Taylor & Francis
Taylor & Francis Group

Check for updates

Bacillus species; a potential source of anti-SARS-CoV-2 main protease inhibitors

Sadia Alam[a], Shahida Sadiqi[b], Maimoona Sabir[a], Sobia Nisa[a], Sajjad Ahmad[c] and Sumra Wajid Abbasi[d]

[a]Department of Microbiology, The University of Haripur, Haripur, Pakistan; [b]Department of Microbiology, Hazara University, Mansehra, Pakistan; [c]National Center for Bioinformatics (NCB), Quaid-i-Azam University, Islamabad, Pakistan; [d]NUMS Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan

Communicated by Ramaswamy H. Sarma

ABSTRACT

The COVID-19 being a preconized global pandemic by the World Health Organization needs persuasive immediate research for possible medications. The present study was carried out with a specific aim to computationally evaluate and identify compounds derived from *Bacillus* species as the plausible inhibitors against 3-chymotrypsin-like main protease (3CLpro) or main protease (M^{Pro}), which is a key enzyme in the life-cycle of coronavirus. The compounds were isolated from the crude extracts of *Bacillus* species. Among the isolated compounds, novel inhibitory leads were identified using *in silico* techniques. Molecular docking revealed that stigmasterol (-8.3 kcal/mol), chondrillasterol (-7.9 kcal/mol) and hexadecnoic acid (-6.9 kcal/mol)) among others bind in the substrate-binding pocket and also interacted with the catalytic dyad of the 3-CLpro. Further evaluation using 50 ns molecular dynamic simulation and MMPB-GBSA indicated that among the top three docking hits, hexadecanoic acid was found to be the most promising anti-COVID-19 lead against the main protease. Hexadecanoic acid might serve as a potent anti-SARS-CoV-2 compound to combat COVID-19, however, *in vitro* and *in vivo* validation and optimization is needed.

Abbreviations: 3CLpro: 3C-Like Protease; 3D: Three Dimensional; AMCs: Antimicrobial Compounds; COVID-19: Coronavirus Disease 2019; EGB: Polar Solvation Energy for GB; Ele: Electrostatic; ESURF: Non Polar Solvation Energy; GB: Generalized Born; GC-MS: Gas Chromatography–Mass Spectrometry; MD: Molecular Dynamics; MERS-CoV: Middle East Respiratory Syndrome Coronavirus; MMGBSA: Molecular Mechanics Generalized Born Surface Area; MMPBSA: Molecular Mechanics Poisson–Boltzmann Surface Area; NHBs: Number of Hydrogen Bonds; NPT: Number of Particles, Pressure and Temperature; NVT: Number of Particles, Volume and Temperature; PB: Poisson–Boltzmann; PDB: Protein Databank; Rg: Radius of Gyration; RMSD: Root Mean Square Deviation; RMSF: Root Mean Square Fluctuation; SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2; UCSF Chimera: University Of California, San Francisco; vdwAALS: Van der Waals; VMD: Visual Molecular Dynamics; WHO: World Health Organization; ΔG_{gas}: Delta Gas Phase Energy; EPB: Polar Solvation Energy for PB; ENPOLAR: Non Polar Solvation Energy for PB; Δtot: Total Energy

1. Introduction

A novel strain of the coronavirus SARS-CoV-2 was detected on December 29, 2019, in Wuhan city of China, which affects the lower respiratory tract and causes severe acute respiratory syndrome in patients (Adhikari et al., 2020). WHO declares coronavirus infection (COVID-19) as a public health emergency of international concern, affecting millions of people, and it is spreading worldwide with a mortality rate of approximately 3% (https://www.who.int/emergencies/diseases/novel-coronavirus-2019). COVID-19 has clinically significant characteristics that cause certain disease conditions such as; pneumonia, fever, difficulty in breathing, infection of the lung. Coronaviruses (CoVs) are enveloped, single-stranded RNA viruses that infect large number of vertebrates. CoVs species are zoonotic and transmitted to humans via droplets of different sizes. A total of six coronavirus species are responsible for

inducing clinical complications in humans. Four out of six named as; 229E, OC43, NL63, and HKU1 are most common and generally cause common cold in immune-compromised individuals (Liu et al., 2020). The remaining two strains named; Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus (SARS-CoV) are zoonotic and cause severe clinical complications (Cui et al., 2019). The 3 C-like protease (3CLpro) along with papain-like protease(s) is the main protease enzyme that is significant for the processing of the viral polyproteins (John et al., 2015). This role of 3CLpro in duplication of the virus makes it an ideal drug target for the identification and designing of potent anti-SARS-CoV-2 inhibitors.

Microorganisms are the most abundant living entity on the earth's surface, which interacts with other organisms that flourish in the biosphere. Both of them create an interactive network that constitutes the basis for life on the planet Earth

(Cani & Knauf, 2016; Nicholson et al., 2012; O'Mahony et al., 2015). Bioactive compounds produced by microorganisms are quite significant for developing novel drugs and diverse substances with various biological activities, including antifungal, antibacterial, anti-inflammatory, anticancer, and antiviral. These new substances and novel drugs are cost-effective and highly efficient (Lavecchia & Di Giovanni, 2013; Tripathi & Misra, 2017). Bacteria of the Bacillus genus are gram-positive, rod-shaped, and can form endospores. Due to their diversity in physiological properties, they have the ability to produced plethora of antimicrobial compounds (AMCs) (at least 4–5% of its genome) to favour their ubiquitous distribution in soil and environment (Caulier et al., 2019).

Keeping this in mind, the current study aimed to investigate the antiviral efficacy of crude extract compounds produced by soil Bacillus species as 3CLpro-inhibitors. This study involves molecular docking and molecular dynamics (MD) simulation analysis to evaluate the potential of secondary metabolites identified through GC-MS in extracts of *Bacillus* species as potential drugs against COVID-19 and as significant standardization markers.

2. Methodology

2.1. Soil sample collection and preparation

Soil samples were collected from different regions of Khyber Pakhtunkhwa (Haripur, Swat, and Charsada). By using sterile polythene bags approximately 4 g of soil samples were collected. All the samples were shifted to the lab under sterile conditions for further processing as previously described by McPherson et al. (2018).

2.2. Isolation of bacillus species

From each sample, one gram of soil was suspended in sterile distilled water (9 ml volume) and strongly shaken for 2-3 min for dissolution. Sterile normal saline (0.85%) was used for serial dilution of soil suspension. Then serial dilutions up to 10^{-6} were prepared using sterile normal saline. From each dilution, an aliquot of 0.1 ml was taken and lawned on Petri plates containing isolation media. Nutrient agar, GYMS medium, and Tryptic Soy Broth with an additional 0.5% yeast extract (modified Tryptic Soy Broth, pH 7.4 at 37 °C under shaking condition) used for the isolation of bacterial colonies from soil samples. The plates were incubated at 30 °C and monitored after 24 h before the selection of the bacterial colonies (McPherson et al., 2018).

Plates were then examined and suspected colonies were Gram staining stained. The rod-shaped, spore-forming gram-positive bacilli were selected for additional identification tests. Consecutive identification tests were done via sugar fermentation, starch hydrolysis, motility, Indole production, Voges-Proskauer VP, nitrate reduction, catalase, and gas production.

2.3. Bacterial extracts preparation and compounds separation

Selected bacterial isolates were inoculated in Erlenmeyer flask of Nutrient broth followed by incubation for 3 to 4 days. Afterward, the fermentation flask was incubated in a rotary shaker at 110 rpm for 7 days. Secondary metabolites were excreted, the broth culture was filtered and added with solvent ethyl acetate. The solvent layer was separated, stored in a sterile flask to allow evaporation of ethyl acetate extract using a rotary evaporator. Before further use, the extract was stored at 4–5 °C.

2.4. Bacterial metabolites analysis by gas Chromatography-Mass spectrometry

Crude extract of each sample was used to investigate their chemical composition by using GC–MS, which has high sensitivity and specificity as previously described by Begum et al. (2016). GC-MS was performed in a Thermo Scientific GC Focus Series DSQ. Helium gas was used as carrier keeping the flow rate constant (1 ml/min) with 1ul injection volume. A 22-min run was accomplished and injector temperature was maintained at 250 C, the oven temperature was set at110 C, with an increase of 10 C/min to 200 C, after that 5 C/min to 280 C, closing with 9 min isothermal at a temperature of 280 C.

Peaks responding to the compounds were eluted from the GC column and recorded along with retention time. The data obtained were compared with mass spectra, and the database was sought for homologous compounds of the same molecular mass and retention time. Biological activities of the reported natural compounds were also studied, and comparability was made to correlate the activities of bacterial extracts with their components.

2.5. In silico analysis

2.5.1. Preparation of ligands and target protein
For this study, compounds with a strong manifestation of biological activities were selected. Chemical structures of selected compounds were drawn using Chem Office 2004 (Li et al., 2004). After that, the structures were converted into. mol2 format using Chem3D Ultra. UCSF Chimera 1.14 was used for structural geometry optimization of the compounds (Pettersen et al., 2004). The three dimensional (3 D) structure of the 3CLpro (PDB ID:6LU7) was retrieved from the protein data bank (https://www.rcsb.org/structure/6LU7). Crystallographic water molecules along with co-crystallized have been removed from the PDB file of the target.

2.5.2. Molecular docking
The docking study was accomplished by using Autodock/vina at the active site of 3CLpro to explain the mechanism of inhibition of selected compounds at the molecular level (Trott & Olson, 2010). Preparatory steps were followed as mentioned; addition of polar hydrogens to the protein structure, checking for missing residues, generation of pdbqt files

for both ligands and receptor along with the setting of grid box around the active site was accomplished by using GUI program of Auto Dock Tool (Morris et al., 2009). The grid size dimensions were allowed as $40 \times 40 \times 40$ along X, Y, and Z axes with 0.375 Å spacing. The pdbqt files with grid box properties were recorded into a configuration file. During the docking run, both protein and ligand were treated as rigid entities. In positional RMSD, the results of less than 1.0 Å were clustered and were characterized in the most conducive free energy of binding. For further analysis, a low binding affinity pose has been extracted and aligned with the protein. The docking protocol was validated by docking co-crystallized N inhibitor at the active site of 3CLpro using the same parameters described above and the predicted pose was found the same as reported in the crystal structure thus validating the docking procedure.

2.5.3. MD Simulations

Dynamics of selected docking complexes, 50 ns MD simulation was carried out using AMBER 18 software (Case et al., 2010) using a similar protocol as defined in our previous studies (Abbasi et al., 2016; Wahedi et al., 2020). The simulation trajectories were further assessed by analyzing various physical properties: Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), and Gyration Radius (Rg). Additionally, hydrogen bond analysis using VMD (https://www.ks.uiuc.edu/Research/vmd/) was also performed.

2.5.4. Binding free energies of the complex

The free energy of the binding of the docked complex was estimated using the methodology reported by Abro and Azam (2016).

3. Results and discussion

3.1. Gc-MS analysis

The results of GC-MS evaluation indicate the presence of several compounds in crude extracts of *Bacillus* species. The most important highest components exhibited in the crude extract analyzed with the help of GC-MS are documented in Tables 1–3.

The crude extracts of Bacillus species show the presence of bioactive compounds and these compounds are alkaloid, steroid, saturated fatty acid, and lipid in nature. Different compounds have been found to possess a wide range of activities. The compounds are known to have anti-bacterial, anti-fungal, anti-viral, anti-ulcer, anti-cancer, anti-oxidant, anti-arthritic, anti-asthma, anti-psychotic, nematocidal, and pesticide activity.

3.2. Molecular docking analysis

A molecular docking study was conducted to better understand the mechanism of binding of compounds to SARS-CoV-2 3CLpro at the molecular level. To predict their binding affinities and modes, interactions between SARS-CoV-2

3CLpro and compounds were analyzed. Top-ranked compounds, based on binding affinities (chondrillasterol −7.9 kcal/mol, cholestan −7.8 kcal/mol, trifluoroacetic acid −7.7 kcal/mol, octadeccenoic-acid −9.9 kcal/mol, stigmasterol −8.3 kcal/mol, 9 octsdecenoic-acid −9.7 kcal/mol, and hexadecnoic acid −6.9 kcal/mol) were selected for further *in silico* analysis. Visual exploration of the binding patterns of the aforesaid compounds identified three bioactive compounds: stigmasterol (-8.3 kcal/mol), chondrillasterol (-7.9 kcal/mol) and hexadecnoic acid (-6.9 kcal/mol)) that are hypothesized and predicted to fit inside the substrate-binding pocket, and also interact with the catalytic dyad comprising of His41 and Cys145 (Figure 1) (Jin et al., 2020; Macchiagodena et al., 2020).

Stigmasterol is produced by different species . of *Bacillus* and have anti-inflammatory activity (Zeb et al., 2017). The molecular interaction studies of stigmasterol showed that it was not only able to bind to the 3CLpro's catalytic dyad (HIS41-CYS145) but also exhibit the affinity for the residues of substrate binding site: MET49, PHE140, LEU141, ASN142, GLY143, SER144, MET165, GLU166, LEU167, PRO168 and GLN 189 (Figure 2). The 3 D analysis of the docked complex using VMD revealed that catalytic dyad HIS41, CYS145 along with other important residues (SER144, GLY143, and LEU141) played a vital role in stabilizing the docked conformation by making eight hydrogen bonds of 2.44 Å, 3.39 Å, 2.73 Å, 2.69 Å, 3.84 Å, 3.25 Å 3.78 Å and 3.87 Å, respectively with the electronegative as well as hydrogen atoms of stigmasterol. Additionally, the carbonyl atoms of ligand and carbon atoms of important interacting residues of 3CLpro, namely: GLU166, MET165, GLN189, THR190, and ALA191 produced a network of hydrophobic interactions (Table 4 and Figure 3a). Two dimensional (2 D) interactions map generated using complex using Discovery studio highlighted additional interactions not evident from the results viewed using VMD. As shown in Figure 3b, HIS41, MET49, LEU141, ASN142, GLY143, HIS164, MET165, and GLU166 were involved in van der Waals interaction. An alkyl interaction with PRO168 and two pi-alkyl interaction with CYS145 and HIS163 were also observed.

Chondrillasterol, which is a lipid sterol and has anti-proliferative and antimicrobial properties (Mozirandi et al., 2019), also binds with the key residues of the active site namely: HIS41, MET49, PRO52, TYR54, PHE140, LEU141, ASN142, CYS145, HIS164, MET165, GLU166, ASP187, ARG188, GLN189 (Figure 4). ARG188, ASP187, and MET49 were involved in the hydrogen bond interactions. ARG188 was also involved in an ionic interaction of 3.35 Å. LEU141, GLU166, MET165, GLN189, and HIS41 were involved in the hydrophobic contacts (Table 4 and Figure 5a). No notable interactions with CYS145 have been observed within 4 Å. 2 D interaction map illustrated that ASN142, GLU166, GLN189, ARG188, ASP187, MET49, and HIS 164 were involved in van der Waals interaction (Figure 5b). A conventional hydrogen bond between TYR54 and the oxygen atom of the compound was observed. A pi-alkyl interaction with the residue of the catalytic dyad, HIS41, and alkyl interactions with LEU141 and MET 165 were also observed.

Table 1. Major compounds identified in methanolic extract of H1 (*Paenibacillus dendridiformis*).

S. No	Name of the Compound	Retention Time (min)	Molecular Mass	Molecular Formula	Pharmacological actions	CAS#	Library	Probability	Percentage of Area
1.	Oxalic Acid	0.67	90 g/mol	C2H2O4	use as Pesticides	144-62-7	mist-msms	18.77	71.11%
2.	9-Octadecenoic acid(z)-, 2-hydroxyl-1,3-propanedyl ester	1.98	620 g/mol	C39H72O5	Anti-bacteria Anti-Fungal	2465-32-9	MAIN LIB	6.38	0.08%
3.	Cholestan- 3-ol,2-methylene (3a,5a)	400	400 g/mol	C28H48O	Antibacterial Anti-ulcer Analgesic Anti-cancer		MAIN LIB	12.09	0.39%
4.	Naphthalene,1,2,3,4-tetrahydro-6-methyl	7.78	146 g/mol	C11H14	Insecticide	1680-51-9	MAIN LIB	27.71	0.88%
5.	Stearic acid	19.42	594 g/mol	C39H78O3	Antimicrobial Antioxidant Hypoglycemic and Thyroid Inhibiting properties, act as a precursor of Progesterone, Anti-Cancer Anti-arthritic Anti-asthma Anti-inflammatory diuretic.	17367-40-7	MAINLIB	10.69	0.02%
6.	Chondrilla sterol	20.88	412 g/mol	C29H48O	Antimicrobial	481-17-4	MAINLIB	19.83	0.24%
7.	Propanoic Acid	0.67	144 g/mol	C8H16O2	Antimicrobial	97-87-0	replib	18.04	71.11%
8.	Trifluoracetic Acid	12.02	366 g/mol	C20H37F3O2	Rodenticide Preservative Anti-Cancer	79392-43-1	MAINLIB	4.99	0.31%
9.	Stigmasterol	20.88	412 g/mol	C29H48O	Antimicrobial	83-48-7	replib	13.2	0.24%

Table 2. Major compounds identified in methanolic extract of H2 (*Bacillus subtilis*).

S. No	Name of the Compound	Retention Time (min)	Molecular Mass	Molecular Formula	Pharmacological actions	CAS#	Library	Probability	Percentage of Area
10.	Octadeccenoic Acid	3.42	884 g/mol	C57H104O6	Antibacterial Anti-fungal	537-39-3	MAINLIB	41.23	0.16%
11.	Hexadecanoic Acid	13.3	256 g/mol	C16H32O2	Antimicrobial Anti-oxidant Hypo-cholesterolemic Nematicide Anti-androgenic Hemolytic Pesticide Lubricant 5-alpha reductase inhibitor Anti-psychotic	57-10-3	MAINLIB	21.83	0.99%
12.	Aspidospermidin	9.49	414g/mol	C23H30N2O5	AdrenergicblockageActivities	2122-26-1	MAINLIB	13.09	0.19%
13.	Steric Acid	17.66	594g/mol	C39H78O3	Anti- Microbial	17367-40-7	MAINLIB	10.79	0.13%
14.	Phenol	4.65	94g.mol	C6H6O	Anti- Microbial	108-95-2	replib	67.81	13.52%

Table 3. Major compounds identified in methanolic extract of H3 (*Brevibacillus formsus*).

S. No	Name of the Compound	Retention Time (min)	Molecular Mass	Molecular Formula	Pharmacological actions	CAS#	Library	Probability	Percentage of Area
15.	1,3,5-Trioxane	0.71	90g/mol	C3H6O3	Anti-malarial	110-88-3	MAINLIB	83.68	61.30%
16.	Cyclobutane	1.33	92g/mol	C7H8	Anti-Microbial Anti-bacterial Anticancer	52097-85-5	MAINLIB	18.99	0.95%
17.	Phenol	9.96	206g/mol	C14H22O	Anti-bacterial	96-76-4	MAINLIB	51.04	1.87%
18.	Dasycarpidan	19.41	326g.mol	C20H26N2O2	Anti-bacterial Anti-fungal Anti- Diabetic Anti-Cancer	55724-48-6	MAINLIB	43.35	0.14%

Like the aforesaid two compounds, hexadecanoic acid also exhibited an affinity for the residue of the catalytic dyad and interacting residues of the active site as shown in Figure 6. The 3 D interactions map generated for hexadecnoic acid-3CLpro docked complex revealed that hexadecnoic acid occupied the substrate-binding site through a network of hydrogen bond interactions between the hydrogen atoms and the electronegative atoms of the interacting residues (GLY143, SER144, HIS163, and LEU141) and the studied compound. It was observed that the sulfur atom of catalytic site
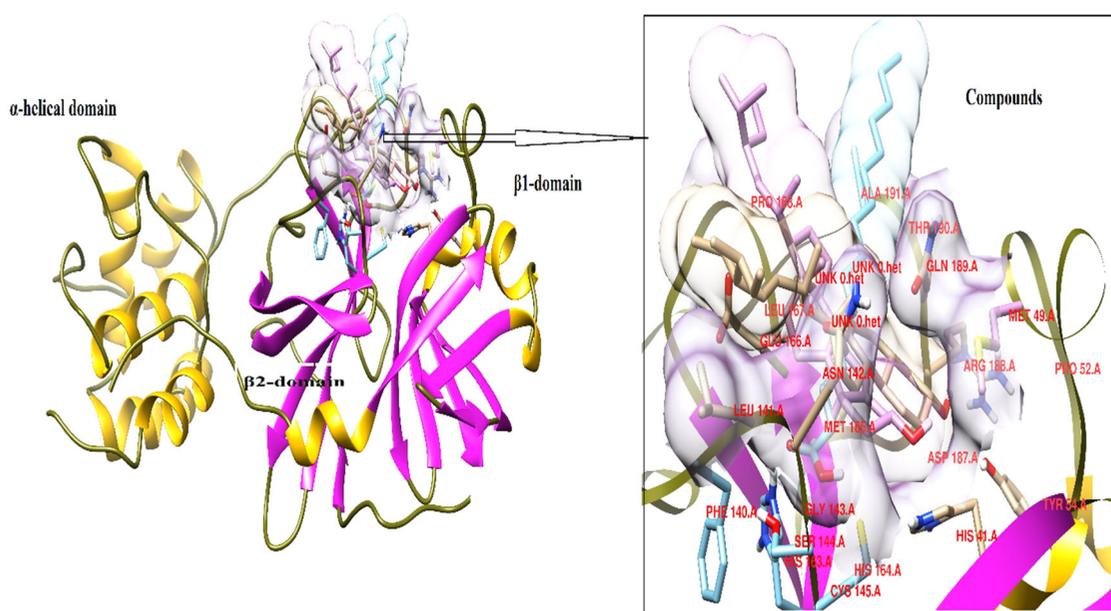
Figure 1. Predicted conformations of selected three potential compounds into the substrate-binding pocket of 3CLpro (comprising of three domains). Key residues of the binding pocket are in red.
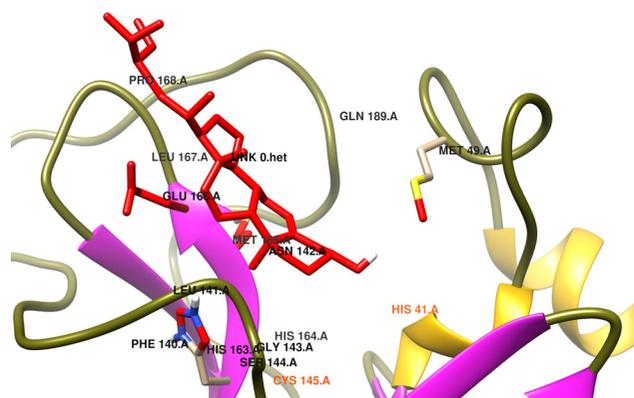
Figure 2. Illustrating the active site residues of best-docked conformation selected for stigmasterol-3CLpro, visualized using Chimera. Residues of substrate binding sites are shown in black color whereas catalytic dyad residues are shown in reddish-orange color.
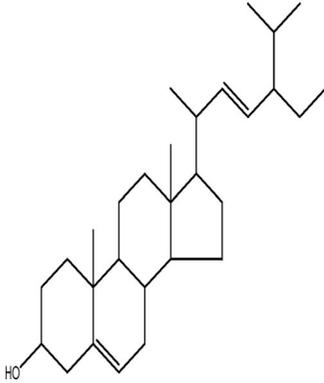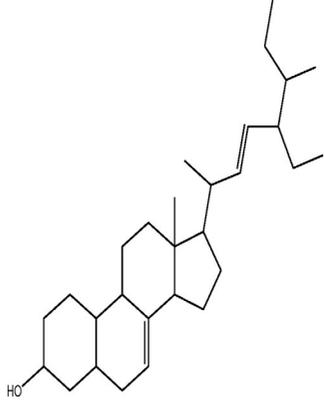
CYS145 along with the amine group also participated in the formation of two hydrogen bonds (CYS145: SG‐‐UNK: H 2.97 Å, CYS145: HN‐‐UNK: O 2.85 Å). Anionic interaction between the nitrogen atom of CYS145 and the oxygen atom of a small compound at a distance of 2.26 Å was also observed. HIS163 and SER144 also interacted with Hexadecnoic-acid through ionic interactions. ALA191, THR190, and GLN 189 were involved in the hydrophobic contacts (Table 4 and Figure 7a). The van der Waals interactions, conventional hydrogen bond interaction, as well as alkyl interactions, were observed using discovery studio (Figure 7b).
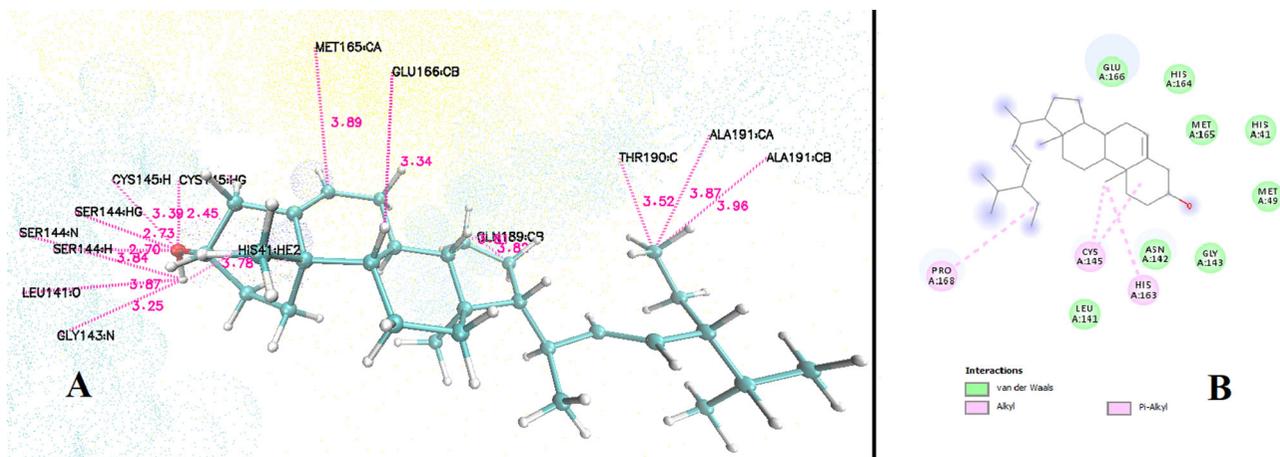
The aforementioned screened compounds not only exhibited good binding affinities but also showed hydrogen bond interactions with the catalytic dyad and other important residues of pocket, recommending these compounds for further atomic level optimization through molecular dynamics simulation.

3.3. MD simulations

We performed MD simulations to validate the docking findings and to further analyze the binding pattern and stability of potential compounds. Also, the docked complexes were rescored using MMPBSA to verify the docking results and to rank the binding affinities accordingly. Based on the results of MMPBSA, it was concluded that hexadecnoic acid-3CLpro is the most stable complex for further study compared to stigmasterol-3CLpro and chondrillasterol-3CLpro docked complexes with better docking affinities. The estimated average potential energy for the hexadecnoic acid-3CLpro complex after 50 ns was −144302 kcal/mol which revealed the complex's stability. The RMSD plot showed that protein attains the equilibrium after 1 ns and remains stable till 50 ns with few minor deviations and an average RMSD value of 1.6 Å (Figure 8a). Snapshots extracted at different intervals revealed that around 10 ns the compound moved a little bit from the initial docked conformation, binds into the substrate-binding pocket, and catalytic dyad (HIS41, CYS145) and stays there till the entire simulation run. Visualization through the using Discovery Studio showed that both the residues of the catalytic dyad (HIS41 and CYS145) along with MET160 were involved in pi-alkyl and alkyl interactions. THR190 and GLN198 were involved in a conventional hydrogen bond as well as a carbon-hydrogen bond, respectively. Thirteen residues, including the residues of substrate binding site (THR25, THR26, LEU27, ASN28, TYR54, ASN119, ASN142, GLY143, SER144, HIE164, ASP187, GLN189, and ALA191) were interacting using van der Waals interactions to form a stable complex (Figure 9a and b). Most of these residues have been reported as crucial because of their involvement in the formation of the substrate-binding pocket (Jin et al., 2020). The plot for the radius of gyration also revealed that the system attained compactness and stability (22.2 Å) and the results achieved are consistent with the results of RMSD stated herein (Figure 8b). The RMSF analysis was conducted to

Table 4. Three bioactive compounds along with their respective structures, binding affinity, and key interactions with the target protein.

Compound Name	2D Structure	Binding Affinity (kcal/mol)	Interaction Type		
			Hydrogen Bond (Distance in Å)	Ionic Interactions (Distance in Å)	Hydrophobic Contacts (Distance in Å)
Stigmasterol		−8.3	**CYS145**:HG—UNK:O 2.44 **CYS145**:H—UNK:O 3.39 SER144:HG—UNK:O 2.73 SER144:H—UNK:O 2.69 SER144:N—UNK: H 3.84 GLY143:HN–UNK:O 3.25 LEU141:O—UNK:H 3.87 **HIS41:HE2**—UNK:O 3.78	—————	ALA191:CB—UNK:C 3.96 ALA191:CA—UNK:C 3.87 THR190:C—UNK:C 3.52 MET165:CA—UNK:C 3.89 GLN198:CB—UNK:C 3.81 GLN198:CB—UNK:C 3.82 GLU166:CB—UNK:C 3.34
Chondrillasterol		−7.9	ARG188:N–UNK:H 2.77 ASP187:O–UNK:H 2.42 MET49:O–UNK:H 3.90	ARG188:N–UNK:O 3.35	LEU141:CD2–UNK:C3.98 LEU141:CB–UNK:C 3.93 GLU166:CD–UNK:C3.96 GLU166:CB–UNK:C3.60 GLU166:CB–UNK:C3.58 MET165:CA–UNK:C3.95 MET165:CB–UNK:C3.75 MET165:CB–UNK:C3.56 MET165:CB–UNK:C3.53 MET165:CB–UNK:C3.87 MET165:CB–UNK:C3.33 GLN189:CG–UNK:C3.39 GLN189:CG–UNK:C3.16 GLN189:CG–UNK:C3.56 GLN198:CD–UNK:C3.74 **HIS41**:CE1–UNK:C3.67 **HIS41**:CG–UNK:C3.46 **HIS41**:CD2–UNK:C3.72
Hexadecanoic acid		−6.9	GLY143:HN–UNK:O 3.28 GLY143N–UNK:H 3.68 SER144:HN–UNK:O 2.60 HIS163:HE2–UNK:O 2.02 HIS163:HE2–UNK:O 3.66 **CYS145**:SG–UNK:H 2.97 **CYS145**:HN–UNK:O 2.85 SER144:HG–UNK:O 3.06 SER144:HG–UNK:O 3.44 LEU141:O–UNK:H 3.59	SER144:N–UNK:O 3.32 HIS163:NE2-UNK:O2.92 **CYS145**:N–UNK:O 2.26 GLY143:N–UNK:O 3.78	ALA191:CA–UNK:C3.94 THR190:C–UNK:C3.43 THR190:C–UNK:C3.74 GLN189:CA–UNK:C3.80 GLN189:CD–UNK:C3.75 GLN189:CD—UNK:C3.76

Figure 3. Docked poses of Stigmasterol with 3CLpro main protease. (a) 3 D representation of interactions using VMD, protein is represented in dotes style whereas compound is in CPK style. (b) 2 D representation of key interactions using Discovery studio.

explore the per residue fluctuations of the system. The binding of Hexadecnoic-Acid has been shown to make proteins less flexible, especially in regions containing catalytic dyad and interacting residues of pocket. The average RMSF for the system was 1.02 Å, especially for HIS41 and CYS145 it was estimated as 0.50 Å and 0.71 Å (Figure 8c). For the residue of domain III, especially residues 300-303, a higher fluctuation of approximately 4.66 Å was observed. The possible reason for this higher fluctuation could be the involvement of domain III in the formation of a homodimer. RMSD plot for ligand showed few deviations with an average RMSD of 2.05 Å (Figure 10a). When snapshots for the ligand were analyzed, it was observed that the deviations in the RMSD plot were because of the flexibility of the carboxyl group. As shown in Figure 10 b at 10th, 15th, and 45th ns this group showed a major deviation from the initial conformation. The chain of carbon atoms remained stable throughout the simulation run playing an anchoring role in the binding of the compound to the protein. This finding could be significant for the future design of potent compounds that use this compound as the lead. The compound's flexible moiety may be substituted by another group that may limit its movement and the ability to increase the potency that optimally inhibits the 3CLpro. Over the entire simulation run, the complex remains stable.

3.3.1. Hydrogen bonds analysis

Hydrogen bonds play a vital role in maintaining the stability of the docked complex, so the hydrogen bond calculation

significantly highlights the stability of the studied complex (Gerlt et al., 1997). To gain insight into the stable complex, it is therefore essential to assess the number of hydrogen bonds between protein-ligand interacting regions. The hydrogen bond plot for hexadecnoic acid-3CLpro docked complex generated by VMD is shown in Figure 11. The figure shows that, on average, two hydrogen bonds were observed between the interactive residues and Hexadecnoic-Acid, during the entire MD spam. The results are consistent with the MD predictions.

3.4. Mmpbsa and MMGBSA calculations

Free binding energies were calculated using the MMPBSA and MMGBSA methods for 100 snapshots of the last 10 ns trajectories. Very strong protein-compound interactions have been demonstrated for both approaches. The entropy has been omitted because it cannot be measured in certain situations due to convergence issues. In both techniques, the contribution of gas-phase energy to the overall net energy of the system is considerably high as compared to the net solvation energy of the system, which shows complex stability primarily due to the receptor's best confirmation with respect to the ligand. For both cases, net gas phase energy was −38.29 kcal/mol. Van der Waals interactions

Figure 4. Illustrating the active site residues of best-docked conformation selected for chondrillasterol-3CLpro, visualized using Chimera. Residues of substrate binding sites are shown in black color whereas catalytic dyad residues are shown in reddish-orange color.

Figure 6. Illustrating the active site residues of best-docked conformation selected for hexadecnoic acid-3CLpro, visualized using Chimera. Residues of substrate binding sites are shown in black color whereas catalytic dyad residues are shown in reddish-orange color.

Figure 5. Docked poses of Chondrillasterol with 3CLpro main protease. (a) 3 D representation of interactions using VMD, protein is represented in dotes style whereas compound is in CPK style. (b) 2 D representation of key interactions using Discovery studio.

Figure 7. Docked poses of hexadecnoic acid with 3CLpro main protease. (a) 3 D representation of interactions using VMD, protein is represented in dotes style whereas compound is in CPK style. (b) 2 D representation of key interactions using Discovery studio.
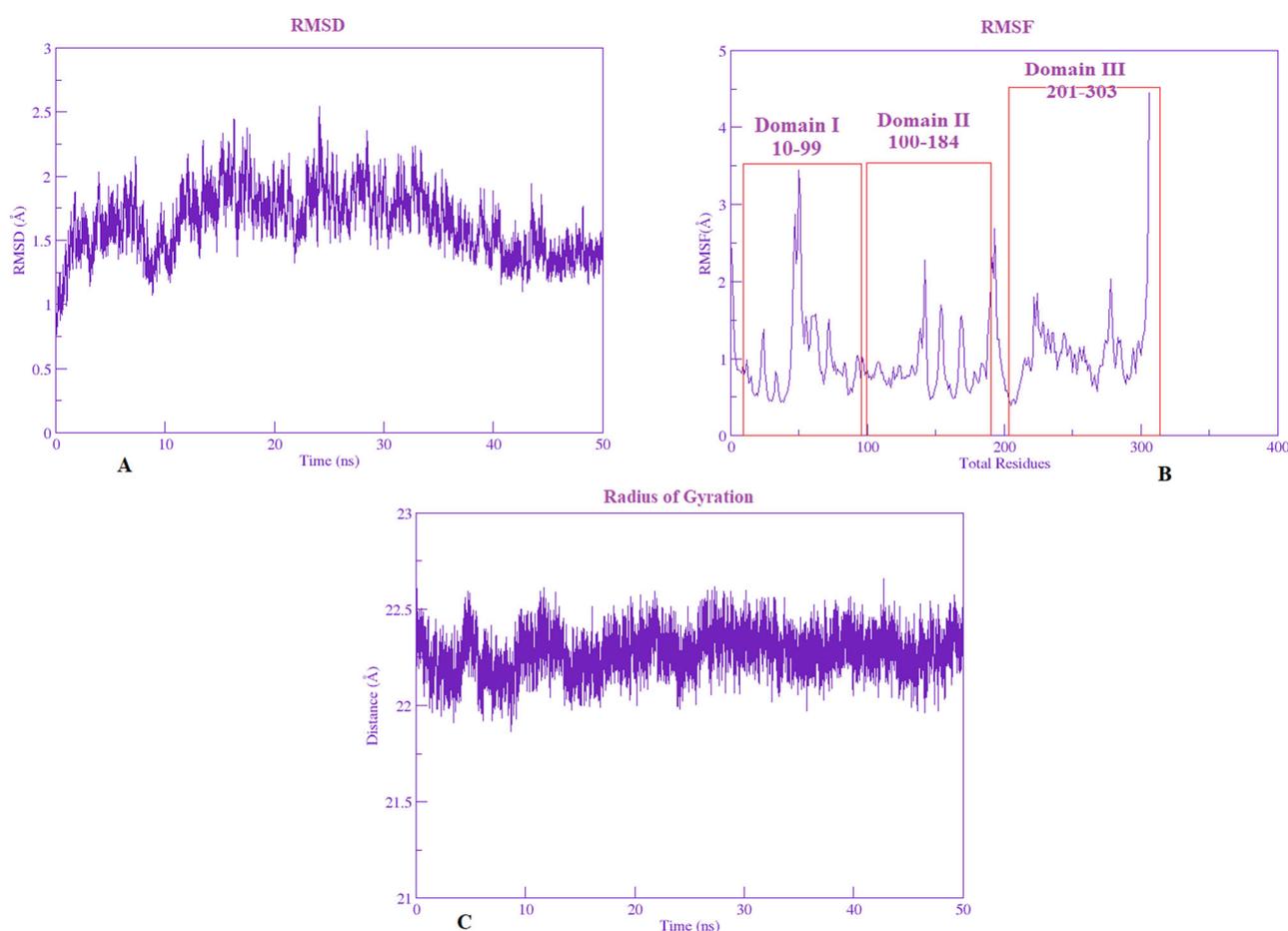
Figure 8. Trajectories analysis vs time. (a) The RMSD profile of backbone atoms of 3CLpro. (b) RMSF profile of C-alpha atoms. (c) The radius of gyration.

(-34.90 kcal/mol) have also contributed favorably in addition to the minor contribution from electrostatic energy towards the gas-phase energy. The total binding free energy calculated for the complex using MMPBSA and MMGBSA methods was -22.42 kcal/mol and -28.67 kcal/mol receptively, shown in Table 5.

4. Conclusions

The accelerated increase of COVID-19 infections and prevailing disease severity compels an urgent need of scheming remedial measures through small compound or peptide drugs to cure the COVID-19. In this regard, using *in silico* approaches, among the nine bacterial compounds extracted from the soil, we identified Hexadecnoic-Acid as a plausible inhibitor against COVID-19 Main Proteases. Hexadecanoic acid is an aliphatic long-chain fatty acid that has shown potential antibacterial and antifungal activity (Begum et al., 2016). It is anticipated that the findings of the current study, after *in vivo* and *in vitro* testing, may prove significant for developing novel antimicrobial therapeutic agents against COVID-19.

Figure 9. Docked Conformation extracted from the trajectory. (a) 2 D depiction of protein-ligand interactions highlighting hotspots. (b) 3 D depiction of interacting residues.

Figure 10. (a) RMSD profile of hexadecnoic acid throughout 50 ns of the simulation run. (b) Superimposed conformations of the compound at different intervals.

Figure 11. Hydrogen bond analysis over the time scale of 50 ns simulations. The presence of hydrogen bond is indicated by black squares whereas white squares correspond to the absence of H-bond.

Table 5. Binding free energy values for the Hexadecnoic-Acid-3CLpro complex.

Energy Component	Value (Kcal/mol)
vdwAALS	−34.90
ele	−3.39
EGB	14.52
ESURF	−4.89
ΔG_{gas},PB	−38.29
ΔG_{solv}, PB	9.62
Δtot,GB	−28.67
vdwAALS	−34.90
ele	−3.39
EPB	19.88
ENPOLAR	−4.00
ΔG_{gas},PB	−38.29
ΔG_{solv}, PB	15.88
Δtot,PB	−22.42

References

Abbasi, S., Raza, S., Azam, S. S., Liedl, K. R., & Fuchs, J. E. (2016). Interaction mechanisms of a melatonergic inhibitor in the melatonin synthesis pathway. *Journal of Molecular Liquids*, 221, 507–517. https://doi.org/10.1016/j.molliq.2016.06.034

Abro, A., & Azam, S. S. (2016). Binding free energy based analysis of arsenic (+ 3 oxidation state) methyltransferase with S-adenosylmethionine. *Journal of Molecular Liquids*, 220, 375–382. https://doi.org/10.1016/j.molliq.2016.04.109

Adhikari, S. P., Meng, S., Wu, Y. J., Mao, Y. P., Ye, R. X., Wang, Q. Z., Sun, C., Sylvia, S., Rozelle, S., Raat, H., & Zhou, H. (2020). Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: A scoping review. *Infectious Diseases of Poverty*, 9(1), 1–12. https://doi.org/10.1186/s40249-020-00646-x

Begum, I. F., Mohankumar, R., Jeevan, M., & Ramani, K. (2016). GC-MS analysis of bio-active molecules derived from paracoccus pantotrophus FMR19 and the antimicrobial activity against bacterial pathogens and MDROs. *Indian J Microbiol*, 56(4), 426–432. https://doi.org/10.1007/s12088-016-0609-1

Cani, P. D., & Knauf, C. (2016). How gut microbes talk to organs: The role of endocrine and nervous routes. *Molecular Metabolism*, 5(9), 743–752. https://doi.org/10.1016/j.molmet.2016.05.011

Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., & Roberts, B.

(2010). AMBER 12; University of California: San Francisco, 2012. There is No Corresponding Record for This Reference, 1–826.

Caulier, S., Nannan, C., Gillis, A., Licciardi, F., Bragard, C., & Mahillon, J. (2019). Overview of the antimicrobial compounds produced by members of the Bacillus subtilis group. *Frontiers in Microbiology*, 10, 302. https://doi.org/10.3389/fmicb.2019.00302

Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature Reviews. Microbiology*, 17(3), 181–192. https://doi.org/10.1038/s41579-018-0118-9

Gerlt, J. A., Kreevoy, M. M., Cleland, W. W., & Frey, P. A. (1997). Understanding enzymic catalysis: The importance of short, strong hydrogen bonds. *Chemistry & Biology*, 4(4), 259–267. https://doi.org/10.1016/S1074-5521(97)90069-7

https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., … Yang, H. (2020). Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 582(7811), 289–293. pphttps://doi.org/10.1038/s41586-020-2223-y

John, S. E. S., Tomar, S., Stauffer, S. R., & Mesecar, A. D. (2015). Targeting zoonotic viruses: Structure-based inhibition of the 3C-like protease from bat coronavirus HKU4-The likely reservoir host to the human coronavirus that causes Middle East Respiratory Syndrome (MERS)). *Bioorganic & Medicinal Chemistry*, 23(17), 6036–6048. https://doi.org/10.1016/j.bmc.2015.06.039

Lavecchia, A., & Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: A critical review. *Current Medicinal Chemistry*, 20(23), 2839–2860. https://doi.org/10.2174/09298673113209990001

Li, Z., Wan, H., Shi, Y., & Ouyang, P. (2004). Personal experience with four kinds of chemical structure drawing software: Review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *Journal of Chemical Information and Computer Sciences*, 44(5), 1886–1890. https://doi.org/10.1021/ci049794h

Liu, Y., Ning, Z., Chen, Y., Guo, M., Liu, Y., Gali, N. K., Sun, L., Duan, Y., Cai, J., Westerdahl, D., Liu, X., Xu, K., Ho, K.-F., Kan, H., Fu, Q., & Lan, K. (2020). Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature*, 582(7813), 557–560. https://doi.org/10.1038/s41586-020-2271-3

Macchiagodena, M., Pagliai, M., & Procacci, P. (2020). Inhibition of the main protease 3cl-pro of the coronavirus disease 19 via structure-based ligand design and molecular modeling. *arXiv Preprint*, arXiv: 2002.09937. https://doi.org/10.1016/j.cplett.2020.137489

McPherson, M. R., Wang, P., Marsh, E. L., Mitchell, R. B., & Schachtman, D. P. (2018). Isolation and analysis of microbial communities in soil, rhizosphere, and roots in perennial grass experiments. *Journal of Visualized Experiments*, (137), e57932. https://doi.org/10.3791/57932

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. https://doi.org/10.1002/jcc.21256

Mozirandi, W., Tagwireyi, D., & Mukanganyama, S. (2019). Evaluation of antimicrobial activity of chondrillasterol isolated from Vernonia adoensis (Asteraceae). *BMC Complementary and Alternative Medicine*, 19(1), 1–11. https://doi.org/10.1186/s12906-019-2657-7

Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., & Pettersson, S. (2012). Host-gut microbiota metabolic interactions. *Science (New York, N.Y.).)*, 336(6086), 1262–1267. https://doi.org/10.1126/science.1223813 (2012).

O'Mahony, S. M., Clarke, G., Borre, Y. E., Dinan, T. G., & Cryan, J. F. (2015). Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behavioural Brain Research*, 277, 32–48. https://doi.org/10.1016/j.bbr.2014.07.027

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Tripathi, A., & Misra, K. (2017). Molecular docking: A structure-based drug designing approach. *JSM Chem*, 5(2), 1042–1047.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. https://doi.org/10.1002/jcc.21334

Wahedi, H. M., Ahmad, S., & Abbasi, S. W. (2020). Stilbene-based natural compounds as promising drug candidates against COVID-19. *Journal of Biomolecular Structure and Dynamics*, 1–10. https://doi.org/10.1080/07391102.2020.1762743

Zeb, M., Khan, S., Rahman, T., Sajid, M., & Seloni, S. (2017). Isolation and Biological Activity of β-Sitosterol and Stigmasterol from the Roots of Indigofera heterantha. *Pharmacy & Pharmacology International Journal*, *5*, 139.