Identification of potent COVID-19 Main Protease (Mpro) inhibitors from Curcumin analogues by Molecular Docking Analysis

Jaydip Bhaliya
bhaliyajaydip6041@gmail.com
ITM Vocational University, Vadodara, Gujarat

Vraj Shah
vraj97966@gmail.com
ITM Vocational University, Vadodara, Gujarat

ABSTRACT

These days, COVID-19, a new strain of coronavirus COVID-19 is rapidly spreading, has affected more than 210 countries and territories received global attention. The lack of efficacious medicines or vaccines in opposition to SARS-CoV-2 has also worsened the situation. Hence, there is a pressing want to increase up research for the improvement of potential therapeutics and low priced diagnostic in opposition to COVID-19. The crystalized form of COVID-19 main protease (Mpro) was illustrated by a Chinese researcher Liu et al. (2020) which is a novel therapeutic drug target. The goal of the study is to identify COVID-19 Mpro potential from mono-carbonyl analogues of curcumin through binding free energy analysis into COVID-19 by utilizing molecular docking. We conducted docking simulation to mono-carbonyl analogues of curcumin as ligands into the main protease of COVID-19 as a protein. The 3D structure of the COVID-19 Mpro was downloaded from PDB (Code ID: 6LU7). The structure of ligands was prepared using Chem Bio Draw Ultra 12.0.02. Docking process, the interaction, and binding of ligands – protein was done using the software Molegro Virtual Docking (MVD) and visualized using the software Molegro Molecular Viewer (MMV). The results showed hydrogen bonding and Steric interaction between compound A2 (curcumin analogues) with COVID-19. Moldock scores of compound A2 is -202.476 kcal/mol. It is predicted that compound A2 has potency as a lead compound to find new antiviral candidates against COVID-19 for possible therapeutic agents.

Keywords — Coronavirus (Covid-19), Curcumin analogues, Docking, Molegro Virtual Docker

1. INTRODUCTION

The outbreak of COVID-19 known as severe acute respiratory syndrome coronavirus 2 originated from the 4 admitted sufferers with pneumonia who had been working in the Wuhan, Huanan seafood wholesale market, China doing commercial enterprise in stay poultry, aquatic products, and some wild animals. The now-closed market being a frequent aspect of infections stimulated the faith that the contamination may additionally be linked with certain animals. The species that harbored the SARS-CoV-2 was illustrated to be in all likelihood bat, containing 96% same at the whole Genome stage [1]. The virus is made by two layers among them one is an outer layer called protein crown and the inner layer which is the core of the virus called genetic material [2]. As of 20 April 2020, a total of 23, 14,621 confirmed cases globally, with 1, 57,847 deaths in 210 countries and territories had been reported by WHO. COVID-19 cases are still steadily growing due to its rapid human to human transmission [3].

Coronaviruses, getting their identity from the viruses' indistinct resemblance to monarchical crowns when imaged the usage of an electron microscope, are a massive family of viruses that purpose sickness in mammals and birds. Coronaviruses can reason ailments that vary from the frequent bloodless to lots greater extreme ailments like SARS, Middle East respiratory syndrome, and COVID-19 [4]. To add to this, many efforts have been made to produce the vaccine, there is still no specific vaccine or antiviral drug for the prevention or treatment of this pandemic (COVID-19). Thus, there is an urgent need to identify and develop effective antivirals against COVID-19 fight this deadly virus [5]. On Jan seven, 2020 researchers obtained a sequenced genome of SARS-CoV-2.(2020) [6].one in every of the research worker from china found the crystalized structure of COVID-19 main protease(Mpro) that it's a possible drug target macromolecule for the inhibition of SARS-CoV-2 replication. A key macromolecule that is required for chemical action maturation of the virus is that the main protease (PDB ID: 6LU7) [7].

Curcumin belongs to the category of polyphenol compounds derived from the South Asian herb turmeric that belongs to Curcuma domestica. The herb C. longa consists of curcuminoids that compose of curcumin, demethoxycurcumin, and bisdemethoxycurcumin. In Ayurvedic drugs, curcumin is widely used for numerous treatment aspects thanks to the therapeutic properties like anti-oxidant, antiviral, anti-septic, analgesic, antimarial, and anti-inflammatory [8]. It is suggested that stability and by eliminating the b-diketone moiety, metabolic profiles of curcumin could be enlarged. Despite this, some are of the
researches inclined to voice that the presence of the b-diketone moiety may additionally be integral for the medicinal activities of curcumin, current research from various unbiased agencies verified that some curcumin analogues containing a 5-carbon enone spacer without b-diketone either retained or increased growth-suppressive activities against several cancer cells. Research also showed that some mono-carbonyl analogues of curcumin without the b-diketone moiety exhibited better anti-bacterial and anti-inflammatory activities than those of curcumin [9]. Compounds chemically known as 1, 5-diaryl-1, 4-pentadien-3-ones are the mono-carbonyl analogues of curcumin. These are structural analogues of the natural product curcumin (1, 7-bis-(4-hydroxy-3-methoxyphenyl)-1, 6-heptadiene-3, 5-Dione) which is found as a major pigment in the Indian spice turmeric Curcuma longa, Zingiberaceae. Curcumin and 1, 5-diaryl-1, 4-pentadien-3-ones share not only similar chemical structures but also similar biological properties [10].

To overcome this world's biggest challenge, choose curcumin and mono-carbonyl analogues of curcumin as a vaccine or antiviral drug to cure this outbreak of COVID-19 due to its biological properties of curcumin. Furthermore, the technique used to detect whether curcumin derivatives give activity or not like vivo and vitro studies is time-consuming and very costly. As a consequence, many types of researches have indicated that computational approaches, such as structural bioinformatics pharmacophore modelling are the first-class choice [11]. Docking a number of ligands to the protein of activity accompanied with the aid of scoring to determine the binding affinity and to reveal the strength of interaction turn out to be appreciably used in digital screening of massive databases and lead optimization [12].

This paper reports screening of various mono-carbonyl analogues of curcumin bound directly or indirectly to COVID-19 extracted from the protein data bank, by utilizing the Molegro Virtual Docker Software. Moreover, mono-carbonyl analogues of curcumin inhibitors that target COVID-19 are considered as effective therapeutic agents to fight against COVID-19.

2. METHODOLOGY
2.1 Protein and ligand structures

The crystal structure of COVID-19 (PDB ID: 6LU7) [7] was directly downloaded to the workspace of MVD from the PDB accessed at the URL: (http://www.rcsb.org/pdb). The two-dimensional (2D) structures of curcumin and mono-carbonyl analogues of curcumin as ligands were obtained using the Chem Bio Draw 12.0.02 computer program. 2D to three dimensional (3D) representations were converted by the use of ChemBio3D 12.0.02 software and then, these were energetically minimized by using methods implemented in the same software, and saved as SDF format (*.sdf). The chemical 2D structures of all ligands are shown in table 1. To make correct predictions, it is essential that the imported structures must be right prepared, that is, the atom connectivity and bond orders are right, and partial PDB archives regularly have the bad or lacking undertaking of express hydrogens, and the PDB file structure could not accommodate bond order information. All vital valency assessments and H atom addition had been as a result carried out the use of the utilities furnished in MVD. The binding site indicates the place of engagement where the docking procedure will have appeared for promising poses (ligand conformations).

2.2 Molecular docking studies

The Molecular docking procedure was performed using the Molegro Virtual Docker (MVD) program [13]. In this study, docking of mono-carbonyl analogues of curcumin against COVID-19 has performed using MVD software. A single crystal structure of COVID-19 held in the Protein Data Bank (PDB) accessed at the URL (http://www.rcsb.org/pdb) under the criteria that they had reasonable resolution 2.16 Å.

After importing a Mpro in the MVD program, mono-carbonyl analogues of curcumin, have been docked against COVID-19 crystal structure and 10 independent runs were performed with the guided differential evolution algorithm, with every of these docking runs returning one answer (pose). The Moldock scoring feature used through MVD is derived from the PLP scoring features initially proposed by way of Gehlhaar et al. and prolonged later through Yang et al. [14]. The 10 solutions got from the 10 unbiased docking runs had been re-ranked, in order to enlarge the docking accuracy, by way of the usage of a greater complicated scoring function. In the MVD, alongside with the docking scoring terms, a Lennard Jones 12-6 plausible and sp2-sp2 torsion phrases had been additionally used [14]. On the foundation of pilot docking studies, the MolDock re-rank ratings had been chosen for rating the inhibitor poses, and for all the mono-carbonyl analogues of curcumin docking carried out here, the poses chosen as the excellent.

2.3 Docking Visualization

Molecular docking Interaction between ligands (analogues of curcumin) & protein (COVID-19) was visualized by Molegro Molecular Viewer 7.0.0 (MMV). MMV is a comprehensive software suite for analyzing and modelling molecular protein-ligand interactions, sequences and structures.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>2D structure of Ligand</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curcumin</td>
<td><img src="image" alt="Curcumin 2D Structure" /></td>
<td>[15],[16]</td>
</tr>
<tr>
<td>Compound</td>
<td>Structure</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>GO-Y086</td>
<td><img src="structure1.png" alt="GO-Y086" /></td>
<td></td>
</tr>
<tr>
<td>GO-Y129</td>
<td><img src="structure2.png" alt="GO-Y129" /></td>
<td></td>
</tr>
<tr>
<td>GO-Y130</td>
<td><img src="structure3.png" alt="GO-Y130" /></td>
<td></td>
</tr>
<tr>
<td>GO-Y147</td>
<td><img src="structure4.png" alt="GO-Y147" /></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td><img src="structure5.png" alt="A1" /></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td><img src="structure6.png" alt="A2" /></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td><img src="structure7.png" alt="B1" /></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td><img src="structure8.png" alt="B2" /></td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td><img src="structure9.png" alt="B3" /></td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td><img src="structure10.png" alt="B4" /></td>
<td></td>
</tr>
</tbody>
</table>
3. RESULT AND DISCUSSION

The research through molecular docking is important, now not solely from a theoretical viewpoint, to explain the relationship between the shape of ligand and the feature of protein however additionally in terms of practical applications, as they allow interpretation of the transporting process and therapeutic effectiveness of drugs [18].

The Molecular docking study is thus an optimization problem, where the aim is to find the best ligand-receptor interaction binding mode with the lowest potential energy. The method of docking includes sampling the coordinate space of the target binding site and scoring every feasible ligand pose inside that site, the perfect scoring pose then taken. There are many distinctive docking packages now on hand and they vary in the nature of the sampling algorithms they employ, in their manner of coping with ligand and protein flexibility, in the scoring features they use, and in the cpu time they required. In the research suggested here, MVD used to be used, due to the fact it confirmed greater docking accuracy when benchmarked in opposition to different reachable docking packages (MD: 87%, Glide: 82%, Surflex: 75%, FlexX:58%) and has been proven to be profitable in numerous latest studies, however additionally for motives of fee and person- friendliness [13].

MVD software program robotically identifies conceivable binding sites (also referred as cavities or active sites) by using the usage of its cavity detection algorithm. The cavities inside a 30 x 30 x 30 Å³ dice founded at the experimentally regarded ligand function had been used. The cavities that are recognized through the cavity detection algorithm are then used by using the guided differential evolution search algorithm to focal point the search, to that unique location during the docking simulation. In the case of the crystal structures for COVID-19, the program generally identified four different binding sites (Figure 1). From these four predicted cavities the one with the highest volume (131.072 Å³) was selected for consideration, as it includes the bound ligand.

One application of molecular docking study is to optimizing targeted lead candidates against a protein in pharmaceutical through silico [13]. Here, mono-carbonyl analogues of curcumin as ligands have been designed to obtain more potent compounds as inhibitors of COVID-19 Mpro. The molecular docking of these ligands was performed with the crystal structure of COVID-19 Mpro, and their calculated volumes (in Å³) (PDB code 6LU7 [6]), detected cavity green, carbon atoms grey, oxygen atoms red, nitrogen atoms are summarized in Table 2. In table 2 show, H-bond is the hydrogen bonding energy between protein and ligand. Steric is the steric interaction energy between protein and ligand [18].

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Moledock score</th>
<th>Rerank score</th>
<th>H bond</th>
<th>Steric</th>
</tr>
</thead>
<tbody>
<tr>
<td>curcumin</td>
<td>-151.545</td>
<td>-103.901</td>
<td>-6.78773</td>
<td>-163.442</td>
</tr>
<tr>
<td>GO-Y086</td>
<td>-184.916</td>
<td>-95.54</td>
<td>-5.0432</td>
<td>-210.085</td>
</tr>
<tr>
<td>GO-Y129</td>
<td>-178.256</td>
<td>-127.931</td>
<td>0</td>
<td>-127.931</td>
</tr>
<tr>
<td>GO-Y130</td>
<td>-176.89</td>
<td>-105.271</td>
<td>-9.9144</td>
<td>-179.269</td>
</tr>
<tr>
<td>GO-Y147</td>
<td>-182.582</td>
<td>-100.986</td>
<td>-8.30835</td>
<td>-220.075</td>
</tr>
<tr>
<td>B1</td>
<td>-179.794</td>
<td>-86.9086</td>
<td>-5.25149</td>
<td>-190.142</td>
</tr>
<tr>
<td>B2</td>
<td>-189.464</td>
<td>-84.1768</td>
<td>-6.01892</td>
<td>-204.037</td>
</tr>
<tr>
<td>B3</td>
<td>-185.619</td>
<td>-83.7425</td>
<td>-9.2213</td>
<td>-194.195</td>
</tr>
<tr>
<td>B4</td>
<td>-183.631</td>
<td>-135.368</td>
<td>-2.99314</td>
<td>-208.592</td>
</tr>
</tbody>
</table>

Docking simulations using MVD revealed MolDock scores are between -202.476 to -176.89 kcal/mol. docking to COVID-19 Mpro was also performed on curcumin. Moldock scores of all ligands are lower than curcumin. A2 compound has the lowest energy lead candidates against COVID-19 as compare to other compounds. This result indicates that the above -mentioned molecules are predicted to be antiviral drug candidates, as observed in the cavity of the crystallographic structure of COVID-19 Mpro and the best conformation obtained theoretically for A2 compound are shown in Figure 2. The result suggests that the software reproduced the appropriate conformation of A2 compound inside its binding site in the COVID-19 Mpro.
Fig. 2: Structural cartoon of COVID-19 (PDB code 6LU7), The α helices and β strands are represented as coils (red) and arrows (blue), respectively. A2 compound is represented in ball and stick (yellow)

The best docking poses obtained on the basis of MVD re-rank score for curcumin and top three ligand A2, B2 and B3 with the crystal structures of COVID-19 are presented in Figures 3 to 6. In which The best score docking solution of ligands and with the selected crystal structure of COVID-19, Amino acids in the active site are presented in wireframe with element colour (where carbon is grey, oxygen is red, nitrogen is blue and sulphur is yellow and hydrogen in white) and ligand is presented in Stick lines with fix colour green. Blue lines represent the hydrogen bonds between the ligand and the active site of Covid-19 Mpro.

Fig. 3: Structural cartoon of COVID-19 (PDB code 6LU7), The α helices and β strands are represented as coils (red) and arrows (blue), respectively. A2 compound is represented in ball and stick (yellow)

The best docking poses obtained on the basis of MVD re-rank score for curcumin and top three ligand A2, B2 and B3 with the crystal structures of COVID-19 are presented in Figures 4 to 7. In which The best score docking solution of ligands and with the selected crystal structure of COVID-19, Amino acids in the active site are presented in wireframe with element colour (where carbon is grey, oxygen is red, nitrogen is blue and sulphur is yellow and hydrogen in white) and ligand is presented in Stick lines with fix colour green. Blue lines represent the hydrogen bonds between the ligand and the active site of Covid-19 Mpro.

Fig. 4: Curcuminligand docked with Covid-19 main protease (Mpro)
Fig. 5: A2 ligand docked with Covid-19 main protease (Mpro)

Fig. 6: B2 ligand docked with Covid-19 main protease (Mpro)

Fig. 7: B3 ligand docked with Covid-19 main protease (Mpro)
Docking studies of curcumin with COVID-19 showed the presence of hydrogen bonding between these compounds with the proteins of ER-positive. The curcumin showed three hydrogen bonding interactions with the active site of COVID-19 residues includes Thr 190, Gln 192, and Leu 141 (Fig 8a). The hydrogen bond interaction pattern was evaluated by atom ‘O’ of Thr 548 interact with atom ‘O18’of ligand, atom ‘N’ of Gln 192 interact with atom ‘O18’ of ligand, and atom ‘O’ of Leu 141 interact with atom ‘O27’of ligand. Docking simulations using MVD revealed MolDock score of -151.545 Kcal/mol and a re-ranking score of -103.901 Kcal/mol for Curcumin.

In the case of A2 ligand, the results showed six hydrogen bonds between the protein and ligand atoms. The interaction residues include Arg 188, Gln 192, Thr 190, Gly 143, Ser 144, and Leu 141 (Fig 8b). The hydrogen bond interaction pattern was evaluated by atom ‘O’ of Arg 188 interact with atom ‘O24’of ligand, atom ‘N’ of Gln 192 interact with atom ‘O24’ of ligand, atom ‘O’ of Thr 190 interact with atom ‘O25’of ligand, atom ‘N’ of Gly 143 interact with atom ‘O43’ of ligand, atom ‘N’ of Ser 144 interact with atom ‘O43’ of ligand and atom ‘O’ of Leu 141 interact with atom ‘O43’ of a ligand. Besides, the A2 ligand was observed with steric interactions and residue Arg 188, Glu 166, His 164, Gly 143, Cys 145and Met 49 involved in interaction. The docking results A2 ligand with COVID-19 reveals no electrostatic interactions but it has a hydrogen bonding and steric interaction between the ligand to an amino acid of Protein. A2 ligand has the lowest moldock score is -202.476 kcal/mol.

B2 and B3 are the amino methyl derivatives of mono-carbonyl analogues of curcumin [16]. In the case of B2 ligand, the results showed six hydrogen bonds between the protein and ligand atoms. The interaction residues include Ser 144, Leu 141, Cys 145, Gly 143, Gln 192, and Thr 190 (Fig 8c). It is revealed that the atom ‘N’ of Ser 144 interact with atom ‘O23’ of ligand, atom ‘O’ of Leu 141 interact with atom ‘O23’ of ligand, atom ‘N’ of Cys 145 interact with atom ‘O23’ of ligand, atom ‘N’ of Gly 143 interact with atom ‘O38’ of ligand and atom ‘O’ of Thr 190 Interact with atom ‘O40’ of a ligand. As compared to hydrogen bonding interaction this compound more favourable with steric interaction. For B2 steric interaction residues involved Cys 145, Gln 189, Glu 166, Leu 167, Gln 192, and Met 165. Docking simulations using MVD revealed MolDock score of -189.464 Kcal/mol and Re-ranking score of -84.176 Kcal/mol.

In the case of B3 ligand, the results showed three hydrogen bonds between the protein and ligand atoms. The interaction residues include Gln 192, Thr 190, Leu 141, Ser 144, and Gly 143 (Fig 8d), the atom ‘N’ of Gln 192 interact with atom ‘O13’ of ligand, atom ‘O’ of Thr 190 interact with atom ‘O15’ of ligand, atom ‘O’ of Leu 190 interact with atom ‘O30’ of ligand, atom ‘O’ of Ser 144 interact with atom ‘O30’ of ligand and atom ‘N’ of Gly 143 interact with atom ‘O28’ and ‘O30’of ligand. In addition, the B3 ligand was observed with 6 steric interactions were observed and residues Met 165, Thr 190, Asn 142, Glu 166, His 164 and Cys 145 involved in the interaction.

Ligand A2 has the lowest Moldockscore, Re-rank score, Hydrogen binding energy, and steric interaction energy as compared to other ligands, with COVID-19 Mpro. These all ligand well fit in the active site of COVID-19. Figure 9 show hydrogen bonding with COVID-19 and Figure 9 show steric interaction with COVID-19. Fig 10 showed the comparison of the binding of Curcumin, A2, B2 and B3 ligands in the active site of Covid-19 main protease(Mpro) (PDB ID :6LU7). It showed clear binding of Curcumin, A2 ligand, B2 ligand and B3 ligand.

![Curcumin](image1)
![A2 ligand](image2)
![B2 ligand](image3)
![B3 ligand](image4)

**Fig.8 (a-d):** Hydrogen bond interaction with Covid-19 of Curcumin (7a), A2 ligand (7b), B2 ligand (7c) and B3 ligand (7d).
4. CONCLUSION

To wrap up the discussion, it can be stated the rapidly spreading outbreak of COVID-19 and still steadily growing due to its quick human to human transmission and it is challenged the healthcare area of the world in the last few months. Throughout, virtual screening based on molecular docking was performed to identify novel compounds having the potential to bind the main protease (Mpro) (PDB ID 6LU7).
of COVID-19. Based on our molecular docking results, it can be concluded that the compounds investigated can interact with important amino acids of COVID-19 Mpro. According to moldock binding score all analogues of curcumin bearing good binding potency against COVID-19 due to its biological properties as antiviral. In which A2 ligand show lowest binding energy as compared to other compounds as well as compounds fit well in the active site of COVID-19 Mpro and also interact with the residues in the active site are essential for their biological activity. Henceforth, analogues of curcumin compounds could be a potential inhibitor of the main protease of COVID-19 and might be used as antiviral drug candidates. Moreover, further studies should be conducted for the validation of these compounds using in vitro and in vivo models which will be helpful for new drug discovery against corona virus.

5. REFERENCES