J Young Pharm, 2022; 14(2) : 179-186
A multifaceted peer reviewed journal in the field of Pharmacy
www.jyoungpharm.org | www.phcog.net

Original Article

Lead Identification against 3C-like Protease of SARS-CoV-2 Via Target-based Virtual Screening and Molecular Dynamics Simulation

Afaf S Alwabli
Department of Biological Sciences, Rabigh College of Science and Arts, King Abdulaziz University, Jeddah, SAUDI ARABIA.

ABSTRACT

Background: The COVID-19 pandemic has prompted the human population's emotional, social, and financial loss worldwide and presents an unprecedented challenge to health, food, and working styles. However, the exceptional implication of vaccinations at every nook and corner of the world has been breaking the rate of infection and mortality to a greater extent; in the case of potential oral drugs, we still need to decipher more palliative and therapeutic measures to develop effective antiviral drug candidates. **Materials and Methods:** The study exploits the structure-based virtual screening (SBVS) approach to identify small molecule inhibitors against 3C-like protease ($3CL^{PRO}$) of SARS-CoV-2 from more than five million compounds of the MCULE database. **Results:** Four basic properties viz., molecular weight (≤ 500 g/m), hydrogen bond donor (≤ 5), hydrogen bond acceptor (≤ 10), and logP (≤ 5) as an initial filter were employed in SBVS workflow that extracted 2,235,82 compounds that were subsequently reduced to 22 ligands showing lesser ΔG values than reference drug nirmatrelvir (-7.9 kcal/mol). Upon toxicity check, 10 ligands were obtained that further curtailed to 9 molecules when passed through the BOILED-Egg model of the ADME. Upon compliance of druglikeness other than Lipinski viz., Ghose, Veber, Egan, Muegge, and bioavailability score 7 molecules were shortlisted in which 5 molecules exhibited zero PAINS and Brenk alert. **Conclusion:** At last, only 1 ligand hit (Mcule-3133395989) was identified that obeyed hydrogen bond selection criterion (ligand-$3CL^{PRO}$ complex ≥ 3 HBs). RMSD, RMSF, SASA, ΔGsolv, Rg, and HBs parameters of MD simulations predict Mcule-3133395989 more stable and promising antiviral agent compared to nirmatrelvir.

Key words: $3CL^{PRO}$, SARS-CoV-2, SBVS, Nirmatrelvir, MD simulation.

Correspondence

Dr. Afaf S Alwabli,

Department of Biological Sciences, Rabigh College of Science and Arts, King Abdulaziz University, Jeddah, SAUDI ARABIA.

Email id: aalwabli@kau.edu.sa

DOI: 10.5530/jyp.2022.14.34

INTRODUCTION

One of the most contagious viruses, better known as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), formerly known by the 2019 novel coronavirus (2019-nCoV), has been recognized as an emerging strain added to the Human coronavirus (CoV) family, which causes the most infectious COVID-19 disease affecting our life drastically. There are four genera of CoVs viz., α, β, γ, and δ, in which former two infect mammals while later two causes infection among birds. Apart from middle east respiratory syndrome (MERS-CoV) and SARS-CoV-2, which was respectively responsible for the respiratory infections in China during 2002–2003 and in the Middle East in 2012, four common types of human CoVs, including 229E (α-CoV), NL63(α-CoV), OC43 (β-CoV), and HKU1 (β-CoV) are known contagious strains which may cause mild respiratory infections in healthy individuals.[1-3] As per the Worldometer statistics on January 27, 2022, 363,941,212 COVID-19 cases with 287,993,289 recoveries, and 5,647,818 deaths were recorded worldwide.[4] Wearing the mask, frequent hand wash, alcohol-based sanitization, social distancing, lockdown, and working from home are the new normal in the ongoing pandemics. No one can predict the future despite adopting all COVID-19 safety protocols and measures, healthy lifestyles, suitable medications, and recommended vaccination dosages. Experts believe that the virus changes its nature according to the environment and acquire mutations that may be mild to severe. The USA, India, Brazil, France, the UK, Russia, Turkey, Italy, Spain, and Germany are ranked among the top ten countries where COVID-19 cases recorded more than any other country.[4] However, the actual COVID-19 infections, recoveries, and deaths computation is still underestimated because of ununiform policies and limited testing facilities in urban and rural areas. Vaccinations also face challenges in making efficient immunization against the virus; however, it increases day by day and takes some more time to come on the floor. The fight with one of the most virulent enemies is still on, and the clinicians, scientists, NGOs, and common people are standing at the forefront. To strengthen and boost the drug discovery pipeline scientific community is continuously working towards identifying new oral drug molecules that can intervene viral replication process. The 3C-like protease ($3CL^{PRO}$), aka main protease (M^{PRO}) or nonstructural protein 5 (NSP5), is an important therapeutic target of the SARS-CoV-2 genome, which is encoded by its polyprotein 1a and 1ab. The viral polyproteins 1a and 1ab consist of 14 NSPs, papain-like cysteine protease (PL^{PRO}), and MPRO. The $3CL^{PRO}$ cleaves the polypeptide at various evolutionary conserved regions, releasing various nonstructural proteins (NSPs) to perform their essential functions and thus, the $3CL^{PRO}$ takes the crucial biocatalytic commands in the viral replication processes.[5-8] The $3CL^{PRO}$ itself is 306 amino acids long polypeptide having three domains, in which first domain covers 8 to 101 amino acids (AAs), the second domain contains 102-184 AAs, and the third domain includes 201-303 AAs, majorly folds into five α-helices that are distributed over a large antiparallel globular fold that joins the second domain through 185-200 residues long loop.[9] Cysteine residue positioned at 145 of the second domain and histidine residue positioned at 41 of the first domain form biocatalytic dyad, which serves as a substrate-binding crevice. During the activation process of the binding crevice, transfer of a proton from cysteine 145 residue

to histidine 41 takes place via acylation and diacylation process of an intermediate step. Due to its paramount importance in viral replication and transcription processes, $3CL^{PRO}$ is a critical therapeutic target against etiological agents, especially SARS-CoV-2.[10-12] In the proposed work, we aim to decipher new small-molecule inhibitor(s) akin to Pfizer's new drug molecule viz., nirmatrelvir, that blocks substrate binding crevice cysteine145, and Histidine 41 residues of domain second and first, respectively. To accomplish the aim, structure-based virtual screening (SBVS) with fundamental search limits, e.g., molecular weight (MW ≤ 500 Da), hydrogen bond donor (HBD ≤ 5), hydrogen bond acceptor (HBA ≤ 10), and lipophilicity (logP ≤ 5) was used to screen potential lead molecule from MCULE's huge digital investigational ligand repository. The MCULE database (https://mcule.com/) is a digital online drug discovery resource portal that contains more than five million synthetic accessible and purchasable molecules towards performing cell-based assays, pre-clinical, and clinical trials. The SBVS ligand was docked with $3CL^{PRO}$ via AutoDock Vina (ADV) inbuilt with MCULE drug discovery platform followed by toxicity assessments. Virtually obtained ligands were passed through the Brain or IntestinaL EstimateD (BOILED)-Egg model to check the human intestinal abortion (HIA) and blood-brain barrier (BBB) permeation based on two physicochemical properties, including topological surface area (TPSA) and WLOGP. The BOILED-Egg filtered compounds were passed through druglikeness features other than Pfizer' Lipinski rule of five, followed by medicinal chemistry's pan assay interference structure (PAINS) and Brenk alert investigation. The stability of ligand hits succeeded through the parameters mentioned above were accomplished via the molecular dynamics (MD) simulation process. A comparative analysis between the reference drug nirmatrelvir and a lead molecule depicted through the abovementioned approaches was made to elucidate the inhibition potential of the identified investigational lead molecule.

MATERIALS AND METHODS

Receptor Retrieval and Optimization

3D crystal structure (2.16 Å) $3CL^{PRO}$ of SARS-CoV-2 (6LU7) was extracted from RCSB PDB.[11] Only apo-part of the protein was taken to build a 3D input file suitable to the selected docking tool, and remaining heteroatoms, ions, and molecules were removed. Optimization of target protein was done by assigning the CHARMm force field. The protonation state of the receptor molecule was maintained by using the PropKa utility at physiologic pH.[13-17]

Target-based Virtual Screening

The digital online drug discovery resource portal MCULE was exploited for target-based virtual screening of investigational ligands from its digital repository containing more than five million synthetically accessible and purchasable ligands. In the SBVS workflow, the basic search limit obeying Pfizer's rule of five (RO5: MW ≤ 500 Da; HBD ≤ 5; HBA ≤ 10; LogP ≤ 5) was taken to screen small molecules. The sample size, diversity range, and resemblance search threshold for the input query were 1000, 100, and 0.90, respectively. The FP2 fingerprint of open babel as a 2D search algorithm was applied to run the SBVS workflow. Others search parameters were kept as default.[18-19]

3D Structure Retrieval of Reference Drug Nirmatrelvir

The 2D structure of nirmatrelvir (CID: 155903259) in SDF (standard data format) was retrieved from the NCBI PubChem database.[20-22] Two dimensional to the three-dimensional conversion of inhibitor was accomplished via Accelrys discovery studio visualizer (DSV). The energy minimization of the selected inhibitor was carried out through the same protocol as done in the receptor molecule.

Molecular Docking using AutoDock Vina

The AutoDock Vina inbuilt to the MCULE online drug discovery portal (https://mcule.com/) was used for docking simulation between $3CL^{PRO}$ and SBVS-ligands. The 3D input protein file in PDB format was provided to the ADV interface of the MCULE portal. A grid size covering the selected binding pocket was drawn to cover the protein binding crevice. ADV parameters for binding modes per ligand and exhaustiveness were kept default. Free energy of binding (ΔG) was selected as the key criterion to identify the best pose of ligand hits that docked into the binding crevice of $3CL^{PRO}$.[23-26]

Assessment of Toxic Moieties

Presence of toxiphoric moieties, fragments, and substructures in virtually-screened ligand hits unsafe to the human and ecosystem was investigated through the toxicity checker facility inbuilt to the MCULE resource portal based on the rigorous and robust SMILES arbitrary target specification (SMARTS) algorithm.[18]

BOILED-Egg Filtration

Brain Or IntestinaL EstimateD (BOILED)-Egg a.k.a. Egan Egg model of SwissADME tool, was employed to investigate the HIA and BBB permeation of predicted ligands. The BOILED-Egg employs the two physicochemical features, i.e., WLOGP (reference value ≤ 5.88) and TPSA (reference value ≤ 131.6) for lipophilicity and apparent polarity, respectively and distinct pictorial elaboration of how far a molecule is from the ideal one for plausible absorption.[27-30]

Medicinal Chemistry Features Assessment

The medicinal chemistry features include the promiscuous compound, a.k.a. frequent hitters were identified through pan assay interference structure (PAINS) alert option of Swiss ADME tools adopted from the Eli Lilly. Evaluation of undesired substructures, dyes, toxic moieties was carried out through Ruth Brenk alert.[31-32]

Stability evaluation through molecular dynamics simulation

Docked complex of the best ligand-$3CL^{PRO}$ and reference drug nirmatrelvir-CL^{PRO} was computationally simulated at 300K at the molecular mechanics level using GROMACS 5.1.2. to assess their stabilities. The ligands were separated from their respective docked complexes via GMX grep module. The CHARMm general force field (CGENFF) server predicted the topology and forcefield parameter files of the selected ligand hits. The topologies were generated for $3CL^{PRO}$ using pdb2gmx modules of GROMACS 5.1.2. The .str file of ligand known inhibitor were downloaded from CGENFF server.[33-35]

All ligand and reference drug bound complexes were drenched in a dodecahedron box of water molecules with a margin of 10 Å. The gmx editconf module was used for creating boundary conditions. By adding sodium and chloride ions using the gmx genion module, the charges on the bound complexes were neutralized to maintain neutrality, preserving the physiological concentration of 0.15 M. The system was then minimized for 2,50,000 steps using the steepest descent algorithm. The system temperature was raised from 0-300 K during their equilibration of 10 ns duration at constant NVT and NPT. After the equilibration phase, the particle mesh was assigned through Ewald method.[36-37] The gmx rms, gmx rmsf, and gmx sasa, gmx ΔGsolv, gmx Rg, and gmx HB modules of GROMACS were used to depict the root mean square deviation (RMSD), root mean square fluctuation (RMSF), solvent accessible surface area (SASA), free energy of solvation (ΔGsolv), and radius of gyration (Rg) and hydrogen bond plots.[33,38]

RESULTS

Structure-based Virtual Screening

We used the SBVS method to screen small molecules from the MCULE's digital repository of investigational ligands. 2,235,82 ligand hits substantially docked into the binding pocket of $3CL^{PRO}$ and satisfied the applied Pfizer's Lipinski RO5 as discussed earlier in methodology portrayed as results of the SBVS. As illustrated in Figure 1, the SBVS-hits were subsequently subjected to filter based on binding free energy (ΔG), 22 ligands were obtained that showed lower or qual ΔG than control drug nirmatrelvir (ΔG: -7.9 kcal/mol).

Docking Simulation and Toxicity Evaluation of Ligand Hits

All virtually screened ligands and reference inhibitor nirmatrelvir were docked into the binding crevice of $3CL^{PRO}$ using MCULE's ADV tool to assess their binding affinities in terms of binding free energy (ΔG) that was in the range of -8.8 kcal/mol to -7.9 kcal/mol. Docked complexes of predicted ligands and $3CL^{PRO}$ were compared with the reference drug nirmatrelvir in terms of ΔG, binding crevice and type of interactions. The nirmatrelvir was docked into the binding pocket of $3CL^{PRO}$ with a ΔG value of -7.9 kcal/mol exhibiting interactions with 17 residues, including the two key residues Cys-145 and His-41, via eight different binding interactions viz., van der Waals, conventional hydrogen bond, carbon-hydrogen bond, Pi-anion, Pi-sigma, Pi-Pi T-shaped, Alkyl, and Pi-alkyl bonds (Figure 2). After toxicity evaluation, only 10 ligands turn out as efficacious drug candidates, and the remaining 11 ligands were rejected. Lipinski RO5 parameters and binding affinity of these 10 ligands with $3CL^{PRO}$ is shown in Table 1.

The BOILED-Egg Filtration

As the name indicates, the model has yellow and white regions predicting the physicochemical positions for substantial BBB penetration and HIA permeation. Among 10 molecules, 1 molecule (MCULE-1987480634), exhibited brain penetration, 8 molecules (MCULE-3437733956, MCULE-7962834856, MCULE-3133395989, MCULE-8456493019,

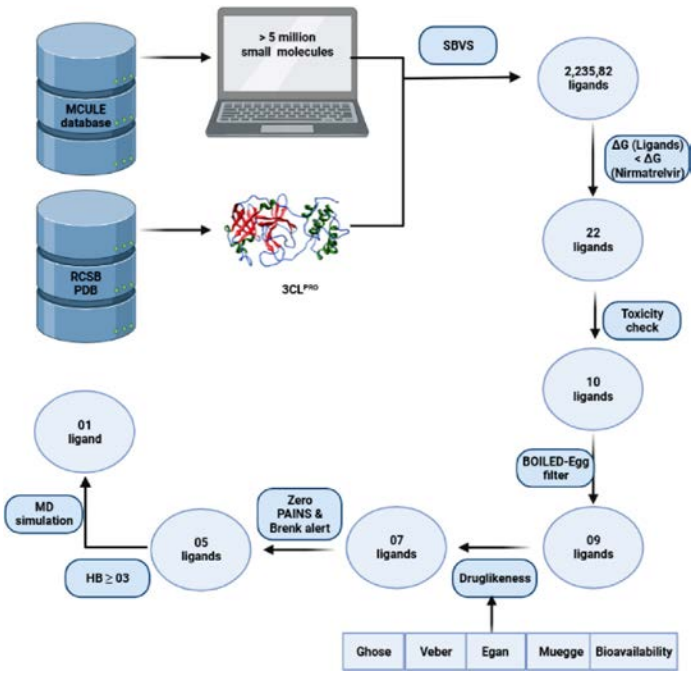

Figure 1: Flowchart of SBVS approach applied for identification of most probable lead molecule against $3CL^{PRO}$ of SARS-CoV-2.

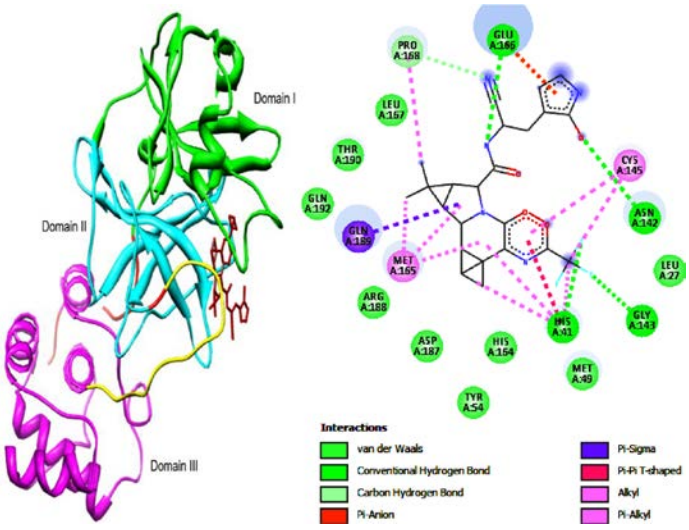

Figure 2: $3CL^{PRO}$-nirmatrelvir complex. Left Figure shows 3D pose of the nirmatrelvir (red sticks) docked to binding crevice of $3CL^{PRO}$. Right Figure shows the 2D pose of nirmatrelvir binding with different residues of $3CL^{PRO}$.

MCULE-8923803255, MCULE-7636440423, MCULE-1179948999, and MCULE-6773799844) showed plausible HIA permeation, and 1 molecule (MCULE-2009284974) showed neither BBB penetration nor HIA permeation. The reference drug nirmatrelvir showed substantial HIA permeation but not BBB. The BOILED-Egg prediction of ligands and reference molecules is shown in Figure 3. The blue and red dots represent P-gp positive and P-gp negative molecules, which means any ligands that are a substrate of P-glycoprotein has kicked out during BBB penetration, and non-substrate ligands can cross the brain membrane.

Druglikeness other than Pfizer's RO5

Ghose (Amgen), Veber (GSK), Egan (Pharmacia), Muegge (Bayer), and Abbott bioavailability score (BS) was used as druglikeness parameters that qualitatively compute the druglikeness for a small molecule to become promising oral lead molecules. Only 2 ligands, namely- MCULE-8456493019 and MCULE-763644042, were found as disobedient molecules of Egan and Ghose rules and the remaining 7 ligands obeyed all the parameters mentioned above. Nirmatrelvir exhibits 1 Ghose and 1 Veber violations. Although all 10 ligands and nirmatrelvir exhibited the Abbott BS value of 0.55, meaning they have substantial oral absorption characteristics (Table 2).

PAINS and Ruth Brenk alert Prediction

MCULE-1987480634 was predicted as a promiscuous compound means it exhibits PAINS alert. MCULE-1179948999 was recognized as a toxiphoric molecule means it shows Brenk alert (Table 3). These 2 molecules were excluded, while 5 molecules that show zero PAINS and Brenk alerts were taken forward for further *in-silico* study.

Hydrogen Bond Selection Criterion

The reference drug nirmatrelvir showed 5 conventional hydrogen bonds during binding interaction with the target protein $3CL^{PRO}$, while PAINS and Brenk-succeeded compounds showed a maximum of 3 hydrogen bonds upon interaction with $3CL^{PRO}$. On the contrary, ligands showed a strong binding affinity with more negative ΔG values and target protein residues. So, arbitrarily, only those compounds were taken forward that reflected maximum hydrogen bonds and residues of binding crevice viz., Cys-145, and His-41 of both domains I and II. Upon applying this

Table 1: Predicted Lipinski RO5 and binding affinity of ligands and reference drug molecule.

Compound	Formula	MW	HBA	HBD	MLOGP	ΔG (kcal/mol)
MCULE-1987480634	$C_{21}H_{17}N_3O$	327.38	3	2	4.28	-8.7
MCULE-3437733956	$C_{19}H_{17}N_7$	343.39	4	2	2.63	-8.5
MCULE-7962834856	$C_{19}H_{21}F_3N_4O_2$	394.39	7	2	2.63	-8.2
MCULE-3133395989	$C_{23}H_{18}F_2N_4O_4$	452.41	7	2	2.22	-8.8
MCULE-8456493019	$C_{22}H_{24}N_4O_5S$	456.51	7	2	1.10	-8.4
MCULE-2009284974	$C_{17}H_{13}NO_8S_2$	423.42	9	4	1.46	-8.1
MCULE-8923803255	$C_{23}H_{18}ClN_3O_2S$	435.93	3	2	4.32	-8.3
MCULE-7636440423	$C_{28}H_{23}N_5O_2$	461.51	4	2	3.25	-8.5
MCULE-1179948999	$C_{21}H_{22}N_4O_4S$	426.49	5	2	2.00	-8.5
MCULE-6773799844	$C_{20}H_{18}N_4O_3S$	394.45	5	2	2.30	-7.9
Nirmatrelvir-reference drug	$C_{23}H_{32}F_3N_5O_4$	499.53	8	3	0.41	-7.9

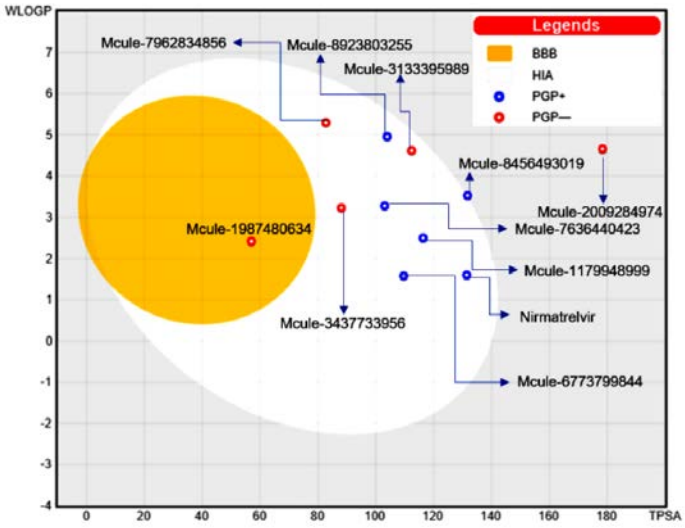

Figure 3: The BOILED-Egg a.k.a. Egan-Egg model evaluates passive HIA permeation and BBB penetration of ligands and nirmatrelvir. Out of 10 ligands, 1 ligand positioned within Egg yolk shows BBB penetration, 8 ligands and reference drug nirmatrelvir show substantial HIA permeation, while 1 molecule is positioned outside the egg, which means that it is neither brain penetrator nor gastrointestinal absorber. Dots are shown in the blue and red respectively exhibited substrate and non-substrate of P-glycoprotein.

Table 2: Predicted druglikeness other than Lipinski RO5 of ligands and reference inhibitor.

Molecule	Ghose	Veber	Egan	Muegge	Abbot BS
	Violation(s)				
MCULE-1987480634	0	0	0	0	0.55
MCULE-3437733956	0	0	0	0	0.55
MCULE-7962834856	0	0	0	0	0.55
MCULE-3133395989	0	0	0	0	0.55
MCULE-8456493019	0	0	1	0	0.55
MCULE-8923803255	0	0	0	0	0.55
MCULE-7636440423	1	0	0	0	0.55
MCULE-1179948999	0	0	0	0	0.55
MCULE-6773799844	0	0	0	0	0.55
Nirmatrelvir-reference drug	1	1	0	0	0.55

Table 3: Predicted medicinal chemistry features of drug likeness-succeeded ligands and nirmatrelvir.

Molecule	PAINS	Brenk
	Alert(s)	
MCULE-1987480634	1	0
MCULE-3437733956	0	0
MCULE-7962834856	0	0
MCULE-3133395989	0	0
MCULE-8923803255	0	0
MCULE-1179948999	0	1
MCULE-6773799844	0	0
Nirmatrelvir-reference drug	1	1

rule only, a single compound named MCULE-3133395989 was passed, showing 3 HBs and similar residues, including Cys-145 and His-41, were found engaged in molecular interactions with $3CL^{PRO}$ having a ΔG value of -8.8 kcal/mol. Moreover, MCULE-3133395989 interacted with 15 residues, including the two crucial residues Cys-145 and His-41 via four different binding interactions viz., van der Waals, conventional hydrogen bond, Pi-anion, and Pi-alkyl bonds (Figure 4).

Stability Assessment through MD Simulation

The stability of docked complexes of the best ligand hit MCULE-3133395989, and refrence drug nirmatrelvir with $3CL^{PRO}$ wea assesses through MD simulations of 10 ns duration using GROMACS package. MD graph for RMSD, RMSF, SASA, ΔGsolv, Rg, and HBs were ploted to evaluate the molecular interaction stability of ligands and protein docked complexes.[14,39-40]

Root-mean-square Deviation

The RMSD elucidate the stability of docked complexes. The mean RMSD for reference inhibitor nirmatrelvir (black) and predicted ligand hits MCULE-3133395989 (red) complexed with $3CL^{PRO}$ was found 0.168 nm and 0.184 nm, respectively. The RMSD plot reveals that the stability of docked complex of $3CL^{PRO}$ and MCULE-3133395989 is comparable to the docked comlex of reference molecule nirmatrelvir and $3CL^{PRO}$ (Figure 5).

Root-mean-square Fluctuation

The residues fluctuations at various regions of the RMSF graph are due to the binding interactions of nirmatrelvir and MCULE-3133395989 with

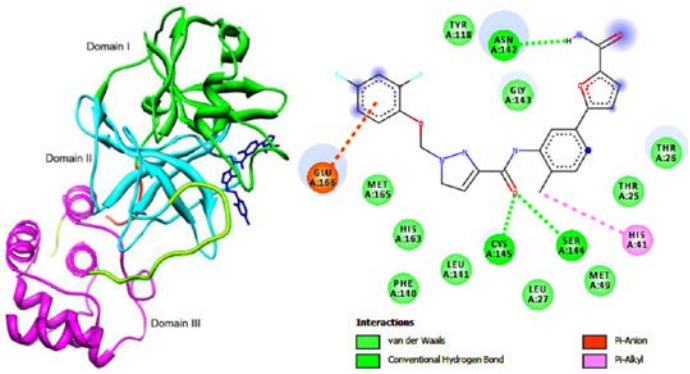

Figure 4: 3CL^{PRO}-MCULE-3133395989 complex. Left Figure shows 3D pose of the MCULE-3133395989 complex (blue sticks) docked to binding crevice of 3CL^{PRO}. Right Figure shows the 2D pose of MCULE-3133395989 binding with different residues of 3CL^{PRO}.

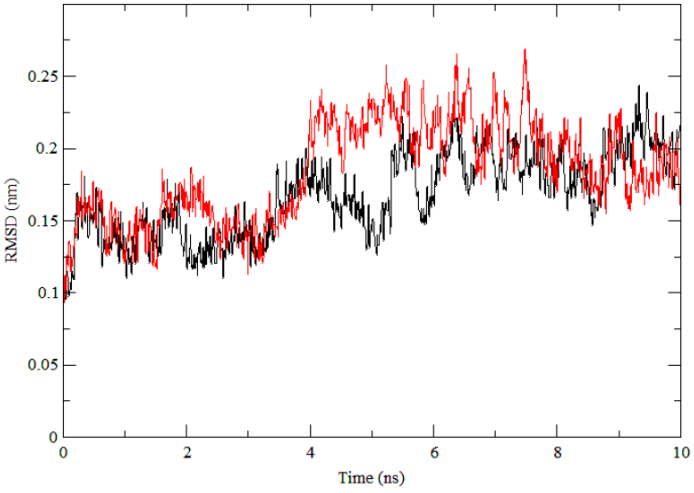

Figure 5: RMSD plot as a function of time. Black and red represent values computed for 3CL^{PRO}-nirmatrelvir and 3CL^{PRO}- MCULE-3133395989, respectively.

3CL^{PRO}. Average residues fluctuation upon binding with nirmatrelvir (black) and MCULE-3133395989 (red) was found as 0.098 nm and 0.116 nm, respectively (Figure 6).

Solvent-accessible Surface Area

The protein's surface area that is accessed by the solvent molecule is represented by the SASA plot. The average SASA values upon binding with nirmatrelvir (black), and MCULE-3133395989 (red) with 3CL^{PRO} was found as 18.499 nm^2 and 18.507 nm^2, respectively (Figure 7).

Free Energy of Solvation

The average free energy of solvation values upon binding with nirmatrelvir (black), and MCULE-3133395989 (red) with 3CL^{PRO} was found as 24.938 kJ/mol and -24.869 kJ/mol, respectively (Figure 8).

Radius of Gyration

The Rg reveals the compactness of docked complexes and it is inversely related to the compactness. The average Rg values of docked complexes of nirmatrelvir (black), and MCULE-3133395989 (red) with 3CL^{PRO} was found as 4.3397 nm and -2.2589 nm, respectively (Figure 9).

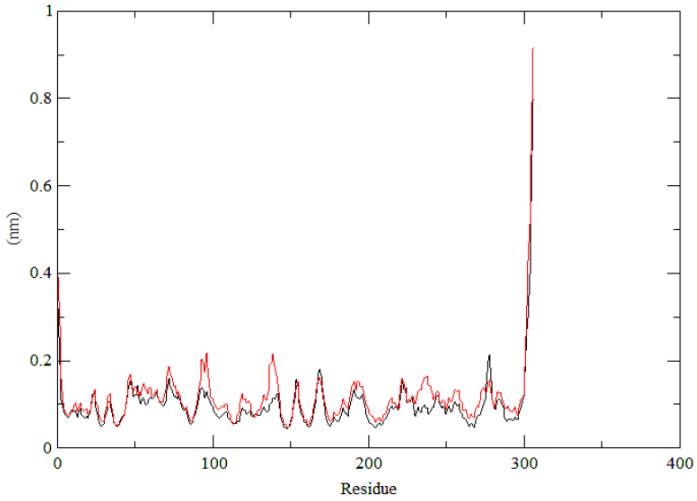

Figure 6: RMSF plot for 3CL^{PRO}-nirmatrelvir (black) and 3CL^{PRO}-MCULE-3133 395989 (Red).

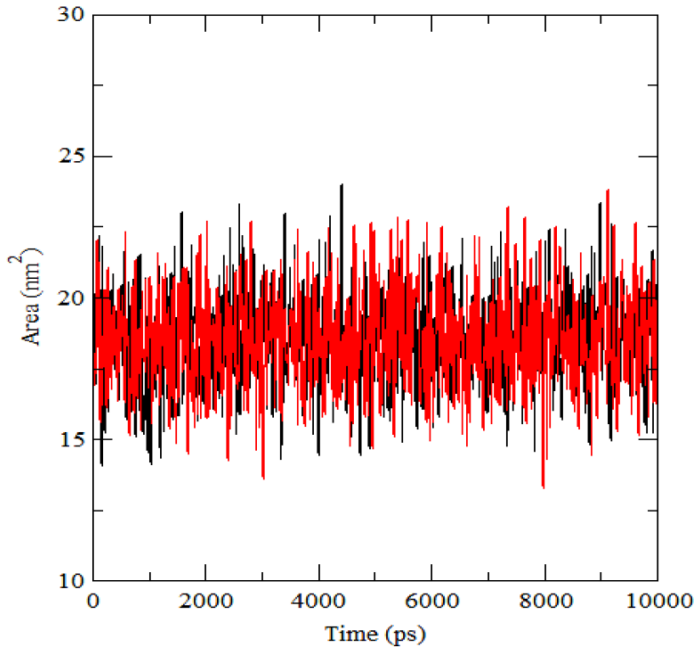

Figure 7: SASA plot for 3CL^{PRO}-nirmatrelvir (black) and 3CL^{PRO}-MCULE -3133 395989 (Red).

Hydrogen Bond Formation and Deformation

The HB plot shows the number of hydrogen bond formations, deformation and stability during the entire process of MD simulations. Figure 9a shows the HB plot for the docked complex of reference drug nirmatrelvir and protein 3CL^{PRO}, while Figure 9b depicts the nature of HB formation and deformation during the entire process of molecular dynamics simulation of 10 ns duration. Five hydrogen bonds are formed in the case of nirmatrelvir and target protein, but they did not remain stable till the entire duration of MD simulation, while in the case of MCULE-3133395989, 8 HBs are being formed, but they are not stable till the end of MD simulation process.

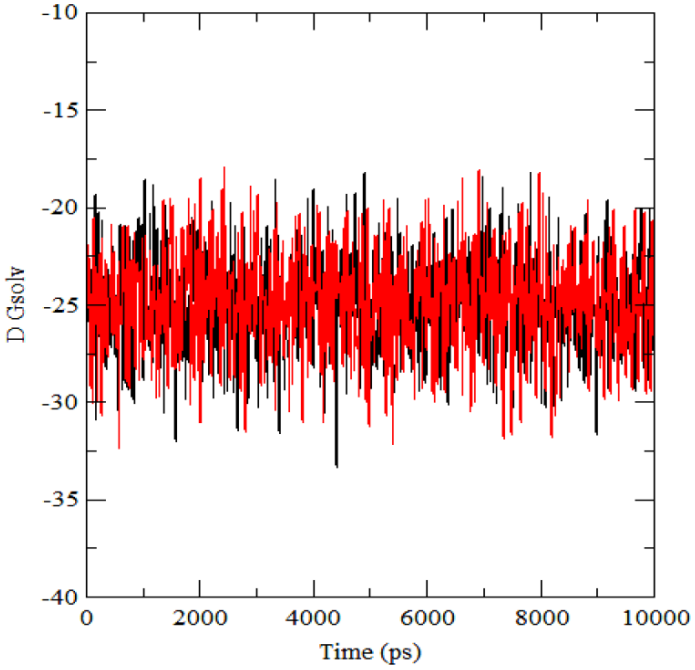

Figure 8: ΔGsolv plot for 3CL^{PRO}-nirmatrelvir (black) and 3CL^{PRO}-MCULE-3133395989 (Red).

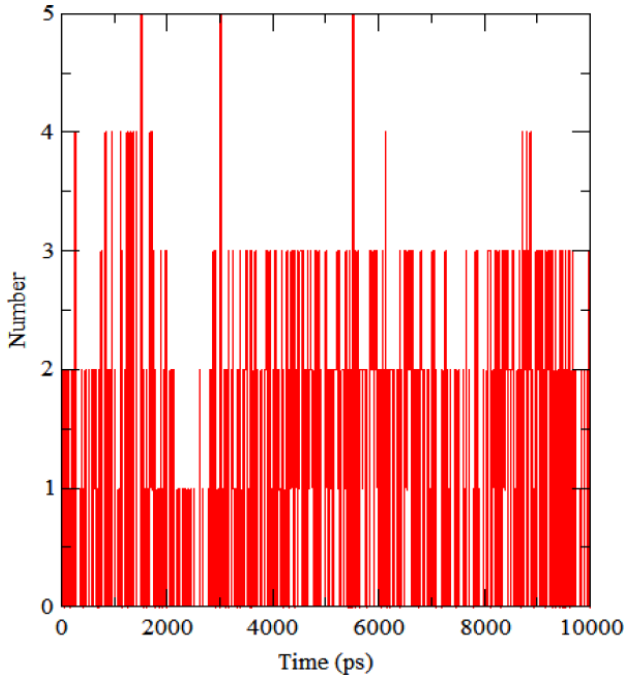

Figure 9a: HB plot shows the formation and deformation of H-bonds during interaction of nirmatrelvir with 3CL^{PRO}.

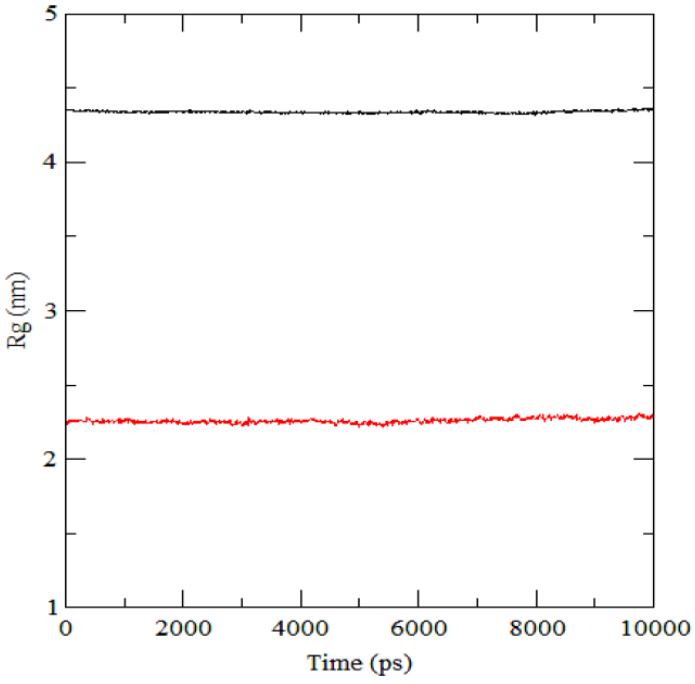

Figure 9: Rg plot for 3CL^{PRO}-nirmatrelvir (black) and 3CL^{PRO}-MCULE-3133395989 (Red).

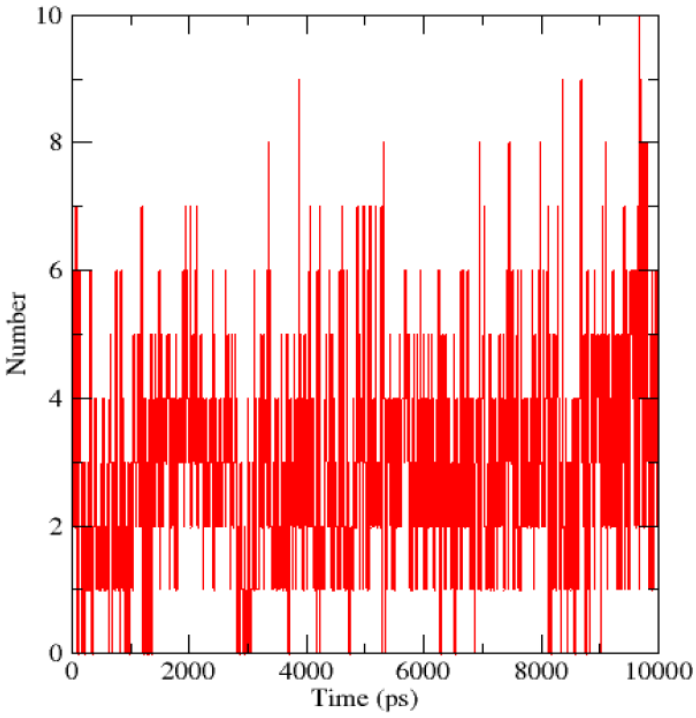

Figure 9b: HB plot shows the formation and deformation of H-bonds during interaction of MCULE-3133395989 with 3CL^{PRO}.

DISCUSSION

Despite the significant advancements in science and technology in recent decades and extensive research studies related to viral diseases, an ultimate strategy and cure design for their treatment still seem to be a resolved issue. Multiple interdisciplinary approaches have been used in antimicrobial research and therapy, from biochemical studies to biotechnological advancements such as high throughput proteogenomics approaches. Amid this scenario, computational approaches and bioinformatics have also come a step ahead in countering the backlash and limitations suffered by existing technologies. From being cost-effective to storing and processing billions of data in parallel,

computational strategies have bestowed easy access and excellent virtual platforms for modelling, designing, and analyzing highly sophisticated research data.

The current study focuses on the novel routes and the latest cascaded computational studies and advancements involving SBVS research and analysis to identify new investigational inhibitor(s) similar to Pfizer's newly lead molecule nirmatrelvir, that inhibits substrate-binding cavity made up of Cysteine145 and Histidine 41 residues of second and first domain, respectively. SBVS yielded 2,235,82 ligand hits from more than five million ligands of MCULE that were docked into the binding cavity of $3CL^{PRO}$ and obeyed the Lipinski RO5 as mentioned in the methodology section followed by sifting through ΔG, that forecasted 22 ligand hits having equal or lower ΔG as compared to reference drug nirmatrelvir (ΔG: -7.9 kcal/mol). Twenty-three ligand hits (out of 2,235,82 ligands) exhibited equal or lower ΔG (-8.8 kcal/mol to -7.9 kcal/mol) compared to nirmatrelvir narrowed down to 10 ligands after toxicity evaluations, and the remaining 11 ligand hits were excluded.

Among ten hits, one single-molecule showed BBB permeation, eight molecules exhibited substantial HIA permeation, and one molecule portrayed neither brain penetration nor GI permeation, while control drug nirmatrelvir depicted plausible HIA permeation but negligible brain permeability. Seven out of ten molecules followed all drug-likeness filtration viz., Ghose (Amgen), Veber (GSK), Egan (Pharmacia), Muegge (Bayer) and these also showed permissible values of Abbott BS. Two out of seven molecules were excluded due to the showing PAINS alert and Brenk alert, while the remaining five ligand hits adhered to the zero PAINS and Brenk alerts.

The known inhibitor nirmatrelvir interacted with the target protein 3CLPRO with five conventional HBs but predicted lead molecules depicted 3 HBs upon binding with 3CLPRO. The predicted ligand showed maximum HBs and lowered ΔG (-8.8 kcal/mol) value and interacted with Cys-145 and His-41 of domains I and II. This single-molecule with Id MCULE-3133395989 exhibited excellent binding into the active site (Cys-145 and His-41) and 13 other vicinal residues via four different binding forces, including van der Waals, conventional hydrogen bond, Pi-anion, and the Pi-alkyl bonds. MD simulation studies based on the RMSD, RMSF, SASA, ΔGsolv, Rg, and HBs parameters favour the MCULE-3133395989 as the most promising inhibitor of $3CL^{PRO}$ compared to control drug nirmatrelvir.

CONCLUSION

High-throughput target-based virtual screening, toxicity profiling, physicochemical properties, lipophilicity, solubility, pharmacokinetics, druglikeness, medicinal chemistry attributes, molecular docking, RMSD, RMSF, SASA, ΔGsolv, Rg and HBs analyses ascertain that MCULE-3133395989 contains all therapeutic features that are comparable with the Pfizer's drug nirmatrelvir and exhibited as an excellent oral drug candidate and might be promising against etiological agents viz., SARS-CoV-2, MERS-CoV, and other human CoVs via inhibiting main viral protease. Although we have performed thorough *in-silico* experimentations and analyses of ligands binding to the $3CL^{PRO}$, computational biology and its allied sciences have several possible research boundaries. Therefore, more rigorous *in-vitro* and *in-vivo* experiments are needed to validate the *in-silico* results of the study.

ACKNOWLEDGEMENT

King Abdulaziz City supports this study for Science and Technology, Riyadh, Saudi Arabia, under Grant No. 1-18-01-009-0035. Additionally, the authors thank the Deanship of Scientific Research (DSR) at King Abdulaziz University.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest.

REFERENCES

1. Tiwari U, Bano A, Ahmad Khan MK. A review on the COVID-19: Facts and current situation. NeuroPharmac J. 2021;30;06:180-91. doi: 10.37881/1.616.
2. Khan MKA, Pokharkar NB, Al-Khodairy FM, Al-Marshad FM, Arif JM. Structural perspective on molecular interaction of IgG and IgA with spike and envelope proteins of SARS-CoV-2 and its implications to non-specific immunity. Biointerface Res Appl Chem. 2020 Nov 15;11(3):10923-39. doi: 10.33263/BRIAC113.1092310939.
3. Kin N, Miszczak F, Lin W, Gouilh MA, Vabret A, EPICOREM Consortium. Genomic analysis of 15 human coronaviruses OC43 (HCoV-OC43s) circulating in France from 2001 to 2013 reveals a high intra-specific diversity with new recombinant genotypes. Viruses. 2015;7(5):2358-77. doi: 10.3390/v7052358, PMID 26008694.
4. https://www.worldometers.info/coronavirus/accessed on January 27, 2022; 19:55 IST.
5. Muramatsu T, Takemoto C, Kim YT, Wang H, Nishii W, Terada T, *et al*. SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. Proc Natl Acad Sci U S A. 2016;113(46):12997-3002. doi: 10.1073/pnas.1601327113, PMID 27799534.
6. Liu X, Wang XJ. Potential inhibitors against 2019-nCoV coronavirus M protease from clinically approved medicines. J Genet Genomics. 2020;47(2):119-21. doi: 10.1016/j.jgg.2020.02.001, PMID 32173287.
7. Ionescu MI. An overview of the crystallized structures of the SARS-CoV-2. Protein J. 2020;39(6):600-18. doi: 10.1007/s10930-020-09933-w, PMID 33098476.
8. Iketani S, Forouhar F, Liu H, Hong SJ, Lin FY, Nair MS, *et al*. Lead compounds for the development of SARS-CoV-2 3CL protease inhibitors. Nat Commun. 2021;12(1):2016. doi: 10.1038/s41467-021-22362-2, PMID 33795671.
9. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, *et al*. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020;582(7811):289-93. doi: 10.1038/s41586-020-2223-y, PMID 32272481.
10. Swiderek K, Moliner V. Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M pro by QM/MM computational methods. Chem Sci. 2020;11(39):10626-30. doi: 10.1039/d0sc02823a, PMID 34094317.
11. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, *et al*. The crystal structure of COVID-19 main protease in complex with an inhibitor N3. Nat 2020. 2020;582(7811):289–93: 5827811. PMID 32272481.
12. Mishra A, Waghela R. A comparative study of approved drugs for SARS-CoV-2 by molecular docking. J mol docking. 2021;1(1):25-31.
13. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, *et al*. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem. 1998;19(14):1639-62. doi: 10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.
14. Khan FI, Lai D, Anwer R, Azim I, Khan MKA. Identifying novel sphingosine kinase 1 inhibitors as therapeutics against breast cancer. J Enzyme Inhib Med Chem. 2020; Jan 1;35(1):172-86. doi: 10.1080/14756366.2019.1692828, PMID 31752564.
15. Ahmad KMK, Salman A, Al-Khodairy Salman F, Al-Marshad Feras M, Alshahrani Abdulrahman M, *et al*. Computational Exploration of dibenzo [a,l] pyrene Interaction to DNA and its Bases: Possible Implications to Human Health. Biointerface Res Appl Chem. 2020;11(4):11272-83. doi: 10.33263/BRIAC114.1127211283.
16. Bas DC, Rogers DM, Jensen JH. Very fast prediction and rationalization of pKa values for protein-ligand complexes. Proteins. 2008;73(3):765-83. doi: 10.1002/prot.22102, PMID 18498103.
17. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem. 1983;4(2):187-217. doi: 10.1002/jcc.540040211.
18. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. J Cheminform. 2012;4(1):17. doi: 10.1186/1758-2946-4-17. PMID 22889332.
19. Shakil S. Molecular interaction of investigational ligands with human brain acetylcholinesterase. J Cell Biochem. 2019; Jul 11;120(7):11820-30. doi: 10.1002/jcb.28461, PMID 30746750.
20. Macchiagodena M, Pagliai M, Procacci P. Characterization of the non-covalent interaction between the PF-07321332 inhibitor and the SARS-CoV-2 main protease. J Mol Graph Model. 2022; Jan;110:108042. doi: 10.1016/j.jmgm.2021.108042, PMID 34653812.
21. Wu CR, Yin WC, Jiang Y, Xu HE. Structure genomics of SARS-CoV-2 and its Omicron variant: Drug design templates for COVID-19. Acta Pharmacol Sin. 2022; Jan 20;1-3. doi: 10.1038/s41401-021-00851-w, PMID 35058587.
22. Heskin J, Pallett SJC, Mughal N, Davies GW, Moore LSP, Rayment M,

et al. Caution required with use of ritonavir-boosted PF-07321332 in COVID-19 management. Lancet. 2022;399(10319):21-2. doi: 10.1016/S0140-6736(21)02657-X, PMID 34973713.

23. Steffen C, Thomas K, Huniar U, Hellweg A, Rubner O, Schroer A. TmoleX–a graphical user interface for TURBOMOLE. J Comput Chem. 2010;31(16):2967-70. doi: 10.1002/jcc.21576. PMID 20928852.
24. Ajijur R, Salman A, Ahmad KMK. Combinatorial design to decipher novel lead molecule against Mycobacterium tuberculosis. Biointerface Res Appl Chem. 2021;11(5):12993-3004. doi: 10.33263/BRIAC115.1299313004.
25. Khan MKA, Akhtar S, Arif JM. Development of *in silico* protocols to predict structural insights into the metabolic activation pathways of xenobiotics. Interdiscip Sci. 2018;10(2):329-45. doi: 10.1007/s12539-017-0237-4, PMID 28527150.
26. Khan MKA, Akhtar S, Arif JM. Structural insight into the mechanism of dibenzo[a,l]pyrene and benzo[a]pyrene-mediated cell proliferation using molecular docking simulations. Interdiscip Sci Comput Life Sci. 2018;10(4):653-73. doi: 10.1007/s12539-017-0226-7, PMID 28374118.
27. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules [sci rep]. Sci Rep. 2017;7(1):42717. doi: 10.1038/srep42717, PMID 28256516.
28. Attique SA, Hassan M, Usman M, Atif RM, Mahboob S, Al-Ghanim KA, *et al*. A molecular docking approach to evaluate the pharmacological properties of natural and synthetic treatment candidates for use against hypertension. Int J Environ Res Public Health. 2019;16(6):923. doi: 10.3390/ijerph16060923, PMID 30875817.
29. Egan WJ, Merz KM, Baldwin JJ. Prediction of drug absorption using multivariate statistics. J Med Chem. 2000;43(21):3867-77. doi: 10.1021/jm000292e, PMID 11052792.
30. Egan WJ, Lauri G. Prediction of intestinal permeability. Adv Drug Deliv Rev. 2002;54(3):273-89. doi: 10.1016/s0169-409x(02)00004-2, PMID 11922948.
31. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem. 2010;53(7):2719-40. doi: 10.1021/jm901137j, PMID 20131845.
32. Brenk R, Schipani A, James D, Krasowski A, Gilbert IH, Frearson J, *et al*. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. ChemMedChem. 2008;3(3):435-44. doi: 10.1002/cmdc.200700139, PMID 18064617.
33. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. J Comput Chem. 2005;26(16):1701-18. doi: 10.1002/jcc.20291, PMID 16211538.
34. Vanommeslaeghe K, MacKerell AD. Automation of the CHARMM general force field (CGenFF) I: Bond perception and atom typing. J Chem Inf Model. 2012;52(12):3144-54. doi: 10.1021/ci300363c, PMID 23146088.
35. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, *et al*. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem. 2010;31(4):671-90. doi: 10.1002/jcc.21367, PMID 19575467.
36. Petersen HG. Accuracy and efficiency of the particle mesh Ewald method. J Chem Phys. 1995;103(9):3668-79. doi: 10.1063/1.470043.
37. Stenberg S, Stenqvist B. An exact Ewald summation method in theory and practice. J Phys Chem A. 2020;124(19):3943-6. doi: 10.1021/acs.jpca.0c01684, PMID 32285671.
38. Fischer NM, Van Maaren PJ, Ditz JC, Yildirim A, Van der Spoel D. Properties of organic liquids when simulated with long-range Lennard-Jones interactions. J Chem Theor Comput. 2015;11(7):2938-44. doi: 10.1021/acs.jctc.5b00190, PMID 26575731.
39. Ali S, Khan FI, Mohammad T, Lan D, Hassan MI, Wang Y. Identification and evaluation of inhibitors of lipase from *Malassezia restricta* using virtual high-throughput screening and Molecular Dynamics studies. Int J Mol Sci. 2019;20(4):884. doi: 10.3390/ijms20040884, PMID 30781686.
40. Kuzmanic A, Zagrovic B. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. Biophys J. 2010;98(5):861-71. doi: 10.1016/j.bpj.2009.11.011, PMID 20197040.

Article History: Received: 05-02-2022; Revised: 06-04-2022; Accepted: 06-05-2022.
Cite this article: Alwabli AS. Lead Identification against 3C-like Protease of SARS-CoV-2 Via Target-based Virtual Screening and Molecular Dynamics Simulation. J Young Pharm. 2022;14(2):179-86.