen et al., 2020; Osman et al., 2020). Although SARS-CoV-2 is considered to be introduced from bats, its specific source, animal reservoir and enzootic patterns of transmission remain unresolved.

The SARS-CoV-2 genome, comprising of ~30,000 nucleotides encodes 4 structural proteins, 16 non-structural proteins and 9 accessory proteins (Chen et al., 2020). By translation of the viral genomic RNA (gRNA), CoVs produce two overlapping polyproteins pp1a (450-500 kDa) and pp1ab (750-800 kDa) (Thiel et al., 2003). These polyproteins undergo extensive proteolytic processing and ultimately the functional polypeptides are released which are crucial for the replication and assembly of the virus. This proteolytic processing is mediated predominately by the main protease (Mpro) also referred to as 3-chymotrypsin-like protease (3CLpro) or non-structural protein-5 (NSP-5), and by papain-like protease (PLpro). Mpro is a cysteine protease that digests the polyprotein within the Leu-Gln↓ (Ser, Ala, Gly) sequence (↓indicates the cleavage site), which appears to be a conserved pattern of this protease. The ability of CoVs to hydrolyze the peptide bond specifically after Gln residue is very unique which is unknown for human enzymes (Pillaiyar et al., 2016; Hilgenfeld, 2014). This characteristic feature along with the functional importance of Mpro makes it a potential target for COVID-19 antiviral drug discovery (Zhang, Lin, Sun, Rox, et al., 2020; Jin et al., 2020).

The X-ray crystallographic structure of SARS-CoV-2 Mpro bound to a covalent inhibitor N3 was resolved by Jin et al. (2020). Mpro has 306 amino acids long chain with three domains. Domain I contains Phe8 –Tyr101 residues, domain II contains Lys102 –Pro184 residues, and domain III contains Thr201 –Val303 amino acid sequence linked with domain II by a long loop region of Phe185 – Ile200 residues. The substrate-binding site with a Cys145 – His41 catalytic dyad is present in a cleft between domains I and II. The major active subsites where the substrates bind to Mpro are well defined. The S1 subsite is composed of Phe140, Leu141, Asn142, His163, Glu166 and His172 amino acids. A small S1’ subsite involves Thr25, Thr26 and Leu27 residues. Hydrophobic S2 subsite is composed of His41, Met49, Tyr54, Met165 and Asp187 residues. The S4 subsite is made up of Met165, Leu167, Phe185, Gln189 and Gln192 residues (Jin et al., 2020; Zhang, Lin, Sun, Curth, et al., 2020).

The Cys145 amino acid present in catalytic dyad functions as a common nucleophile in the proteolytic cleavage of the natural substrate of Mpro. The proteolytic cleavage is believed to be performed in a multiple-step mechanism (Figure 1). Once the Cys145 side-chain proton is abstracted by the imidazole nitrogen of His41 (Step-I), the resulting thiolate nucleophile attacks the carbonyl amide group of the natural substrate (Step-II). The N-terminal peptide product is released with the abstraction of a proton from His41 (Step-III). Then, the thioester is hydrolyzed (Step-V) and the C-terminal product is released which restores the active catalytic dyad (Step-VI) (Pillaiyar et al., 2016).

Covalent inhibitors hold an important place in the history of drug discovery, beginning in the late 19th century with aspirin and continuing with the current surge of rationally designed kinase inhibitors as anti-neoplastic agents (Sutanto et al., 2020). Till recently, these covalent-binding drugs were not invented intentionally, but their covalent binding modes were invented only after their development. Targeted covalent inhibition has come out as a validated strategy to drug discovery after the FDA approval so fibrutinib (2013), afatanib (2013) and osimertinib (2015), drugs that were intended to
undergo an irreversible Michael addition reaction with a specific cysteine amino acid residue of a target protein (Gehringer & Laufer, 2019). Studies as reported in the literature have demonstrated that Cys145 is a key residue in the active site of Mpro which makes this residue an attractive target for covalent binding by Mpro inhibitors (Figure 2) (Mengist et al., 2020; Ullrich & Nitsche, 2020; Kanhed et al., 2020).

Many research groups have carried out the repurposing of FDA-approved drugs, and to date, many FDA-approved drugs are being investigated in more than 100 clinical trials (Pawar, 2020; Joshi et al., 2020; Serafin et al., 2020; Abuo-Rahma et al., 2020). Some of them have turned out to be effective against SARS-CoV-2 and used as COVID-19 treatments (Pandey et al., 2020). At the same time, several companies have initiated clinical trials of vaccines against COVID-19. Keeping in mind the severity of the disease and urgent medical need for an efficient drug, it was envisaged to identify potential leads for SARS-CoV-2 Mpro inhibitors amongst the reported curcumin derivatives for possible COVID-19 treatment.

This study aims to identify novel potential Mpro inhibitors that could inhibit the intended protein irreversibly through covalent modification of its active site cysteine residue. Our quest for a potential lead is to identify compounds having Michael acceptor-like structures as they are strong electrophiles which react covalently with the nucleophilic cysteine thiolate group in the enzyme active site. In the current report, a systematic virtual screening of a library of curcumin derivatives was carried out using the Glide and CovDock modules of Schrodinger Suite. The promising compounds are discussed here in detail that could be used as possible hits for further drug development for getting an adequate drug treatment of COVID-19.

**RATIONALE FOR SELECTION OF CURCUMIN SCAFFOLD**

In an incisive crystal structure analysis of the SARS-CoV-2 Mpro (PDB Code: 6LU7), it was observed that there is Cys145 amino acid residue in the active site which was covalently attached to the β-position of the peptide-like α,β-unsaturated carbonyl compound N3 (Figures 3A and 3B) as a Michael adduct (Jin et al., 2020). In addition to this, it was also noted that the ligand N3 had adequate hydrogen bonding and stacking interactions with different hydrophilic and hydrophobic regions of
the protein. Michael-type of addition reaction of thiol to α,β-unsaturated ketone functional group is very well-known in biological systems. These observations made us to select molecules possessing α,β-unsaturated ketone as sufficiently electrophilic warhead in the ligand for covalent interactions with the enzyme along with additional structural features for noncovalent interactions.

Turmeric is crowned with captivating medicinal properties in traditional Indian literature (Prasad & Aggarwal, 2011; Chattopadhyay et al., 2004). It is used routinely as a spice, food preservative, and also for various minor and major illnesses as a medication by the Indians. These impressive therapeutic attributes of turmeric inspired the researchers to explore the therapeutic potential of its principal constituent curcumin. Curcumin (1,7-bis(4-hydroxy-3-methoxyphenyl)-1,6-heptadiene-3,5-dione), also termed as diferuloylmethane, is the main natural polyphenol found in the rhizomes of turmeric (Curcuma longa and in others Curcuma spp.) belonging to the Zingiberaceae family. There are three common pharmacophoric features present in the structure of curcumin i.e. the α,β-unsaturated keto-enol system, two phenolic moieties and the seven-carbon skeleton linking both the phenolic groups which are crucial for covalent and/or noncovalent interactions with biological macromolecules. The aromatic functionality can have π-π interactions while the phenolic hydroxyl groups and keto-enol groups can participate in hydrogen bonding interactions. The seven-carbon linker provides flexibility to the molecule so that it can easily adopt suitable conformation to maximize the intermolecular interactions. The covalent interactions of curcumin are mostly due to the reaction of the thiol group present in the proteins with the β-position of the α, β-unsaturated ketone present in curcumin structure. It is these covalent and non-covalent interactions with biomolecules that are responsible to elicit a specific biological activity (Nelson et al., 2017). These valuable structural features (Figure 3C) in a single scaffold inspired us to take advantage of their presence in the search for potential Mpro inhibitors for the treatment of COVID-19.
RESULTS AND DISCUSSION

Protein Preparation

The 3D-crystal structure of cysteine protease Mpro was retrieved from RCSB protein data bank (PDB ID: 6LU7) (Jin et al., 2020) and the covalent bond between the co-crystallized ligand N3 and amino acid Cys145 residue was cleaved. After bond cleavage, structures of the ligand N3 and Cys145 residue of the enzyme were reconstructed by making suitable chemical changes and the resulting ligand-protein complex was refined using Protein Preparation Wizard in Schrödinger (Schrödinger, LLC, New York, NY, 2020). Receptor grid was generated on the active site of Mpro protein by considering the centroid of ligand molecule N3 as the centre of the grid. The grid coordinates (i.e. X, Y, and Z) were -10.47, 12.23, and 68.7, respectively. To validate the generated grid, the co-crystallized ligand (N3) was re-docked into the active site of the protein using the generated grid. Here, the N3 molecule showed similar pattern of orientation and interactions, such as hydrogen bonding with Glu166, Gln189 and Thr190 residues of the active site. The XP docking score between N3 and Mpro protein was −7.85 kcal/mol. This cognate docking was analyzed further by identifying the all atom RMSD value of the re-docked N3 ligand vis-a-vis the co-crystallized ligand, and it was found to be 0.452 Å, which validated the docking protocol.

Library Designing and Docking-Based in Silico Screening

A library of reported curcumin derivatives was prepared using the Scifinder similarity search of curcumin structure. The resulting structures (12,421 compounds, with a similarity of 70 – 74% with curcumin) were retrieved. Among these, compounds possessing α,β-unsaturated carbonyl group (warhead) were searched. The selected compounds (5,000 compounds) having this warhead were structurally polished for docking by Ligprep module of the Schrodinger Suite. These compounds were screened in silico systematically using sequential conformational precision approaches which included HTVS, SP, XP and CovDock protocols of Schrodinger Suite on the Mpro protein and the molecules obtained in the final step were analyzed manually for their predicted binding modes and
scores (Toledo Warshaviak et al., 2014; Al-Khafaji et al., 2020). Amongst them, the top 10 potential covalent inhibitors of the Mpro of SAR-CoV-2 have been described here (Table 1). For all these compounds, the distance between the sulfur atom of Cys145 residue and nucleophilic β-position carbon of curcumin derivatives is shown in Table 1. In Mpro-N3 enzyme-ligand complex, the distance between the sulfur atom of Cys145 and the covalent carbon atom of N3 ligand is 1.8 Å, whereas this distance is around 1.8 Å for all the reported curcumin derivatives (except for compound 6). So, the proximity between these two atoms (i.e. sulfur and carbon) suggests high potential for covalent bonding between the chosen compounds and Cys145 residue of Mpro active site.

**MM-GBSA Binding Energy Calculation**

The docked ligand-receptor complexes were further analyzed for Molecular Mechanism-Generalized Born Surface Area (MM-GBSA) analysis to predict the free binding energy of ligand-receptor complexes (Genheden & Ryde, 2015). These binding free energies were used for re-scoring the top leads obtained after CovDock screening as binding energy is considered to be more accurate than any other scoring parameter of molecular docking (Rastelli & Pinzi, 2019). The total free energy binding (MM-GBSA ΔG_{bind}) was estimated as follows using the software:

\[
\text{MMGBSA } \Delta G_{\text{bind}} = G_{\text{Complex}} - (G_{\text{Receptor}} + G_{\text{Ligand}})
\]

where \(G_{\text{Complex}}\), \(G_{\text{Receptor}}\), and \(G_{\text{Ligand}}\) represent the energies of optimized ligand-receptor complex, optimized receptor and optimized ligand, respectively.

The chemical structures of the top-ranked molecules with curcumin are depicted in Figure 4. Molecular docking scores along with other parameters and MM-GBSA energy components are represented in Tables 1 and 2.

**Interaction Analysis of the Identified Curcumin Derivatives**

After the docking study, top-ranked compounds were arranged according to their free binding energies. Lower binding energy represents a more favorable binding of a ligand with protein. The best poses were visualized for H-bonding interactions, π-π interactions using PyMol (DeLano L, 2002). By default, the covalent docking gives only a single docked pose per ligand, which is the lowest-energy pose. However, as shown in the curcumin structure, there are two matches of the ligand SMARTS pattern for reactive residue, so the covalent docking gave two docking poses (Figure 5). Ranking of poses can be done according to their docking scores and MM-GBSA binding energies. Here the pose-I where curcumin (1) is represented in the cyan color sticks showed better docking score (-7.028 kcal/mol), MM-GBSA ΔG_{bind} (-51.07 kcal/mol) and cdock affinity (-7.028 kcal/mol) compared to the pose-II represented in grey color. The interactions of curcumin (1) and other derivatives (2-10) with the active site of Mpro discussed below are for the lowest-energy poses.

The molecular docking of curcumin (1), the basic scaffold within the active site of Mpro is shown in Figure 6. It shows that curcumin is fitting snugly inside the substrate-binding pocket of Mpro, interacting covalently with Cys145 residue (covalent binding affinity value of -7.028 kcal/mol). Curcumin showed good binding free energy (-51.07 kcal/mol). Main contributors to this notable binding of curcumin to Mpro are the exceptionally strong van der Waals interactions (ΔG_{vdW}) followed by electrostatic interactions (ΔG_{Coulomb}), lipophilic interactions (ΔG_{Lip}), and hydrogen bond interactions (ΔG_{Hbond}). The phenolic hydroxyl group showed H-bonding with Thr26 residue (1.66 Å) and oxygen of methoxyl group on phenyl ring showed H-bonding with Gly143 residue (2.10 Å). Curcumin is found to involve in an important covalent interaction with Cys145 residue of Mpro through Michael addition. The α,β-unsaturated ketone in curcumin showed a covalent bond (1.83 Å) between the β-position and Cys145 residue. LE value indicates the effectiveness of the molecule to use its structural features for binding to the target. Curcumin showed LE value of 0.27 kcal mol^{-1} per heavy atom.
Compound (2) having hydroxypropyl acrylate side chains attached to two hydroxyl groups of curcumin showed prominent binding interaction to Mpro with free binding energy and covalent binding affinity of -67.44 and -8.341 kcal/mol, respectively. As shown in Figures 7A and 7B, the hydroxy groups present at C2 position of the propyl acrylate side chains showed H-bondings with Thr25 residue (1.93 Å) and Glu166 residue (1.72 Å). The covalent bonding between the β-position of the α,β-unsaturated ketone group with Cys145 residue (1.84 Å) and H-bonding between the oxygen of α,β-unsaturated ketone and Gly143 residue (2.21 Å) provide stability to the ligand-receptor complex.

Similar to compound (2), compound (3) showed notable binding to Mpro with a covalent binding affinity value of -8.821 kcal/mol (Figures 7C and 7D). Hydroxyl groups present in the compound showed H-bondings with Thr24 (2.27 Å), Thr25 (1.84 Å) and Thr190 (1.61 Å) residues. Carbonyl group oxygen of acrylate moiety showed H-bonding with Ser46 residue (2.28 Å). The covalent bonding between the β-position of α,β-unsaturated ketone with Cys145 residue (1.83 Å) and H-bonding between the oxygen of α,β-unsaturated ketone and His41 residue (2.05 Å) provided stability to the ligand-receptor complex.

Compound (4) having 1-adamantylamine moiety attached to ethyl curcumin through a propionyl bridge showed strong interaction with Mpro having free binding energy and covalent binding affinity of -61.60 and -7.696 kcal/mol, respectively (Figures 7E and 7F). The –NH of 1-adamantylamine

Figure 4. Chemical structures of top-ranked curcumin (1) and its derivatives (2-10)
moiety gets protonated at physiological pH and showed H-bondings with Phe140 residue (2.63 Å) and Glu166 residue (1.89 Å). The protonated nitrogen showed strong salt bridge interaction with Glu166 residue (2.84 Å). Carbonyl oxygen of propionyl bridge showed H-bonding with Glu166 residue (1.69 Å). The covalent bonding between the β-position of the α,β-unsaturated ketone with Cys145 residue (1.75 Å) and H-bonding between α,β-unsaturated ketone and Gly143 residue (1.75 Å) provide stability to the ligand-receptor complex.

Table 1. Molecular docking results for the hit compounds from the library of curcumin derivatives

<table>
<thead>
<tr>
<th>Compd</th>
<th>Rank</th>
<th>Docking Score</th>
<th>Atom Distance</th>
<th>cdock Affinity</th>
<th>MM-GBSA ΔG Bind</th>
<th>Glide LE</th>
<th>RMSD (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>-8.156</td>
<td>1.84 Å</td>
<td>-8.341</td>
<td>-67.44</td>
<td>-0.181</td>
<td>0.042</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-9.091</td>
<td>1.83 Å</td>
<td>-8.821</td>
<td>-65.99</td>
<td>-0.198</td>
<td>0.027</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-6.134</td>
<td>1.75 Å</td>
<td>-7.696</td>
<td>-61.60</td>
<td>-0.104</td>
<td>0.040</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>-9.656</td>
<td>1.84 Å</td>
<td>-8.562</td>
<td>-59.84</td>
<td>-0.268</td>
<td>0.040</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>-7.791</td>
<td>2.15 Å</td>
<td>-8.204</td>
<td>-59.02</td>
<td>-0.190</td>
<td>0.043</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>-8.031</td>
<td>1.83 Å</td>
<td>-8.917</td>
<td>-58.36</td>
<td>-0.136</td>
<td>0.044</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>-7.045</td>
<td>1.85 Å</td>
<td>-7.190</td>
<td>-56.48</td>
<td>-0.157</td>
<td>0.040</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>-7.198</td>
<td>1.84 Å</td>
<td>-7.174</td>
<td>-55.95</td>
<td>-0.248</td>
<td>0.047</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>-7.334</td>
<td>1.84 Å</td>
<td>-7.259</td>
<td>-53.87</td>
<td>-0.216</td>
<td>0.047</td>
</tr>
<tr>
<td>Cur (1)</td>
<td>10</td>
<td>-7.045</td>
<td>1.83 Å</td>
<td>-7.028</td>
<td>-51.07</td>
<td>-0.27</td>
<td>0.048</td>
</tr>
<tr>
<td>N3 lig.</td>
<td>-</td>
<td>-7.466</td>
<td>1.84 Å</td>
<td>-7.466</td>
<td>-74.91</td>
<td>-0.167</td>
<td>0.043</td>
</tr>
</tbody>
</table>

Docking scores and other values are listed for the identified pose. Atom distance is the distance between the putative reactive atom of the drug and Cys145. cdock affinity: covalent docking affinity, LE: ligand efficiency, binding free energy per heavy atom count (LE = ΔG/HA), Cur: curcumin, N3 lig: N3 ligand.

Table 2. Energy components of the Mpro-ligand complexes from MMGBSA calculation

<table>
<thead>
<tr>
<th>Compd</th>
<th>ΔG&lt;sub&gt;Coulomb&lt;/sub&gt;</th>
<th>ΔG&lt;sub&gt;Coulomb&lt;/sub&gt;</th>
<th>ΔG&lt;sub&gt;Bond&lt;/sub&gt;</th>
<th>ΔG&lt;sub&gt;Soln GB&lt;/sub&gt;</th>
<th>ΔG&lt;sub&gt;Lipophilic&lt;/sub&gt;</th>
<th>ΔG&lt;sub&gt;Packing&lt;/sub&gt;</th>
<th>ΔG&lt;sub&gt;vdW&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-42.72</td>
<td>10.72</td>
<td>-3.25</td>
<td>34.91</td>
<td>-13.3</td>
<td>-0.95</td>
<td>-52.86</td>
</tr>
<tr>
<td>3</td>
<td>-32.41</td>
<td>1.99</td>
<td>-3.72</td>
<td>35.41</td>
<td>-14.24</td>
<td>-0.97</td>
<td>-52.04</td>
</tr>
<tr>
<td>4</td>
<td>-38.13</td>
<td>12.11</td>
<td>-3.09</td>
<td>54.71</td>
<td>-18.57</td>
<td>-1.68</td>
<td>-66.95</td>
</tr>
<tr>
<td>5</td>
<td>-28.61</td>
<td>-0.24</td>
<td>-2.5</td>
<td>40.71</td>
<td>-15.93</td>
<td>-2.19</td>
<td>-51.09</td>
</tr>
<tr>
<td>6</td>
<td>-13.6</td>
<td>3.04</td>
<td>-1.33</td>
<td>24.16</td>
<td>-9.36</td>
<td>-3.11</td>
<td>-58.82</td>
</tr>
<tr>
<td>7</td>
<td>-32.06</td>
<td>0.66</td>
<td>-2.59</td>
<td>45.35</td>
<td>-11.55</td>
<td>-2.75</td>
<td>-55.43</td>
</tr>
<tr>
<td>8</td>
<td>-38.66</td>
<td>16.46</td>
<td>-1.4</td>
<td>52.84</td>
<td>-20.05</td>
<td>-3.08</td>
<td>-62.57</td>
</tr>
<tr>
<td>9</td>
<td>-22.44</td>
<td>-0.95</td>
<td>-0.99</td>
<td>27.33</td>
<td>-15.23</td>
<td>-0.09</td>
<td>-43.58</td>
</tr>
<tr>
<td>10</td>
<td>-12.69</td>
<td>0.3</td>
<td>-0.57</td>
<td>23.65</td>
<td>-17.15</td>
<td>-3.36</td>
<td>-44.05</td>
</tr>
<tr>
<td>Cur (1)</td>
<td>-18.9</td>
<td>1.9</td>
<td>-0.97</td>
<td>26.84</td>
<td>-15.47</td>
<td>-0.11</td>
<td>-44.36</td>
</tr>
<tr>
<td>N3 lig.</td>
<td>-29.96</td>
<td>5.92</td>
<td>-3.2</td>
<td>44.21</td>
<td>-14.70</td>
<td>0.00</td>
<td>-77.17</td>
</tr>
</tbody>
</table>

Figure 5. Overlap of two binding poses of curcumin (1) with SARS-CoV-2 Mpro. Curcumin is shown in cyan color and grey color. The docking scores and energies reported here are in kcal/mol.

<table>
<thead>
<tr>
<th></th>
<th>Pose-I</th>
<th>Pose-II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligand colour</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Docking score</td>
<td>-7.028</td>
<td>-6.269</td>
</tr>
<tr>
<td>Prime energy</td>
<td>-12613.21</td>
<td>-12509.60</td>
</tr>
<tr>
<td>MM-GBSA ΔG&lt;sub&gt;Bind&lt;/sub&gt;</td>
<td>-51.07</td>
<td>-41.79</td>
</tr>
<tr>
<td>cdock affinity</td>
<td>-7.028</td>
<td>-6.269</td>
</tr>
<tr>
<td>RMSD (Å)</td>
<td>0.048</td>
<td>0.049</td>
</tr>
</tbody>
</table>

Figure 6. Binding mode and molecular interactions of curcumin (1) with SARS-CoV-2 Mpro (PDB ID: 6LU7). A: 3D structure of Mpro having three domains as shown in three different colors. The ligand-binding site region is enclosed within the square. B: Magnified view of the catalytical center. Curcumin is shown as a stick model with cyan color. The dashed circle highlights the C-S covalent bond. C: 2D-ligand interaction diagram of curcumin with Mpro active site.
Curcumin mono-α-D-arabinopyranoside (5) is able to covalently dock inside the Mpro with a covalent binding affinity value of -8.562 kcal/mol. As shown in Figures 8A and 8B, the phenolic hydroxyl group showed H-bonding with Glu166 residue (2.53 Å). The α-D-arabinopyranosyl ring was also found to fit closely inside the polar region of Gln189, Thr190 and Gln192 amino acid residues. The hydroxy group on the 4th position of the α-D-arabinopyranosyl ring showed H-bonding with Thr190 residue (1.62 Å). The α,β-unsaturated ketone group of curcumin moiety showed a covalent bond (1.84 Å) between the β-position and Cys145 residue.

Insertion of the thiophenyl group at C5 position of the phenyl ring in curcumin showed improvement in the covalent binding affinity (-8.204 kcal/mol). The phenolic hydroxyl group of the compound (6) showed H-bonding with Glu166 residue (2.23 Å) and oxygen of methoxyl group on phenyl ring showed H-bonding with Thr24 residue (2.36 Å) (Figures 8C and 8D). The covalent

Figure 7. Docking interactions of compounds (2-4: A-E) in the active site of Mpro. A, C and E: Binding modes of compounds 2, 3 and 4 within Mpro active site, respectively. Ligands are shown as cyan sticks. Mpro residues are shown as atom type color sticks. Hydrogen bonds formed between ligands and receptor are indicated by dotted lines. The circle highlights the C-S covalent bond. B, D and F: 2D-ligand interaction diagram of compounds 2, 3 and 4 with Mpro active site, respectively.
bonding between the β-position of the α,β-unsaturated ketone with Cys145 residue (1.83 Å) and H-bonding between α,β-unsaturated ketone with Cys145 residue (2.15 Å) provide additional stability to the ligand-receptor complex.

Some quercetin-curcumin hybrids have been reported as antioxidant agents by Yanase et al (2010). Amongst these hybrids, compound (7) showed notable binding to Mpro with covalent binding affinity value of -8.917 kcal/mol as shown in Figures 8E and 8F. The phenolic hydroxyl group at C3′ position and oxygen of the chromone ring of quercetin moiety showed H-bondings with Glu166 residue (1.75 Å) and Asn142 residue (2.62 Å), respectively. Carbonyl group oxygen of curcumin moiety showed H-bonding with His41 residue (2.02 Å). The covalent bonding between the β-position of the α,β-unsaturated ketone in curcumin with Cys145 residue (1.83 Å) provide stability to the ligand-receptor complex.

Liu et al. (2013) reported novel antiproliferative curcumin analogs having o-aminoalkyl moieties attached to the curcumin scaffold. Compound (8) with o-pyrrolidinopropyl chains attached to the curcumin scaffold showed good binding interactions with the Mpro active site as shown in Figures 9A and 9B. The aromatic phenyl rings of curcumin scaffold allow the formation of more favorable π−π stackings with His41 (4.92 Å) and His163 (5.03 Å). The nitrogen atom of the pyrrolidine ring got protonated at physiological pH. The salt bridge between −NH of the pyrrolidine and Glu166 (4.56 Å), and H-bonding interaction between the nitrogen of pyrrolidine and Glu166 (2.10 Å) provided stability to the ligand−receptor complex. It also showed covalent bonding between the β-position of the α,β-unsaturated ketone with Cys145 residue (1.85 Å) and H-bonding between α,β-unsaturated ketone with Cys145 residue (2.22 Å).

Insertion of the fluoro group at C5 positions of the phenyl ring in curcumin as in compound (9) showed a slight improvement in the covalent binding affinity (~7.174 kcal/mol) (Figures 9C and 9D). This might be due to the electronegative nature of the fluoro group which made the β-position of the α,β-unsaturated ketone more electrophilic for Michael addition. The β-position of the α,β-unsaturated ketone showed covalent bonding with Cys145 residue (1.84 Å). The phenolic hydroxyl group showed H-bonding with Thr26 residue (1.66 Å).

Compound (10) with a 4-fluorophenyl ring in R-configuration showed good binding interactions within the active site of Mpro having a free binding energy value of -53.87 kcal/mol and covalent binding affinity value of -7.259 kcal/mol. As shown in Figures 9E and 9F, the covalent bonding between the β-position of the α,β-unsaturated ketone with Cys145 residue (1.84 Å) and H-bonding between α,β-unsaturated ketone with Gly166 residue (2.58 Å) provide stability to the ligand-receptor complex. A compound having the 4-fluorophenyl ring in S-configuration showed free binding energy (~46.33) and covalent binding affinity (~6.345 kcal/mol) for Mpro.

**Molecular Dynamics (MD) Simulation Studies**

The interactions of a ligand with receptor in the non-covalent mode help us to understand the possibility of covalent interactions between the amino acid of the enzyme and the warhead group of the ligands. When the sulfur atom of thiol group of Cys145 is in close proximity to the α,β-unsaturated carbon of the ligand, it is able to bind covalently through Michael addition reaction. Here, the distance between these two centres was monitored using the MD simulations. The two complexes of the top-ranked compounds (2 and 3) with the active site fragment of Mpro were evaluated using MD analysis. The covalent bond was broken and necessary atoms were added to reconstruct the structures. This complex was considered as the starting pose for carrying MD analysis. A 10 ns MD analysis was carried out. To understand the binding stability of the ligand-receptor complex over the simulation period, statistical properties like RMSD-P, RMSF-P, and RMSD-L (P = protein; L = ligand), van der Waals and electrostatic interaction energies were examined to cross-check and support the stability of the interactions.

The RMSD-P for Mpro in complexation with compound (2) was in the range of 0.1–0.23 nm with an average of 0.15 nm (Figure 10A). This suggests the stability of the protein while having
compound (2) in the active site over this period of simulation. Despite having freely rotatable bonds, the average RMSD-L was 0.54 nm with a range of 0.31 nm to 0.68 nm (Figure 10B). This suggests the protein–ligand stability without a major change in orientation of the ligand in the active site over the period of simulation time. The RMSF value explains the structural integrity and residual mobility of the structure. Here again, all the residues including terminal residues of the protein structure showed RMSF-P below 0.45 nm, while having compound (2) in the active site (Figure 10C). The distance module from GROMACS was used to determine the distance between the thiol group of Cys145 and the α,β-unsaturated carbon of the ligand. The average distance between the sulfur atom of thiol group of Cys145 and the α,β-unsaturated carbon of the ligand was observed to be 0.40 nm with a range of 0.15 nm to 0.64 nm (Figure 10D). The short-range electrostatic (Coul-SR) and van der Waals/
hydrophobic (LJ-SR) interaction energies between the protein and compound (2) explained promising electrostatic as well as hydrophobic interactions. The average values of Coul-SR -56.94 ± 11 kJ/mol and LJ-SR -138.58 ± 10 kJ/mol were observed. This explains that the role of hydrophobic interaction was more important than the electrostatic interactions in stabilizing the complex.

Similar evaluation was done for Mpro in complexation with compound (3). The RMSD-P was observed in the range of 0.11–0.25 nm with an average of 0.19 nm (Figure 11A). Despite having multiple rotatable bonds, the RMSD-L was observed in the range of 0.38–0.61 nm with an average value of 0.46 nm (Figure 11B). The observed RMSF-P for all the residues except for the terminal residues was below 0.2 nm, whereas for the terminal residues, it was below 0.6 nm (Figure 11C). The average distance between the sulfur atom of thiol group of Cys145 and the α,β-unsaturated carbon of...
the ligand was 0.37 nm with a range of 0.21 nm to 0.58 nm (Figure 11D). The short-range electrostatic (Coul-SR, energy: \(-49.73 \pm 6 \text{ kJ/mol}\)) and van der Waals/hydrophobic (LJ-SR, energy: \(-153.45 \pm 4 \text{ kJ/mol}\)) interaction energies suggested promising interactions between the ligand and the protein.

**EXPERIMENTAL**

**Preparation of the In Silico Compound Library**

A library of curcumin derivatives was generated by similarity search of curcumin scaffold in Scifinder. These molecules were retrieved (12421 compounds, 70 – 74% similarity) and the search for the compounds having \(\alpha,\beta\)-unsaturated carbonyl group (as a chemical warhead) was performed manually.
on the collected compounds’ library. The selected compounds (5000 compounds) having this warhead were prepared for computational study at physiological pH condition retaining the chiralities by using LigPrep module of Schrodinger (Schrodinger, LLC, New York, NY, 2020).

The 3D crystal structure of Mpro retrieved from RCSB Protein Data Bank (PDB ID: 6LU7) (Jin et al., 2020) was prepared to ensure structural correctness for hydrogen consistency, bond orders, steric clashes and charges using Protein Preparation Wizard in Schrodinger Suite supported by OPLS3e force field. Thus, the prepared structure was used for receptor grid generation required for the docking protocol. The receptor grid was generated considering the position of the co-crystallized ligand N3 in the active site.

Molecular Docking and Interaction Analysis

The noncovalent molecular docking protocol was applied prior to covalent docking in the virtual screening workflow based on the hypothesis that the covalent inhibitors enter into the binding site of the target protein first, and establish physical interactions with it before forming a covalent bond with the protein (Figure 12). For that the generated curcumin derivatives library having 5000 molecules was first screened using high-throughput virtual screening method. This is considerably fast but raw screening method. From the result, top 1500 molecules were picked up for the next standard precision (SP) docking. From the SP docking results, 500 molecules were carried forward for the extra-precision (XP) docking. From the XP analysis, top 50 compounds were considered for further covalent docking analysis. A covalent docking exercise was performed to identify curcumin derivatives that can target Cys145 residue of Mpro within the substrate-binding pocket, especially keeping in mind that the co-crystallized inhibitor N3 exists in covalent binding mode with Mpro. The targeted covalent docking of 50 compounds through Cys145 residue was carried out by CovDock module of Schrodinger Suite. In CovDock protocol, Cys145 was specified as reactive residue in the receptor, Michael addition as reaction type, α,β-unsaturated carbonyl group in the ligand as the active functional group represented by a SMARTS pattern \([\text{C,c]=[C,c]-[C,c,S,s]=}[\text{O}]\) and docking was performed in pose prediction mode. For the scoring of the docked poses, MM-GBSA calculations were performed within CovDock protocol. Finally, the top 20% molecules were selected for manual analysis of their predicted binding scores and interactions with the active site of Mpro using PyMol.

Molecular Dynamics Simulation Study

The MD studies between the selected compounds and Mpro (PDB code: 6LU7) were performed for a period of 10 ns by using GROMACS 2020.1 software as per our previous report (Kanhed et al., 2020). The covalent bond in ligand-enzyme complex was broken and necessary atoms were added to reconstruct the structures. This complex was considered as the starting pose for carrying out the MD analysis. The distance module from GROMACS was used to determine the distance between the thiol group of Cys145 and the α,β-unsaturated carbon of ligand.

CONCLUSION

In view of the long history of curcumin use in controlling various pathophysiological conditions and the unmet need for the drug development for COVID-19, we envisaged to assess the anti-SARS-CoV-2 potential of some reported curcumin derivatives as Mpro inhibitors using in silico approach. As a result, we identified curcumin derivatives as non-peptidomimetic covalent-binding inhibitors of SAR-CoV-2 Mpro by a systematic virtual screening method. The identified curcumin and its derivatives could modify Cys145 residue of Mpro by covalent bonding which could inhibit the proteolytic processing of viral polyproteins. Apart from covalent bonding with Cys145, these derivatives are also observed to form non-covalent interactions with other amino acid residues of the active site of Mpro. Collectively, these covalent as well as non-covalent interactions made these curcumin derivatives as potential lead molecules as Mpro inhibitors. Although experimental biological assessments are needed to support
these *in silico* findings; findings of the computational studies presented in this report are expected to assist the active researchers in the development of covalent Mpro inhibitors.
CORRESPONDING AUTHOR

Mange Ram Yadav*
Director (R & D), Centre of Research for Development
Parul University, Vadodara, Gujarat, India
E-mail address: mryadav11@yahoo.co.in

AUTHOR CONTRIBUTIONS

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

CONFLICTS OF INTEREST

There are no conflicts of interest to disclose.

ABBREVIATIONS

3CLpro, 3-Chymotrypsin-like protease; FDA, Food and Drug Administration; gRNA, Genomic RNA; HTVS, High-throughput virtual screening; LE, Ligand efficiency; MERS, Middle east respiratory syndrome; MM-GBSA, Molecular mechanism-generalized born surface area; Mpro, Main protease; NSP-5, Non-structural protein-5; PDB, Protein data bank; PLpro, Papain-like protease; RMSD, Root-mean-square deviation; SARS-Cov-2, Severe acute respiratory syndrome coronavirus-2; SP, standard precision; WHO, World Health Organization; XP, Extra precision.
REFERENCES


Dushyant Patel is working as doctoral fellow under the guidance of Prof. M. R. Yadav and Prof. A. V. Bedekar at the Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, Vadodara, India. He obtained his graduation and post-graduation (Pharmaceutical Chemistry) degrees from the same institute in 2012 and 2014, respectively. He has been awarded INSPIRE fellowship by Department of Science and Technology (DST) during doctoral study for research on Alzheimer’s disease. His research interests are mainly focused on designing and synthesis of various novel enzyme inhibitors for the treatment of Alzheimer’s and CVS disorders. Mr. Patel has 6 research publications, 2 book chapters, and 1 review article to his credit.

Divya M. Teli is currently working as Assistant Professor at L. M. College of Pharmacy. She holds her B.Pharm and M.Pharm degrees from L. M. College of Pharmacy. She is having 5 years of industrial experience and 2 years of academic experience.

Ashish M. Kanhed has obtained his PhD from The Maharaja Sayajirao University of Baroda in year 2015-16. Post PhD he has worked as Research Associate and Senior Research Associate at Navin Saxena Research and Technology Pvt. Ltd. (Research center of Rusan Pharma) in API and CADD department, where his work area was Narcotic API process development, Non-narcotic NCE design and synthesis and reaction process optimization on Continuous Flow Reactors (CFR). He also was a post-doctoral Fellow at University of KwaZulu Natal, Durban, South Africa (March to Nov. 2018) where the work area was design of novel computational models for the discovery of Hits against Tuberculosis, HIV and KSP target in Cancer treatment and their synthesis. Apart from these Dr. Kanhed has also worked as lecturer at Rahul Dharker College of Pharmacy, Mumbai University. Duration- 1st January 2011 to 31st January 2012. He is associated with SPPSPTM, NMIMS University, Mumbai Campus since July 2019.

Nirav Patel is working as doctoral fellow under the guidance of Prof. M. R. Yadav at the Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, Vadodara, India. He obtained his post-graduation (Pharmaceutical Chemistry) degree from the same institute in 2014. He has been awarded UGC-BSR fellowship during doctoral study for research on CVS disorders. His research interests are mainly focused on synthesis of novel molecules for the treatment of CVS disorders. Mr. Patel has 5 research publications, 2 book chapters, and 1 review article to his credit.

Bhavik Shah is studying in M.Pharm as Research Scholar (Pharmaceutical Chemistry) under the guidance of Dr. Navnit Prajapati at the Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, Vadodara, India. He obtained his graduation degree from the Gujarat Technological University in 2018. He has been awarded AICTE scholarship during Master’s study.

Amisha K. Vora is an assistant professor at Department of Pharmaceutical Chemistry, NMIMS, Mumbai, Maharashtra. She has 15 years’ experience in teaching undergraduates’ and postgraduates’ subjects like Medicinal Chemistry, Pharmaceutical Analysis and Pharmaceutical Organic Chemistry. She has 10 publications, 1 book chapter, and 3 government-funded projects in her credit.

Mahesh T. Chhabria is currently Principal of L. M. College of Pharmacy, Ahmedabad. His research interests include drug design and discovery of small molecules as Anti TB, Anticancer and antiviral agents.

M.R. Yadav is the Founder Dean, Faculty of Pharmacy, The Maharaja Sayajirao University of Baroda, Vadodara. He has worked as UGC-BSR Faculty Fellow, and Head and Professor, Pharmacy Department in the same university. Currently he is working as Director (R & D), at Parul University, Vadodara. Dr. Yadav has a research and teaching experience of almost 35 years with more than 200 National and International publications in the field of Medicinal chemistry and 19 Indian patent applications to his credit. His research interests include drug designing strategies and synthesis of medicinally active compounds. Dr Yadav has been awarded with 'Teacher of the year 2010 Award' by Association of Pharmacy Teachers of India and ‘Eminent Teacher of the Year 2013’ by Association of Pharmacy Professionals, ‘Professor V. M. Kulkami Visiting Fellow Award in Pharmaceutical Sciences’ by Institute of Chemical Technology (Autonomous University), Mumbai and ‘UGC-Visiting Fellow Award’ by University Institute of Pharmaceutical Sciences, Panjib University, Chandigarh. Currently he is member, Academic Council, Central University of Rajasthan and Scientific Advisor, Office of the Controller General of Patents, Design and Trade Marks, Government of India.