Pharmaceutical Sciences, 2021, 27(Suppl 1), S122-S134
doi:10.34172/PS.2021.4
https://ps.tbzmed.ac.ir/

Research Article

Qualitative and Quantitative Analysis of Anti-viral Compounds against SARS-CoV-2 Protease Enzyme by Molecular Dynamics Simulation and MM/PBSA Method

Saghi Sepehri[1,2*], Niloufar Hashemidanesh[1,3], Karim Mahnam[4], Hila Asham[1,3]

[1]*Department of Medicinal Chemistry, School of Pharmacy, Ardabil University of Medical Sciences, Ardabil, Iran.*
[2]*Pharmaceutical Sciences Research Center, Ardabil University of Medical Sciences, Ardabil, Iran.*
[3]*Students Research Committee, School of Pharmacy, Ardabil University of Medical Sciences, Ardabil, Iran.*
[4]*Department of Biology, Faculty of Sciences, Shahrekord University, Shahrekord, Iran.*

Article Info

Article History:
Received: 9 September 2020
Accepted: 13 January 2021
ePublished: 29 January 2021

Keywords:
-Anti-viral drug
-Corona virus
-COVID-19
-Drug repurposing
-Molecular dynamics

Abstract

Background: A significant worry for global public health is the international spread of the coronavirus disease-19 triggered through the new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Herein, an attempt was performed to qualitative and quantitative analysis of a series of compounds against SARS-CoV-2 main protease (M<[pro]) by *in silico* studies.

Methods: About one hundred anti-viral compounds were collected from DrugBank database. In the second stage, molecular docking simulation was carried out to identify interactions of the molecules with the key residues in the M<[pro] active site. Finally, the molecular dynamics simulation (MD) of four top-ranked compounds and X77 as co-crystal ligand were investigated.

Results: Based on molecular docking studies, four compounds DB00224, DB00220, DB01232 and DB08873 exhibited the best results among compounds against M<[pro] enzyme. Additionally, molecular dynamic simulation and free binding energy were accomplished to compute the interaction energies and stability of the top-ranked compounds at the active site. The binding energy portions of the compounds into the enzyme active site exposed that Van der Waals and non-polar interactions were fundamental factors in the molecule binding. The ligand connections were steadied via hydrophobic interactions and several key hydrogen bonds especially with Glu166 and His41 residues into the active site.

Conclusion: According to calculations of docking and MD, it was observed that the active site is mostly hydrophobic. Additionally, the results showed the steady of selected ligands binding with SARS-CoV-2 M<[pro] active site.

Introduction

One of the causes of severe infections in animals and humans is coronaviruses (CoVs), causing severe problems in the respiratory and digestive tracts.[1] CoVs belong to the *Coronaviridae* family and when detected through an electron microscope, appear just like spiked rings. This viral disease has a wide range of symptoms which can be appeared in the form of simple common cold illness to become severe diseases similar to Middle East Respiratory Syndrome (MERS-CoV) and Severe Acute Respiratory Syndrome (SARS-CoV).[2] The first CoV was found in the birds and later in humans in 1937 and 1960s, respectively.[2] In December 2019, a fast and extensive epidemic of a newly emerged human coronavirus 2019 designated COVID-19 which was first reported from the Chinese city of Wuhan, Hubei province to spread around the world having the potential to become a pandemic.[3] Since COVID-19 is very similar to SARS-CoV, it was classified as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) on February 11, 2020.[3]

COVID-19 is a highly infectious disease with high mortality rate and unfortunately there are not any approved drugs or vaccines for this disease yet. The number of COVID-19 patients is accumulating each day and more than 89,707,115 patients and 1,940,352 cases of death by COVID-19 has been confirmed at the time of this writing.[4] Numerous options such as monoclonal antibodies, oligonucleotide-based therapies, vaccines, interferon therapies, peptides and small-molecule drugs can be

*Corresponding Author: Saghi Sepehri, E-mail: sepehri.saghi@yahoo.com; s.sepehri@pharmacy.arums.ac.ir

suggested to control or treat emerging infections of COVID-19.[5] Many drugs, such as ribavirin, lopinavir and interferon have been tried against SARS or MERS, while the efficiencies of some drugs remained controversial.[6] One of the efficient treatment methods against the COVID-19 pandemic is taking remdesivir alone or in combination with chloroquine or interferon beta. For this treatment, any side effects have not been reported yet. Other recommended viral chemotherapy of human pathogenic CoVs contains neuraminidase inhibitors, nucleoside analogues, lamivudine, tenofovir disoproxil, and umifenovir (arbidol).[7]

The genome of CoVs is a single-stranded RNA viral, positive-sense (27 to 34 kilobases) and nucleocapsid of helical symmetry. Usually, the size of the CoVs is ~20 nm covered with a big petal or club-shaped surface look. Structural proteins of CoVs consist of Spike (S), Nucleocapsid (N), Matrix (M), and Envelope (E) and their non-structural proteins are papain-like protease (PLpro), RNA-dependent RNA polymerase (RdRp), main protease (M<[pro], also called 3CLpro) and helicase.[8] Therefore, these proteins or enzymes can be considered as important therapeutic targets against COVID-19 infection.

One of the most important drug targets of CoVs is the M<[pro]. This protein along with the PLpro has an essential role in viral replication, transcription and packaging within the host cells. Thus, one of the significant drug targets to inhibit coronavirus can be the M<[pro].[9]

Given the important role of SARS-CoV-2 M<[pro] in viral infection, recent research studies have focused on the identification of *in silico* hit compounds as potential SARS-CoV-2 inhibitors. For instance, Mirza and Froeyen reported some hits through a structure-based virtual screening approach.[10] Macchiagodena *et al.*[11] used virtual screening strategy and molecular dynamics technique to recognize the possible lead compounds as M<[pro] inhibitors of the SARS-CoV-2.[11] Chandel *et al.*[12] identified nineteen potential inhibitors through a drug repurposing approach and virtual screening method as strong inhibitors against SARS-CoV-2 M<[pro]. The study conducted by Xu *et al.*[7], in which a series of drugs against SARS-CoV-2 M<[pro] was investigated using MM/GBSA and solvated interaction energy (SIE) methods and introduced nelfinavir as the best potential inhibitor.

In this study, we investigated qualitative and quantitative analysis of series of anti-viral compounds via high-throughput docking and molecular dynamic simulation studies in order to recognize possible inhibitors to combat the recent dangerous SARS-CoV-2. To get more information, the flow chart of the multi-step simulations is depicted in Figure 1.

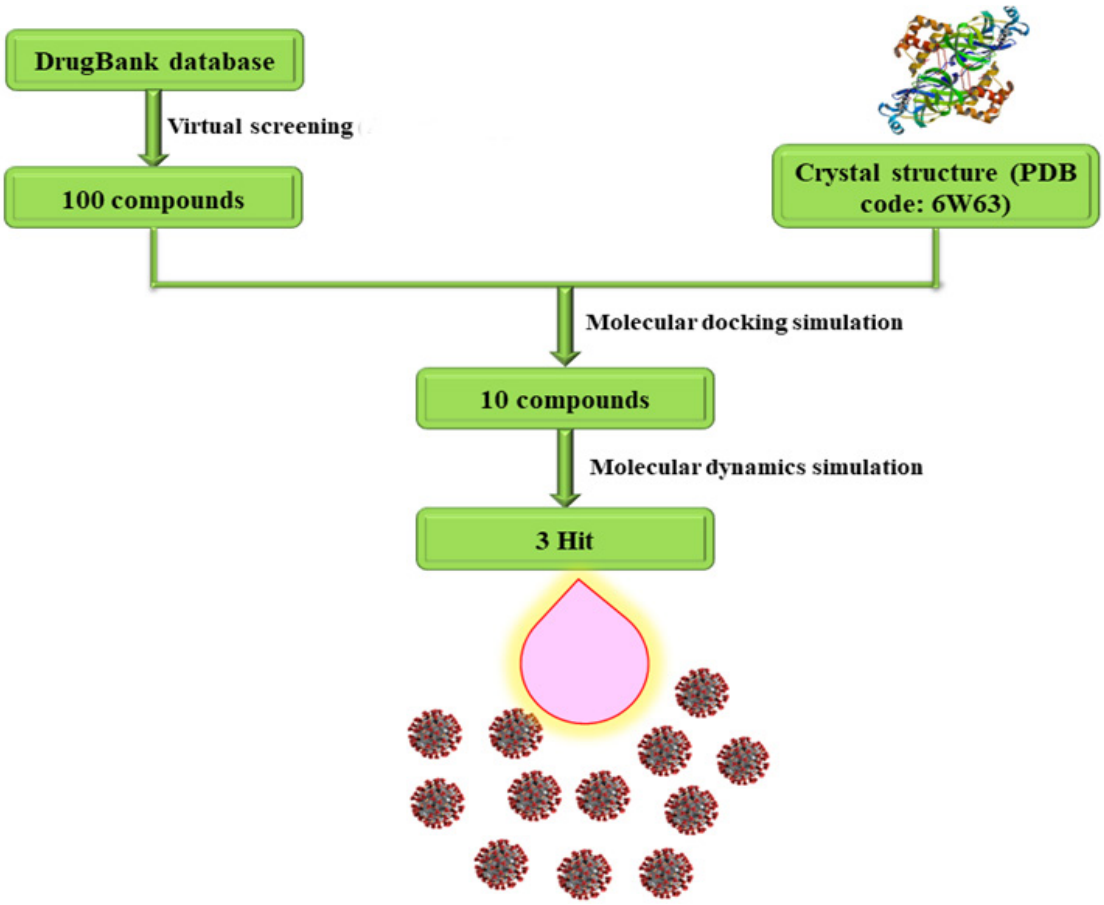

Figure 1. The schematic diagram of hit compounds discovery based on virtual screening and molecular dynamics.

Materials and Methods

SARS-CoV-2 M<[pro] protein preparation and construction of compounds dataset

The three-dimensional (3D) crystal structure of SARS-CoV-2 M<[pro] was regained from the Protein Data Bank (PDB ID: 6W63). Thus, the crystallized water molecules and Co-crystal ligand were eliminated and polar hydrogen and charges were added to the protein. To perform virtual screening a molecules library is required. So, a library was made by counting some anti-viral molecules from the DrugBank databases (https://www.drugbank.ca). Selected ligands were ready in the protein data bank (pdb) format. Then, the partial charges by Gasteiger-Marsili technique were added to molecules in AutoDock Tools 1.5.6 (ADT) package.[13] Lastly, molecules were converted to PDBQT format to perform docking calculations.

High-throughput docking studies

In this step, to forecast favorable binding modes and affinity of selected molecules from the previous step was carried out molecular docking simulation. Molecular docking studies into SARS-CoV-2 M<[pro] enzyme were performed by AutoDock 4.2 software. Genetic Algorithm (LGA) approach was chosen as the search algorithm for the global optimum binding position search and for each docking calculation, the number of runs was set to 150. A grid map of 70 × 70 × 70 points and a grid point spacing of 0.375 Å were used by the Autogrid. After docking calculations, the molecules were ranked based on score docking. Enzyme-molecule interactions were all visualized using ADT and discovery studio visualizer 4 softwares (Accelrys lnc, San Diego, CA, USA).

Molecular dynamics simulation studies

GROningen machine for chemical simulations V4.5.5 (GROMACS) software package was applied for the molecular dynamic simulation (MDS). The forcefield parameters and topology files were provided by PRODRG server for the filtered molecules. Water molecules and suitable forcefield were characterized by SPC216 model and GROMOS96 43a1, respectively. MD simulation was carried out based on the described procedure in the previous article.[14] A 50 ns MD simulation was accomplished by checking equilibration through root-mean-squared deviations (RMSDs) of the backbone atoms.

Binding free energy analyses

Free binding energy values of complexes were estimated by the molecular mechanic/poisson-boltzmann surface area (MM/PBSA) technique from MD trajectories. G_mmpbsa tool was applied to calculate the binding free energy of these compounds.[15] In this procedure, ΔG_{bind} is measured from the free energies of the ligand-protein system:

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

In the recent research study, the free binding energy of the four complexes and Co-crystal ligand (**X77**) were investigated throughout the last stable 20 ns period of MD simulation analysis.

Virtual screening procedure

The screening procedure was performed in three phases. First, docking calculations of compounds was accomplished into SARS-CoV-2 M<[pro] active site to study their binding modes and affinity with amino acids. Second, all molecules were ranked based on their free binding energies. Finally, the studies of molecular dynamics simulation of hit compounds were carried out in order to determine the stability of the ligand-enzyme complexes (Figure 1).

Results and discussion

SARS-CoV-2 M<[pro] protein preparation and construction of drugs dataset

DrugBank is a complete, freely accessible web server comprising FDA-approved drugs and investigational drugs going via the FDA approval procedure. The DrugBank was introduced in 2006 and is sustained to progress over the past 12 years. Now, there are about 2358 drugs, 4501 stages evolutional drug, 3620 drugs with experimentally obtained nuclear magnetic resonance and mass spectrometry spectra, 365984 the number of drug-drug interactions and 5993 the number of pharmacogenomic and SNP-associated drug effects in the DrugBank database.[16] In this study, a collection of about 100 approved anti-viral compounds were obtained from the DrugBank database. Then, the virtual screening of selected molecules was made using AutoDock 4.2 software. The molecules were ranked according to score docking in the enzyme active site. In this regard, ten molecules with the highest binding energies were chosen for the next step.

High-throughput docking studies

Before investigating the compounds, the validation of the docking procedure was evaluated using re-docking ligand **X77** at the active site of the SARS-CoV-2 M<[pro] enzyme. To validate the docking test, a ligand from X-ray structure of the protein is come out and re-docked into its active site. The validation of docking procedure is successful when the RMSD is lower than a determined value (usually<2.0 Å). The RMSD obtained for **X77** when it was docked into the SARS-CoV-2 M<[pro] was within this cutoff limit (0.13 Å). By checking up the conformation of top sorted compounds, hydrophobic and hydrogen bonding interactions were the important factors for connection.

The interactions of the ligand **X77** could be distinguished as follows:

-Cyclohexanamine moiety showed hydrophobic interactions with Gly143 and Glu166 residues.

-*Tert*-butylphenyl and pyridine moieties exhibited hydrophobic interactions with Leu141, His164, Asn142, Asp187, Ser144, Phe140, Arg188 and Glu189 residues.

-The oxygen atoms of two carbonyl groups formed two hydrogen bonds with Glu166 and Gly143 residues. In

addition, the nitrogen atom of the pyridine ring formed a hydrogen bond with His163 residue.

-Imidazole ring demonstrated a π-cation interaction with His41 amino acid. The superimposition of ligand **X77** resulting from the *in silico* calculation and X-ray crystallography into the binding pocket of SARS-CoV-2 M<[pro] enzyme has been shown in Figure 2.

Ten molecules with the highest estimated free binding energy (Tables 1 and 2) that passed the previous step were subjected to high screening molecular docking into the SARS-CoV-2 M<[pro] active site (Figure 3). The free binding energies of these compounds ranged from -8.50 to -11.03 kcal/mol.

Compounds were sorted using their score docking and were analyzed to find the best binding mode in the active site. Four molecules, *viz.* **DB01232**, **DB08873**, **DB00220** and **DB00224** showed the highest score docking toward other compounds and **X77** in SARS-CoV-2 M<[pro] (Tables 1 and 2).

Glu166 and/or His41 were found to form hydrophobic interactions and hydrogen bonds with all of the four molecules (Tables 1 and 2) and hence may be regarded as key residues in maintaining relevant complexes.

Binding mode of **DB00224** in SARS-CoV-2 M<[pro] active site showed a hydrogen bond between the hydroxyl group of ligand and NH of Glu166 (Figure 4). Also, **DB08873** formed two hydrogen bonds between urea moiety NHs of ligand and the carbonyl oxygen atom of this amino acid. Similarly, a hydrogen bond pattern could be detected for **DB01232** between amide NH of ligand and the carbonyl

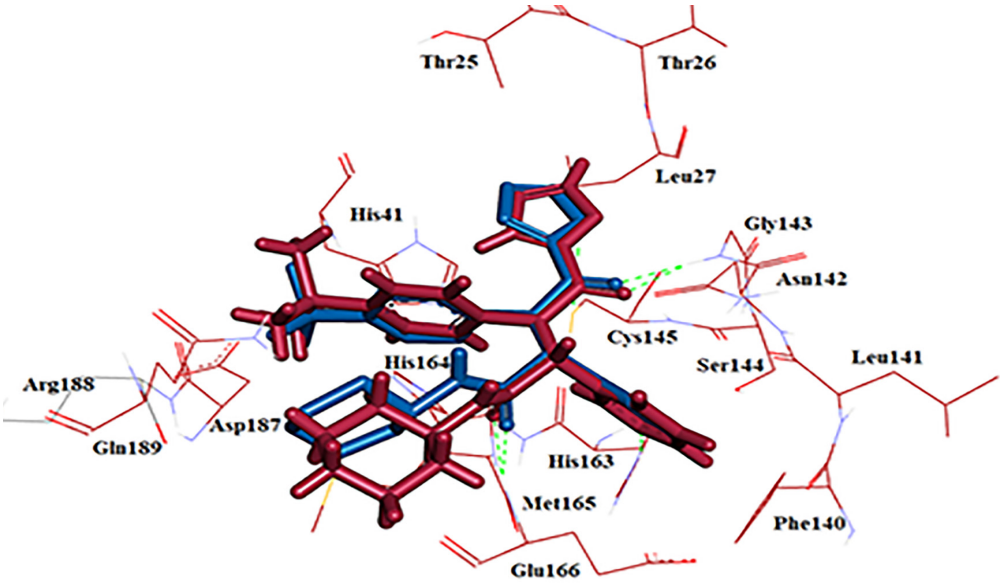

Figure 2. The superimposition of ligand X77 resulting from the docking simulation (Blue) and ligand X77 resulting from X-ray crystallography (Red) in the active site of SARS-CoV-2 M<[pro] enzyme (PDB ID: 6W63).

Table 1. Docking results of selected compounds docked into the SARS-CoV-2 M<[pro].

Compound	$\Delta G_{binding}$ (Kcal/mol)	VHDE[a] (Kcal/mol)	EE[b] (Kcal/mol)	IE[c] (Kcal/mol)
DB01601	-10.19	-12.60	+0.02	-12.58
DB01264	-9.87	-11.42	-0.24	-11.66
DB00220	-10.76	-13.10	-0.04	-13.14
DB00701	-9.86	-11.09	-1.46	-12.55
DB00224	-10.61	-12.65	-0.05	-12.70
DB08864	-9.15	-9.50	-0.25	-9.75
DB01232	-11.03	-13.99	-0.92	-14.91
DB08873	-10.49	-13.46	-0.01	-13.47
DB13997	-8.63	-10.07	-0.05	-10.13
DB12026	-8.50	-10.51	-0.08	-10.59
X77	-9.92	-11.83	-0.18	-12.00

[a] Van der Waals-H bond-Desolvation-Energy
[b] Electrostatic Energy
[c] Intermolecular Energy

Table 2. Hydrogen, hydrophobic, π-π stacking and cation-π interactions for the docked compounds with SARS-CoV-2 M<[pro].

Compound	Hydrogen bonds	Hydrophobic	π-π	Cation-π
DB01601	Glu166 with oxygen of C=O (1.82 Å), Glu166 with NH (2.20 Å)	Thr25, Tyr54, His41, Asp187, Gln189, Arg188, Thr190, Gln192, Pro168, His164, Met49, Met165, Asn142, Glu166, Cys44, Gly143, His163, Cys145, Phe140, Ser144, Leu141	-	-
DB01264	Glu192 with NH_2 (2.38 Å), Thr190 with NH_2 (1.88 Å), His164 with NH (2.29 Å), Gln189 with oxygen of sulfonyl (2.19 Å)	Leu141, His163, Ser144, Gly143, Asn142, Glu166, Cys145, Gln189, His164, Phe140, His41, Met165, Met49, Thr190, Ala191, Pro168, Gln192, Leu167, Arg188, Asp187, Tyr54	-	-
DB00220	Gln189 with NH (2.30 Å) His41 with OH (1.95 Å)	Gly143, Leu141, Ser144, Asn142, Cys145, Glu166, Met165, Pro168, Thr190, Arg188, Leu27, Cys44, His163, His41, His164, His172, Leu167, Asp187, Gln192, His172, Val42, Arg188, Thr25, Met49, Tyr54	Thr25 with Ph ring	His172 with Ph ring
DB00701	Gly143 with oxygen of C=O (2.35 Å), His164 with OH (2.04 Å), His163 with oxygen of tetrahydrofuran (1.94 Å), Thr190 with NH_2 (2.45 Å)	Arg188, Gln189, Thr190, Leu167, His164, Glu166, Phe140, Cys145, His163, Asn142, Ser144, Leu141, Gly143, Glu192, Asp187, Pro168, His41, Met165	-	-
DB00224	Glu166 with OH (2.26 & 2.15 Å), Gln192 with OH (2.12 Å), Thr190 with OH (1.88 Å)	Ala191, Pro168, Thr190, Gln192, Arg188, Gln189, Met49, Thr25, Gly143, Ser144, Glu166, Asn142, Cys145, His41, Cys44, Leu141, His163, Met165, His164	-	-
DB08864	Glu166 with NHs (2.12, 2.09 & 2.45 Å), Gly143 with NH (2.02 Å)	Cys44, Tyr54, Gly143, Ser144, Leu141, Cys145, Asn142, Glu166, Leu167, Met165, Pro168, Gln189, Asp187, Met49, Arg188, Gln192	-	-
DB01232	Thr190 with NH_2 (1.94 Å), Arg188 with NH_2 (1.81 Å), Glu166 with NH of amide (1.66 Å)	Gln192, Thr190, Arg188, Leu167, Gly170, Gln189, Pro168, Glu166, Asp187, Met 165, Tyr54, Met49, Asn142, Leu141, Cys145, His41, Gly143, His163, Ser144	His41 with benzyl ring	-
DB08873	His164 with NH_2 (2.15 Å), Glu166 with NHs (1.96 & 2.15 Å),	Arg188, Leu167, His41, Met165, Cys145, Leu141, His163, Phe140, Ser144, Gly143, Asn142, Met49, Gln189, Gln192, Pro168, Thr190	-	-
DB13997	Cys145 with oxygen of C=O (2.43 Å), His163 with oxygen of C=O (2.14 Å), His163 with oxygen of ester (2.07 Å), Gln166 with oxygen of C=O (2.09 Å)	Met49, Arg188, Gln189, His164, Leu27, Thr25, His41, Glu166, Met165, Cys44, Thr26, Cys145, His163, His172, Phe140, Asn142, Gly143, Leu141, Ser144	-	-
DB12026	Glu166 with NH (2.22 Å), Gly143 with oxygen of C=O (2.14 Å), Pro168 with NH (1.95 Å)	Ser144, Cys145, Asn142, His163, His164, Met165, Thr190, Leu167, Gln192, Arg188, Pro168, Gly143, Glu166, Gly170, Thr169, Asp187, His41, Thr25, Gln189, Pro52, Met49, Tyr54, Cys44	-	-
X77	Glu166 with oxygen of C=O (1.76 Å), Gly143 with oxygen of C=O (2.19 Å), Cys145 with NH (3.13 Å)	Thr26, His41, Thr25, Leu27, His164, Met165, Asn142, Cys145, Glu166, Gly143, Arg188, Gln189, Ser144, His163, Leu141, Asp187, Phe140	-	-

oxygen of Glu66 residue. A hydrogen bond interaction was recorded between Glu166 NH group and amid NH substituent of **X77** as well.

The model obtained from the interaction of **DB00224** with SARS-CoV-2 M<[pro] active site is described here. Four hydrogen bonds were provided between the molecule

and active site amino acids. The hydroxyl group of ligand formed two hydrogen bonds with the NH group and the carbonyl oxygen atom of Glu166 amino acid. Moreover, NH_2 moiety of Gln192 and the carbonyl oxygen atom of Thr190 residues formed the hydrogen bonds with hydroxyl group substituted to dihydroindene ring of the molecule. So, Ala191, Pro168, Thr190, Gln192, Arg188, Gln189, Met49, Thr25, Gly143, Ser144, Glu166, Asn142, Cys145, His41, Cys44, Leu141, His163, Met165 and His164 made hydrophobic contacts with the ligand. This ligand exhibited no π-π or π-cation interaction (Figure 4).

The investigation of all binding modes of **DB00220** exhibited that this ligand is placed in the active site. Gly143, Leu141, Ser144, Asn142, Cys145, Glu166, Met165, Pro168, Thr190, Gln189, Leu27, Cys44, His163, His41, His164, His172, Leu167, Asp187, Gln192, His172, Val42, Arg188, Thr25, Met49 and Tyr54 residues of the binding packet made hydrophobic contacts with this molecule. The molecule also made two hydrogen bonds with Gln189 and His41 amino acids. However, this molecule showed no hydrogen bond with Gln166 residue. Furthermore, it revealed π-π and cation-π stacking interactions with His172 and Thr25 residues, respectively. The orientations and interactions of **DB00220** at SARS-CoV-2 M<[pro] active site are depicted in Figure 5.

The orientation and binding mode of **DB01232** demonstrated the formation of a hydrogen bond between amide NH of ligand and the carbonyl oxygen atom of

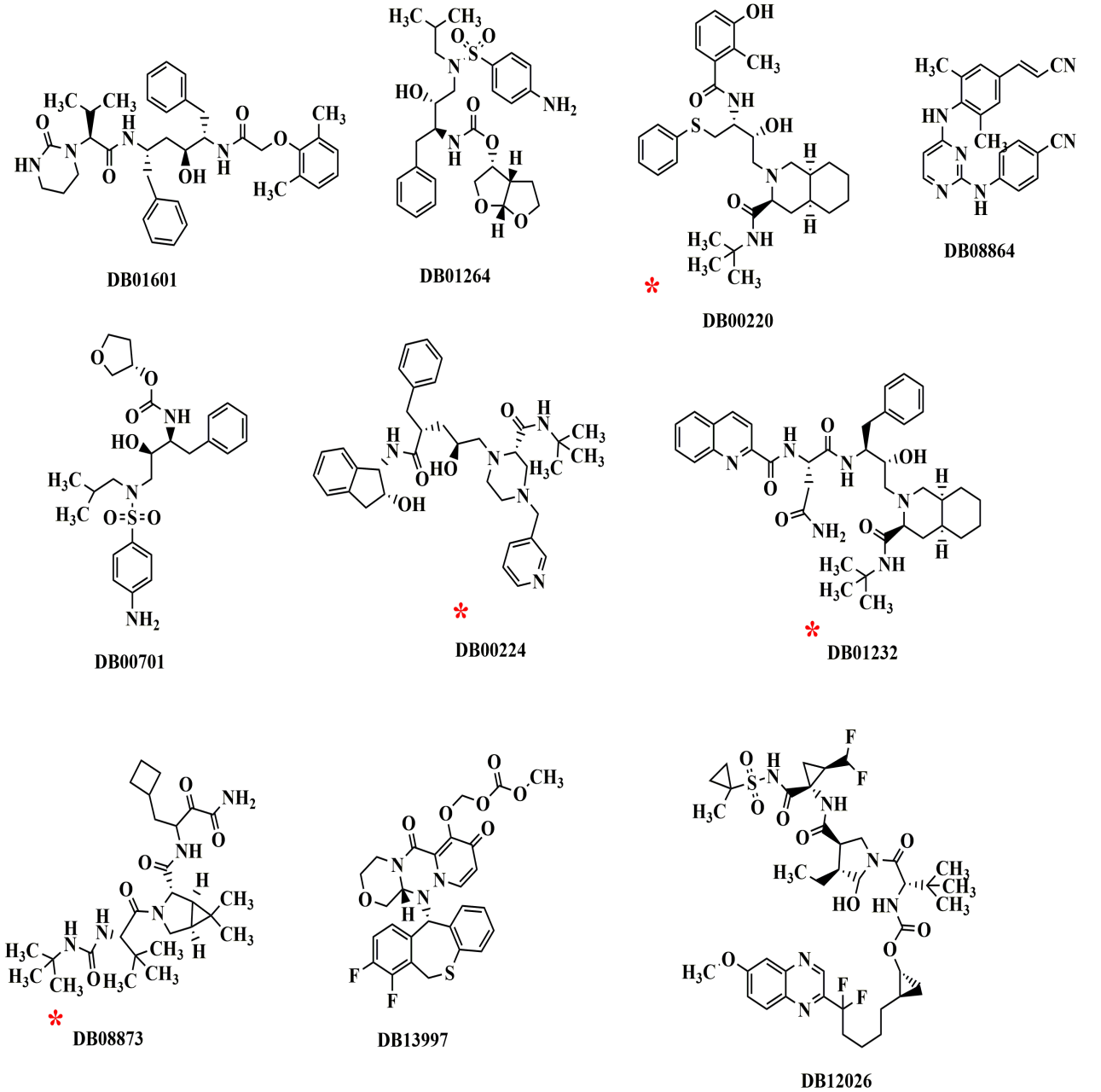

Figure 3. Ten selected compounds from one hundred compounds by molecular docking.

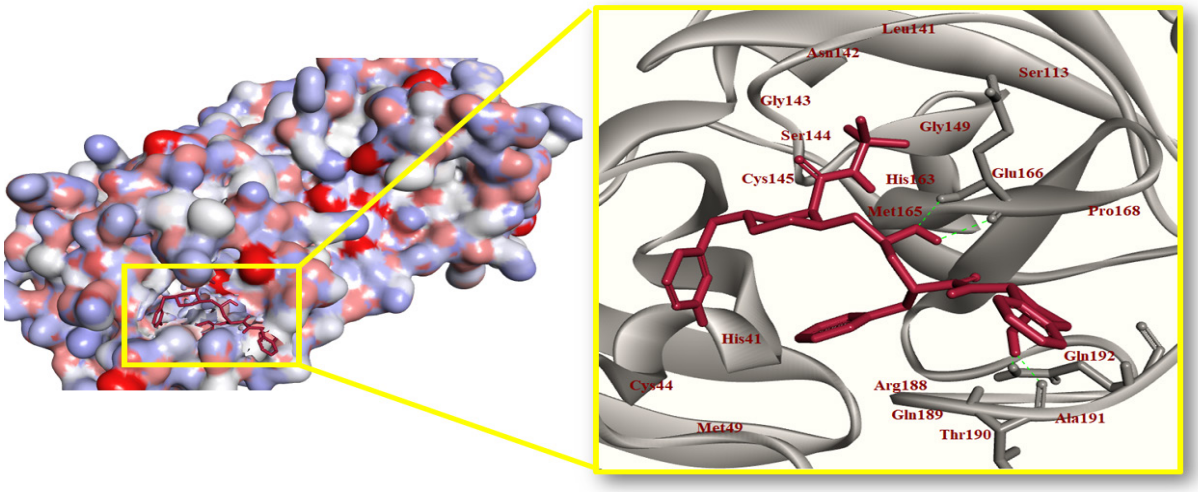

Figure 4. The best binding pose and interactions of **DB00224** in the active site of SARS-CoV-2 M<[pro].

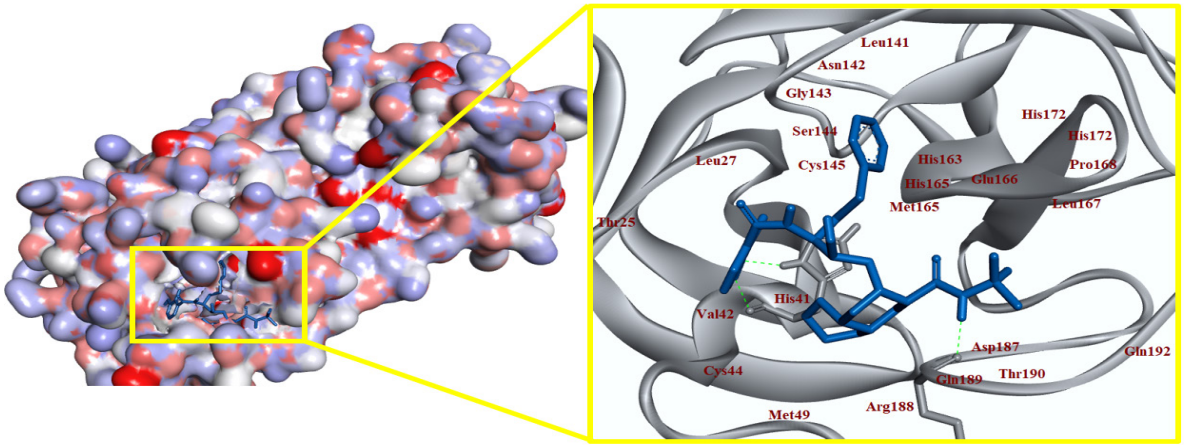

Figure 5. The best binding pose and interactions of **DB00220** in the active site of SARS-CoV-2 M<[pro].

Glu166 residue. Two hydrogen bonds were made between amino moiety and the carbonyl oxygen atom of Thr190 and Arg188 residues, respectively. The compound made hydrophobic interactions with Gln192, Thr190, Arg188, Leu167, Gly170, Gln189, Pro168, Glu166, Asp187, Met 165, Tyr54, Met49, Asn142, Leu141, Cys145, His41, Gly143, His163, Ser144 amino acids, too. In addition, **DB01232** showed a π-π stacking interaction with His41 amino acid but the cation-π stacking interaction was not detected (Figure 6). His41 is one of the important catalytic dyad residues building active site of the enzyme.[17]

Compound **DB08873** exhibited hydrophobic interactions with Arg188, Leu167, Glu166, His41, Met165, His164, Cys145, Leu141, His163, Phe140, Ser144, Gly143, Asn142, Met49, Gln189, Gln192, Pro168, Thr190 residues. Also, this compound showed three hydrogen bonds including: the carbonyl oxygen atom of His164 with NH_2 group of ligand and two hydrogen bonds between the carbonyl oxygen atom of Glu166 and NHs of amide moiety of compound (Figure 7). However, **DB08873** formed no π-π staking interaction.

According to docking calculations, Glu166, His41, His164, Cys145, Gln189, Met165, His163, Thr190, and Gln192 were vital amino acids in keeping the suitable binding mode of the compounds into SARS-CoV-2 M<[pro] because they participated in hydrogen bonds and hydrophobic interactions with most of the molecules. Top-ranked SARS-CoV-2 M<[pro] hits show interactions with the catalytic dyad (at least one strong hydrogen bond or hydrophobic interaction with either His41 or Cys145). Mutation of His41

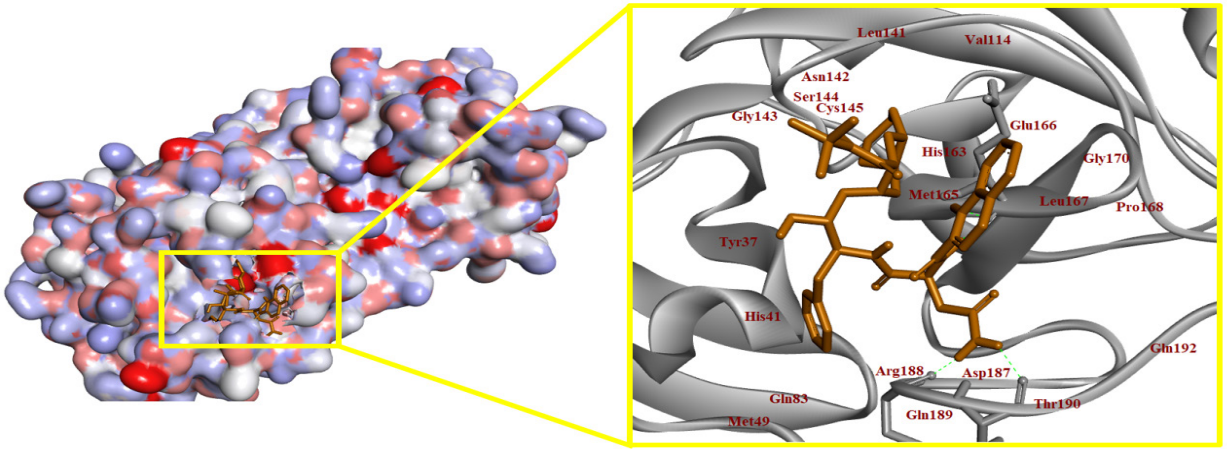

Figure 6. The best binding pose and interactions of **DB01232** in the active site of SARS-CoV-2 M<[pro].

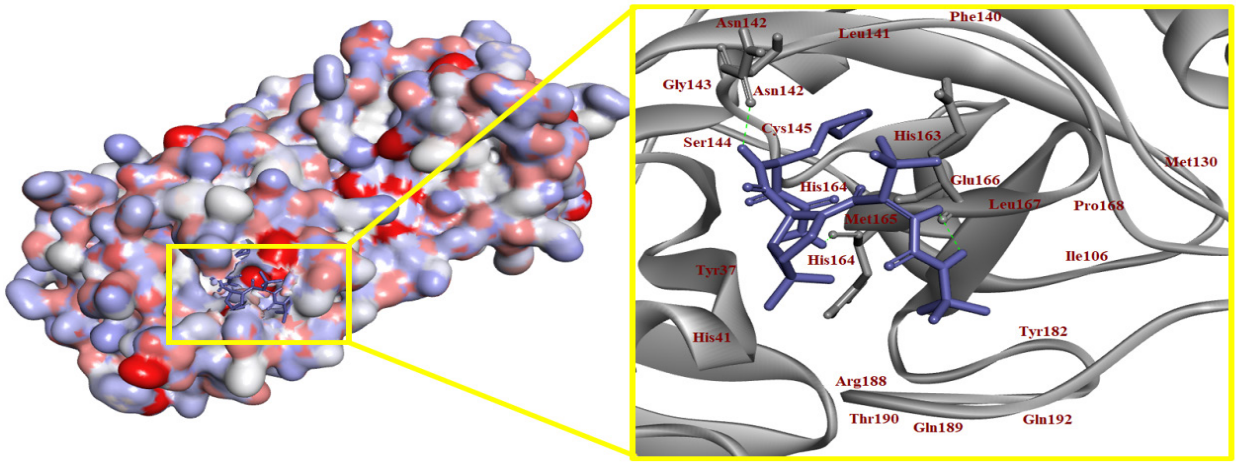

Figure 7. The best binding pose and interactions of **DB08873** in the active site of SARS-CoV-2 M<[pro].

or Cys145 is caused in the lake of enzymatic activity, which establishes the role of the catalytic dyad.[17] In the substrate-binding pocket, catalytic dyad residues, extremely preserved residues of the subsite, Gln189 and Met165 of S4, exposed comparative significance in ligand binding. Additionally, optimum hydrogen bonds, hydrophobic and π-π interactions showed a vital and significant role in the ligand-enzyme interaction.

Molecular dynamics simulation analysis

One of the computational methods to investigate the operation of biological and chemical systems when measuring their trajectory throughout a determined period is the MDS. By this technique, the correlation between the function of macromolecules and structure can be efficiently recognized.[18]

The behavior of **DB00220**, **DB00224**, **DB01232** and **DB08873** and Co-crystal ligand (**X77**) complexes were investigated by molecular dynamics simulations studies for a period of 50 ns. The aim of these studies was to observe and explore the dynamic behavior of molecules in the active site of SARS-CoV-2 M<[pro], in order to confirm the docking results.

Root mean square deviation (RMSD)

RMSD is the calculation of the average distance between the atoms of superimposed proteins, which prepares a preliminary assessment of structural drift. RMSD amounts of both ligand and protein backbone atoms were considered during simulations. RMSD values during the simulations are displayed in Figure 8. Constant backbone atom RMSD and small fluctuations are a good indication of system steadiness.[19]

In this regard, all complexes reached equilibrium and

remained steady throughout the MD calculation, it means that the complex achieved a more steady state than the initial structure. The systems with molecules **DB08873**, **DB00220**, **DB00224**, **DB01232** and **X77** reached equilibrium after 6, 3, 2, 4.5 and 5.2 ns of simulation, with the mean RMSD values of 3.10, 3.46, 3.47, 3.54, and 4.33 Å, respectively.

Overall, RMSD values showed that four complexes were more stable than the Co-crystal ligand. The less RMSD values of four complexes that exposed these compounds are stable into SARS-CoV-2 M<[pro] enzyme. The complex **DB08873** showed the lowest mean RMSD value so it is the most stable molecule during the simulation (Figure S1 in supplementary information).

The complex of **DB01232** indicated an important fluctuation within 15-22 ns and 24-27 ns representing conformational change because of spatial appropriate of the molecule in the binding packet. After the fluctuation, it kept a steady equilibrium up to the finish of the simulation time. These calculations suggest that the stabilities of the dynamic equilibriums for the complexes were reliable.

Likewise, the RMSD of molecules was also calculated individually (Figure S2 in supplementary information). The ligands **DB00220**, **DB08873**, **DB00224**, **DB01232** and **X77** in SARS-CoV-2 M<[pro] showed a stable RMSD profile with marginal deviations 2.22, 2.38, 3.75, 2.42 and 2.62 Å, respectively.

Ligand RMSD exposed that **DB08873** and **DB01232** were very stable during the simulation without large fluctuation. Molecule **DB00220** exhibited stability from 18ns to 50ns after primary fluctuations. Ligand **DB00224** exhibited the highest RMSD value among compounds. Also, **X77** showed stability similar to **DB08873**, **DB01232** and **DB00220** (Figure S2 in supplementary information).

Root mean square fluctuation (RMSF)

Averaging all the atoms of the given residue was calculated by the root mean square fluctuation (RMSF) of a given residue in the MD trajectories. The flexibility of the residues of each complex was compared to Co-crystal using the RMSF values. The flexibility of the SARS-CoV-2 M<[pro] active site (Glu166, His41, Cys145, Gln189, Met165, His163, Thr190 and Gln192) were important and they were taken into account.

The most mobile regions related to the amino acids were near the N-terminal and the loops areas. The α-helix and β-sheet areas as well as the amino acids at the binding packet areas, existing lower RMSF values, showed to be the steadiest areas of the complex. Figure S3 illustrates the considered RMSF for all complexes.

As revealed in Figure S3, all complexes had similar RMSF. For example, in **DB08873**, Glu166, His41, Cys145, Gln189, Met165, His163, Thr190 and Gln192 had maximum RMSFs of 0.79, 0.61, 0.77, 1.17, 0.70, 0.56, 1.35 and 1.20 Å, respectively. This slight range of RMSFs of the amino acids for four complexes revealed that the ligands were able to make stable bindings with the enzyme throughout MD, similar to **X77**.

According to RMSF results, His41 and Cys145 residues in the **DB00224**-M<[pro] complex fluctuated lower than other complexes, reflecting that **DB00224** could form stronger hydrophobic interactions with these residues than other complexes. For **DB00220**, the RMSF fluctuations of the amino acids Met165, His163, Thr190 and Gln192 were lower than other complexes. It could form stronger hydrophobic interactions with these key residues than other complexes. The RMSF fluctuations of the amino acids Met165 and His163 in presence of **DB08873** were the same as the Co-crystal (**X77**), indicating that the compound **DB08873** has the same function to inhibit enzyme as done by **X77**. Similarly, the RMSF fluctuations of these amino acids with **DB00224** were the same as **DB01232** demonstrating that the compound **DB00224** has the same function to inhibit enzyme as done by **DB01232** (Figure S3 in supplementary information). According to the RMSF fluctuations of amino acids, His163 showed the lowest fluctuation among key amino acids and Gln189, Thr190 and Gln192 displayed the highest fluctuation among key amino acids. Glu166, His41, Cys145 and Met165 residues revealed low fluctuation.

Radius of gyration (Rg)

The radius of gyration (Rg) technique was carried out to estimate the protein compactness amount. Lower and stable Rg value is an indicative of suitable stability and folding of the protein structure as well as an amount of its compactness. Conformational flexibility and the lack of appropriate folding of the protein is resulted from a highly fluctuating Rg.[19] Rg of SARS-CoV-2 M<[pro] was shown to know among compaction of the enzyme structure in the existence of the ligands. The average Rg values of **DB00220**, **DB00224**, **DB01232**, **DB08873** and **X77** were 2.11 nm, 2.07 nm, 2.11 nm, 2.18 and 2.10 nm, respectively. Rg values of four selected compounds as well as Co-crystal ligand were lowered after about 12ns. The Rg value of **DB08873** was increased from 3ns to 12ns then decreased after 12ns. In summary, the Rg of the backbone atoms of SARS-CoV-2 M<[pro] in the existence of chosen molecules slightly decreased throughout simulation time. It can be concluded that although ligands binding alters the flexibility of active site amino acids, they do not induce considerable totally domain motions in M<[pro], and the enzyme compactness will remain unchanged. In addition, the protein compactness reveals a suitable folding and stability of the protein structure. The Rg is illustrated in Figure S4 of supplementary information.

Hydrogen bonds analysis

The ligands-M<[pro] hydrogen bonds were evaluated through the 50 ns. The number of the hydrogen bonds is important to recognize the affinity of the molecules because hydrogen bonds have the main role in keeping and stabilizing the molecular conformations in the active site. In this study, the hydrogen bonds were observed

between each compound and SARS-CoV-2 M<[pro]. The hydrogen bonds of complexes are shown in Figure S5 of supplementary information. The analysis showed that the average number of hydrogen bonds made by ligands **DB00220**, **DB08873**, **DB00224**, **DB01232** and **X77** were 3.37 ± 1.02, 1.48 ± 0.94, 1.63 ± 1.04, 2.23 ± 1.13 and 2.01 ± 0.77, respectively.

The highest number of hydrogen bonds described is 7, 9, 5, 9 and 7 for **DB00220**, **DB08873**, **DB00224**, **DB01232** and **X77**, respectively.

Key residues of Glu166 and His41 showed hydrogen bonding contacts for ligands **DB00220** (2% and 98%), **DB00224** (63% and 25 %), **DB01232** (100% and 60%), **DB08873** (50% and 80%) and **X77** (100% and 10%), respectively. In addition, **DB01232** had stable hydrogen bond interactions with both two Glu166 and His41 residues.

To identify the binding modes and interactions of **DB00220**, **DB00224**, **DB01232**, **DB08873** and **X77** throughout MD, the 3D plans of diverse times of simulation were displayed. For example, Figures 8 and 9 showed the comparison between the compounds **DB08873** and **DB00220** in the binding site of SARS-CoV-2 M<[pro] throughout 0, 10, 20, 30, 40 and 50 ns of the simulation.

For ligand **DB08873** at the beginning of the simulation (0 ns), a hydrogen bond was formed between the O atom of the carbonyl of the ligand and amide NH group of Glu166. The residues Gly143, Ser144, Cys145, His163, His41, Glu166, His164, Met165, Leu167, Pro168, Gln189 and Arg188 perched around the ligand. During the first 10 ns, additionally formation of a hydrogen bond with Glu166, other hydrogen bonds were made with Ser144, His41 and His164 residues. The orientation changed in this ligand. After 20 ns, four hydrogen bonds were maintained with the ligand. In 30 ns, the orientation of ligand changed and hydrogen bonds were formed with three residues Ser144, His41 and His164. While, the hydrogen bond with Glu166 residue was removed. Throughout 40 ns, the binding mode of the ligand into active site formed two hydrogen bonds with His41 and His164 residues. Whereas, there are no hydrogen bonds with Glu166 and Ser144 residues; It can be due to a change in the binding mode of the ligand at this time. At the end of the simulation (50 ns), the NH moiety of imidazole ring and the O atom of the amide carbonyl group of ligand made a hydrogen bond with His41. Moreover, a hydrogen bond was formed between the amine group of ligand and the O atom of the carbonyl group of His164 residue. But base on the orientation of ligand into active site, two hydrogen bonds with Glu166 and Ser144 residurs was not formed. According to MD studies, His41 and His164 formed strong hydrogen bonds with ligand because they are in all MD times but Glu166 and Ser144 made weak hydrogen bonds led to they remained until 20 ns and 30 ns, respectively (Figure 8).

For ligand **DB00220** at the beginning of simulation (0 ns), the hydroxyl group of the ligand formed a hydrogen bond with Gln189. In this complex, the residues Asn142, Gly143, Ser144, Cys145, His163, His164, Glu166, His41, Met165, Pro168, Leu167, Gln189, Arg188, Thr190 and Gln192 perched around the ligand. In 10 ns, oxygen atom of carbonyl and hydroxyl group formed hydrogen bonds with His41 and His164 residues, respectively. During 20 ns, change of conformation of ligand led to formation of four hydrogen bonds with His41, Arg188, Thr190 and Gln192 residues. After 30 ns, due to change of orientation of ligand, two hydrogen bonds with His41 and Gln192 remained and formed a hydrogen bond with Gln189 residue. In 40 ns, there are three hydrogen bonds with His41, Gln192 and

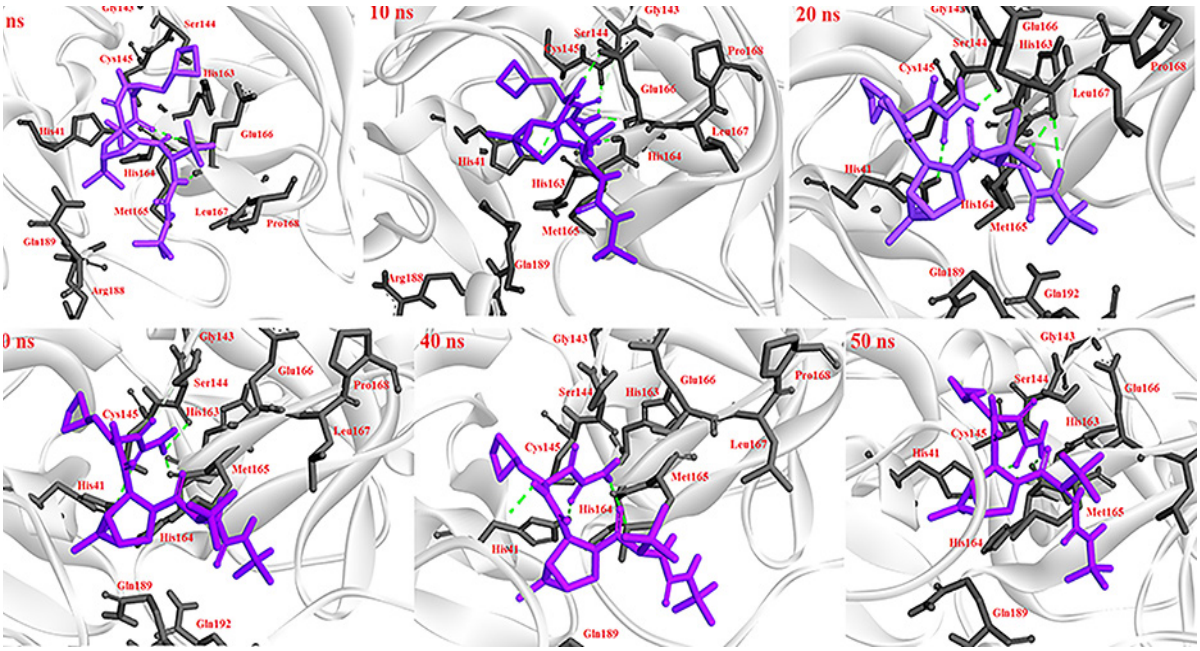

Figure 8. 3D plots of the interaction between ligand **DB08873** and M<[pro] at different times during the MD simulation.

Gln189 residues. Finally, in 50 ns, two hydrogen bonds of His41 and Gln189 with ligand remained. But a hydrogen bond with Gln192 was removed while a hydrogen bond was formed with Asp187 residue. In summary, His41 and Gln189 formed strong hydrogen bonds with ligand because they exist in almost MD times. Also, His41 formed a stroger hydrogen bond than Gln189. But other residues made weak the hydrogen bonds as they were in some MD times (Figure 9).

Solvent accessible surface area (SASA)

The surface area of the protein which is available to solvent molecules was evaluated using Solvent Accessible Surface Area (SASA). SASA has a vital role in the upkeep of protein steadiness, conformational variations and protein folding. Increase in flexibility of the protein is determined by higher values of SASA. While, reduction in flexibility and increase in compactness of the protein structure is determined by lowering of SASA value.[20] The average SASA values for SARS-CoV-2 M<[pro] bound to **DB00220**, **DB08873**, **DB00224** and **DB01232** were 134.97 nm^2, 133.22 nm^2, 133.36 nm^2, 137.67 nm^2, respectively. The SASA for the Co-crystal (**X77**) protein complex was 134.59 nm^2. Based on SASA results and comparison of molecules bound SASA with Co-crystal, it is identified that the enzyme has no remarkable changes in SASA value and the ligands bound enzyme remains stable. The SASA of most of the residues of the hits-bound to SARS-CoV-2 M<[pro] were similar to the Co-crystal-M<[pro], suggesting that hits binding did not alter the SASA of the M<[pro] residues. The SASA plot is represented in Figure S6 of supplementary information.

The binding of **DB01232**, **DB00224** and **DB08873** increased the SASA of Glu166 (0.084, 0.097 and 0.087 nm^2) and His41 (0.0419, 0.0382 and 0.0253 nm^2) in compared to **X77**. The SASA of Thr190 increased in presence of **DB00224**, **DB08873** and **DB00220** binding up to 0.0948, 0.1011 and 0.1164 nm^2, respectively. The binding of **DB01232**, **DB08873** and **DB00220** increased the SASA of Met165 up to 0.0487, 0.0558 and 0.0894 nm^2. Graphical representation of the SASA of the residues of **DB00220**, **DB08873**, **DB00224**, **DB01232** and **X77** in complex with SARS-CoV-2 M<[pro] have been shown in Figure S7 of supplementary information.

The average ligand SASA values for **DB00220**, **DB08873**, **DB00224**, **DB01232** and **X77** in SARS-CoV-2 M<[pro] were 7.84 nm^2, 7.54 nm^2, 9.54 nm^2, 8.76 nm^2 and 7.14 nm^2, respectively. The ligand SASA plot is represented in Figure S8 of supplementary information.

Binding affinity estimation

The binding energy is estimated to quantify the tendency of the molecule to interact with the active site of M<[pro]. The energy components such as Van der Waals, electrostatic, SASA and polar are listed in Table 3. The free binding energies for SARS-CoV-2 M<[pro] were calculated for selected four compounds and **X77** using the MM/PBSA technique. The estimated free binding energies were comparable: -286.053, -182.107, -208.482, -311.317 and -265.039 kJ/mol for **DB00220**, **DB**08873, **DB00224**, **DB01232** and **X77**, respectively.

The molecule **DB01232** had the lowest non-polar portion among molecules, i.e., ΔE_{vdw} = -358.218 and ΔE_{SASA} = - 27.007 kJ/mol. This difference can be relevant to more appropriate Van der Waals interactions and somewhat bigger solvent accessibility through the effects of inhibition using this molecule. A lower electrostatic portion was observed for derivative **DB01232** (ΔE_{elect} = -56.550 kJ/mol) in comparison to **DB00220** (ΔE_{elect} = -75.611 kJ/mol).

The free polar solvation energy was added certainly to the total free binding energy with ΔE_{polar} = +78.185, +81.467, +84.090, +130.166 and +94.718 kJ/mol for **DB08873**, **DB00220**, **DB00224**, **DB01232** and **X77**, respectively. This

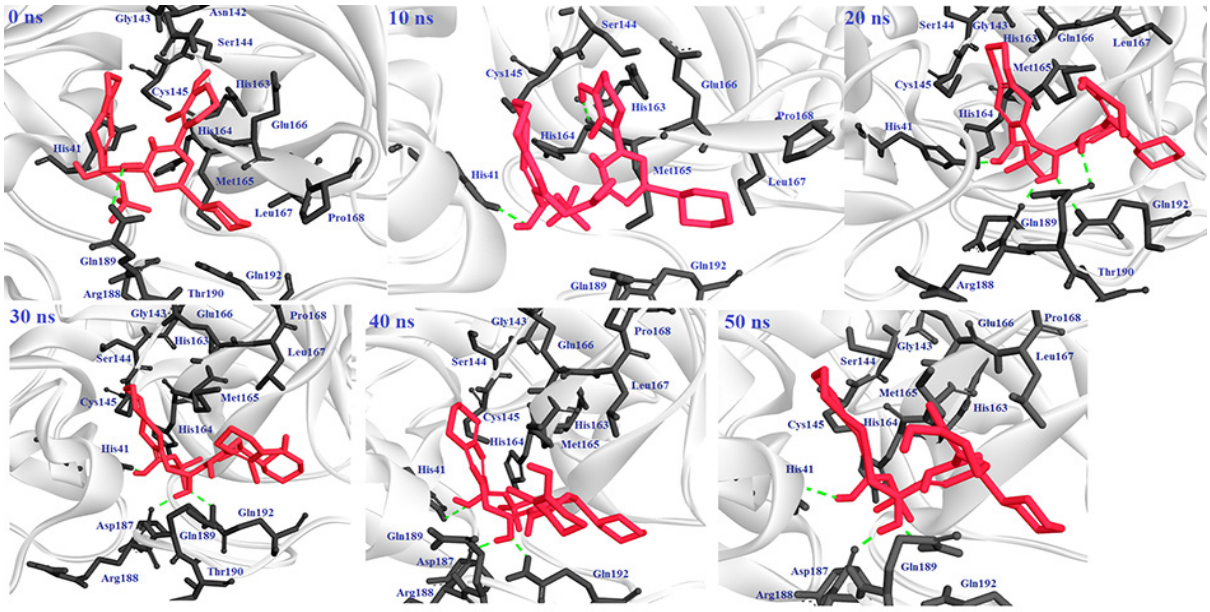

Figure 9. 3D plots of the interaction between ligand **DB00220** and M<[pro] at different times during the MD simulation.

Table 3. Molecular energy terms for interactions of four compounds and **X77** with SARS-CoV-2 M<[pro].

Terms	Energy (KJ/mol)				
	DB00220	**DB08873**	**DB00224**	**DB01232**	**X77**
ΔG_{bind}	-286.053	-182.107	-208.482	-311.317	-265.039
ΔE_{elect}	-75.611	-31.647	-46.702	-56.550	-28.760
ΔE_{vdw}	-270.348	-208.516	-226.270	-358.218	-307.808
ΔE_{polar}	81.467	78.185	84.090	130.166	94.718
ΔE_{SASA}	-21.640	-19.970	-19.606	-27.007	-23.156

showed easier solvation of **DB08873** through inhibition than other ligands. The total polar portions, $\Delta E_{elect} + \Delta E_{polar}$, were inappropriate for binding with SARS-CoV-2 M<[pro] for all molecules. However, the non-polar portions, $\Delta E_{vdw} + \Delta E_{SASA}$, contributed more positively resulting in a totally suitable compound binding. The appropriate non-polar contribution was most probable because of the hydrophobic surrounding of the packet. The MM/PBSA investigations indicate that **DB01232** can equally bind appropriable and powerfully to the M<[pro] active site toward other compounds.

Conclusion

In summary, first structure-based virtual screening was performed on the crystal structure of SARS-CoV-2 M<[pro]. Some anti-viral compounds were given from DrugBank database. Thus, molecular docking was carried out to discover the details of the interactions in the candidate molecules into the active site. Lastly, MD simulations on **DB00220**, **DB00224**, **DB01232**, **DB08873** as well as **X77** (Co-crystal) in complex with SARS-CoV-2 M<[pro] were performed at 50 ns. Additionally, RMSD, RMSF, hydrogen bonds, Rg, SASA and energy analysis through MD surely showed the steady binding of candidate molecules with SARS-CoV-2 M<[pro] structure. According to calculations of docking and MD, it was observed that the active site is mostly hydrophobic, where the value of the ΔE_{vdw} is higher than that of the ΔE_{ele}. Finally, the MM/PBSA results correlated with the analyses of docking and MD.

Acknowledgements

Supports of this project by Ardabil University of Medical Sciences are welcomed. Funding Number is IR.ARUMS.REC.1399.469.

Author Contributions

NH and HA: Carried out the experimental studies; KM: Carried out the experimental studies and participated in analysis of the data; SS: Supervised and designed of the project, analyzed the data and prepared manuscript. All authors approved the final version of the manuscript.

Conflict of Interest

Authors declare that there are no conflict of interests.

Supplementary Data

Supplementary data (Figures S1-S4) are available on the journal's web site along with the published article.

References

1. Khaerunnisa S, Kurniawan H, Awaluddin R, Suhartati S, Soetjipto S. Potential inhibitor of COVID-19 main protease (Mpro) from several medicinal plant compounds by molecular docking study. Preprint. 2020;2020030226. doi:10.20944/preprints202003.0226.v1
2. Ali I, Alharbi OML. COVID-19: Disease, management, treatment, and social impact. Sci Total Environ. 2020;728:138861. doi:10.1016/j.scitotenv.2020.138861
3. Elfiky AA. Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. Life Sci. 2020;248:117477. doi:10.1016/j.lfs.2020.117477
4. Coronavirus disease (COVID-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019.
5. Li G, Clercq ED. Therapeutic options for the 2019 novel coronavirus (2019-nCoV). Nat Rev. 2020;19(3):149-150. doi:10.1038/d41573-020-00016-0
6. Zumla A, Chan JF, Azhar EI, Hui DS, Yuen KY. Coronaviruses-drug discovery and therapeutic options. Nat Rev Drug Discov. 2016;15(5):327-47. doi: 10.1038/nrd.2015.37
7. Xu Z, Peng C, Shi Y, Zhu Z, Mu K, Wang X. Nelfinavir was predicted to be a potential inhibitor of 2019-nCov main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. BioRxiv. 2020. doi: 10.1101/2020.01.27.921627
8. Elfiky AA. Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir against SARSCoV-2 RNA dependent RNA polymerase (RdRp): A molecular docking study. Life Sci 2020;253:117592. doi:10.1016/j.lfs.2020.117592
9. Rahman MM, Saha T, Islam KJ, Suman RH, Biswas S, Rahat EU, et al. Virtual screening, molecular dynamics and structure–activity relationship studies to identify potent approved drugs for Covid-19 treatment. J Biomol Struct Dyn. 2020;21:1-11. doi: 10.1080/07391102.2020.1794974
10. Mirza MU, Froeyen M. Structural elucidation of SARS-

CoV-2 vital proteins: Computational methods reveal potential drug candidates against main protease, Nsp12 polymerase and Nsp13 helicase. J Pharm Anal. 2020;10(4):320-8. doi:10.1016/j.jpha.2020.04.008
11. Macchiagodena M, Pagliai M, Procacci P. Identification of potential binders of the main protease 3CL^{pro} of the COVID19 via structure-based ligand design and molecular modeling. Chem Phys Lett. 2020;750:137489. doi:10.1016/j.cplett.2020.137489
12. Chandel V, Raj S, Rathi B, Kumar D. In silico identification of potent FDA approved drugs against Coronavirus COVID-19 main protease: A drug repurposing approach. Chem Biol Lett. 2020;7(3):166-175.
13. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009;30:2785-91. doi:10.1002/jcc.21256
14. Sepehri S, Saghaie L, Fassihi A. Anti-HIV-1 activity prediction of novel gp41 inhibitors using structure-based virtual screening and molecular dynamics simulation. Mol Inform. 2017;36:1600060. doi:10.1002/minf.201600060
15. Razzaghi-Asl N, Mirzayi S, Mahnam K, Sepehri S. Identification of COX-2 inhibitors via structure-based virtual screening and molecular dynamics simulation. J Mol Graph Model. 2018;83:138-152. doi:10.1016/j.jmgm.2018.05.010
16. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018;46(D1):D1074-82. doi:10.1093/nar/gkx1037
17. Yang H, Yang M, Ding Y, Liu Y, Lou Z, Zhou Z, et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc Natl Acad Sci U S A. 2003;100(23):13190-5. doi:10.1073/pnas.1835675100
18. Ropon-Palacios G, Chenet-Zuta ME, Olivos-Ramirez GE, Otazu K, Acurio-Saavedra J, Camps I. Potential novel inhibitors against emerging zoonotic pathogen nipah virus: A virtual screening and molecular dynamics approach. J Biomol Str Dyn. 2020;38(11):3225-34. doi:10.1080/07391102.2019.1655480
19. Zhou H, Wang C, Ye J, Chena H, Tao R. Design, virtual screening, molecular docking and molecular dynamics studies of novel urushiol derivatives as potential HDAC2 selective inhibitors. Gene. 2017;637:63-71. doi:10.1016/j.gene.2017.09.034
20. Yang WL, Wang J, Chan CH, Lee SW, Campos AD, Lamothe B, et al. The E3 ligase TRAF6 regulates akt ubiquitination and activation. Science. 2009;325:1134-38. doi:10.1126/science.1175065