



# ***In-silico* study MM/GBSA binding free energy and molecular dynamics simulation of some designed remdesivir derivatives as the inhibitory potential of SARS-CoV-2 main protease**

**Maryam Abbasi<sup>1</sup>, Mahboubeh Mansourian<sup>2,3</sup>, Afsaneh Arefi Oskouie<sup>4,\*</sup>, Salman Taheri<sup>5</sup>, and Karim Mahnam<sup>6</sup>**

<sup>1</sup>Department of Pharmaceutical Chemistry, School of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran.

<sup>2</sup>Medicinal Plants Research Center, Yasuj University of Medical Sciences, Yasuj, I.R. Iran.

<sup>3</sup>Department of Pharmacology, Faculty of Medicine, Yasuj University of Medical Sciences, Yasuj, I.R. Iran.

<sup>4</sup>Department of Basic Sciences, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran.

<sup>5</sup>Chemistry & Chemical Engineering Research Center of Iran, P.O. Box 14335-186, Tehran, I.R. Iran.

<sup>6</sup>Faculty of Science, Department of Biology, Shahrekord University, Shahrekord, I.R. Iran.

## **Abstract**

**Background and purpose:** Coronavirus disease (COVID-19) is one of the greatest challenges of the twentieth century. Recently, *in silico* tools help to predict new inhibitors of SARS-CoV-2. In this study, the new compounds based on the remdesivir structure (12 compounds) were designed.

**Experimental approach:** The main interactions of remdesivir and designed compounds were investigated in the 3CL<sup>pro</sup> active site. The binding free energy of compounds by the MM-GBSA method was calculated and the best compound (compound 12 with the value of -88.173 kcal/mol) was introduced to the molecular dynamics simulation study.

**Findings/Results:** The simulation results were compared with the results of protein simulation without the presence of an inhibitor and in the presence of remdesivir. Additionally, the RMSD results for the protein backbone showed that compound 12 in the second 50 nanoseconds has less fluctuation than the protein alone and in the presence of remdesivir, which indicates the stability of the compound in the active site of the M<sup>pro</sup> protein. Furthermore, protein compactness was investigated in the absence of compounds and the presence of compound 12 and remdesivir. The Rg diagram shows a fluctuation of approximately 0.05 Å, which indicates the compressibility of the protein in the presence and absence of compounds. The results of the RMSF plot also show the stability of essential amino acids during protein binding.

**Conclusion and implications:** Supported by the theoretical results, compound 12 could have the potential to inhibit the 3CL<sup>pro</sup> enzyme, which requires further *in vitro* studies and enzyme inhibition must also be confirmed at protein levels.

**Keywords:** 3CL<sup>pro</sup>; Main protease; MM-GBSA; Molecular docking; Molecular dynamics simulation; Remdesivir.

## **INTRODUCTION**

Coronaviruses are most commonly found in birds and mammals and mainly cause respiratory, gastrointestinal, and sometimes neurological diseases or hepatitis. Infections are usually transmitted through the respiratory tract or mouth, and infections can be acute or

chronic. Documents show coronaviruses emerged as a new viral family in the 1960s, following the discovery of several pathogens (1).

### **Access this article online**



**Website:** <http://rps.mui.ac.ir>

**DOI:** 10.4103/1735-5362.394818

\*Corresponding author: A. Arefi Oskouie

Tel: +98-2122718531, Fax: +98-2122721150

Email: a.arefi@sbmu.ac.ir

Historically, in late 2002, the appearance of the coronavirus caused severe respiratory problems in human's so-called severe acute respiratory syndrome coronavirus (SARS-CoV) became famous. The sudden emergence of SARS led to new research to understand the main mechanisms of reproduction and the pathogenicity of members of this viral family (2). The new human coronavirus, first announced at the end of 2019, is called COVID-19 and causes severe respiratory syndrome as SARS-CoV-2 (3).

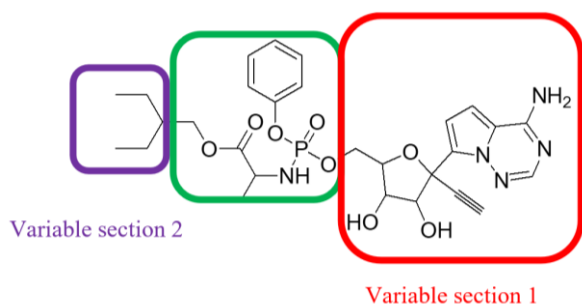
Researchers have shown how the virus enters the host body. The coated single-stranded RNA virus binds to host cells from the spikes of S protein to the angiotensin-converting enzyme 2 (ACE2) receptor. The virus connects to the host cell and endosome receptors enter the cells (4). The SARS-CoV-2 can attach to the peptidase domain in ACE2 directly. Transmembrane protease serine 2 facilitates virus entry through the S protein. ACE2 degradation also confirms the high affinity of the ACE to S protein (5). Viral polyproteins are made inside the cell, and then these proteins are encoded for the transcriptase-replicase complex. After that, RNA of structural proteins synthesized by RNA-dependent RNA polymerase (RdRp) leads to the completion of the assembly, synthesis, and release of viral particles. Potential targets for drug therapy are designed based on stages of the viral life cycle (6,7). The site of action and binding of the drugs introduced and used so far for COVID-19 include non-structural proteins such as 3-chymotrypsin-like protease (3CL<sup>pro</sup>) as main protease (M<sup>pro</sup>), papain-like protease, RdRp, ACE2 receptor, spike receptor-binding domain, hemagglutinin esterase, 2'-O-methyltransferase, helicase, structural and glucose-regulated proteins (8).

Moreover, the main protease (3CL<sup>pro</sup>) significantly coordinated the transcription of the virus life cycle and viral replications. This protease is involved in the cleavage of the main fragment of viral polyprotein so that it releases proteins with replicative functions like RdRp and RNA processing domain (9). It has been proven to be highly conservative sequential between SARS-CoV and SARS-CoV-2 and is currently the most widely used prime target for drugs for SARS-CoV-2 (10-14). 3CL<sup>pro</sup> is a

cysteine protease that is composed of three different domains (15). Other studies have shown the antiparallel structures of the  $\beta$ -sheet of domains I (8-101 residues) and II (102-184 residues) and the existence of 5  $\alpha$ -helices with the parallel connection with domain II through a long loop (185-200 residues) in domain III (201-303 residues) (15). In addition, the location of the substrate binding site of 3CL<sup>pro</sup> has been confirmed to be between the domains I and II. Protein dimerization is performed by the residue Glu166 and catalytic dyads Cys145 and His41 (14,16,17). Some studies have also demonstrated the existence of six sub-units in the active site of the enzyme (S1-S6), and the existence of residues 140-145 as well as 163-166 in domain II in the residues of the active site (14,18).

Numerous drugs have been reported and used to date, including oseltamivir and zanamivir as neuraminidase inhibitors (19). lopinavir (20), ritonavir (21), remdesivir (22), favipiravir (23), ribavirin (24), chloroquine (25), and hydroxychloroquine (26). Most of the drugs introduced are HIV protease inhibitors. Comostat mesylate, as an inhibitor of the transmembrane protease serine 2 receptor, blocks the entry of SARS-CoV-2 into human cells, representing its therapeutic ability as a drug in contrast to COVID-19 (27). Today, the U.S. Food and Drug Administration (FDA) issued an emergency use authorization (EUA) for Pfizer's Paxlovid and molnupiravir (EIDD-2801) as a new medicine for treating mild-to-moderate COVID-19; although, approved the effectiveness of conventional several drug medications such as remdesivir (28), avifavir (14) and dexamethasone (29).

Remdesivir, a small molecule, is a monophosphoramidate prodrug that metabolizes the nucleoside C-adenosine triphosphate analog that inhibits viral RNA polymerase with RdRp as its target (30). Gilead Sciences first revealed remdesivir in 2009 as an antimicrobial drug with activity against RNA viruses (like Coronaviridae and Flaviviridae). Remdesivir effectively inhibits the activity of SARS-CoV-2 *in vitro* (31). Remdesivir is the first treatment for COVID-19 that was approved by the FDA (32), but its effectiveness is disputed (33), emphasizing the need to develop new antiviral drugs (34).



**Fig. 1.** The chemical structure of remdesivir.

Nguyen *et al.* performed molecular docking, steered molecular dynamics, and umbrella sampling for remdesivir (14). According to the documents obtained from molecular simulation, remdesivir has a strong binding to both 3CL<sup>pro</sup> and RdRp of SARS-CoV-2. It showed relatively weak binding to 3CL<sup>pro</sup> compared to RdRp (14,35). The chemical structure of remdesivir (Fig. 1) is the same as ATP (the natural substrate of RdRp), and the remdesivir active form competes with ATP for incorporation into the growing RNA chain (36).

In this study, the new compounds based on the remdesivir structure were designed. The main interactions of remdesivir and designed compounds were investigated in the 3CL<sup>pro</sup> active site. The binding free energy of compounds was calculated and the best compound was introduced to molecular dynamics (MD) simulation.

## METHOD AND MODELING

This study was performed in two parts. In the first part, after designing molecular compounds, screening was performed using molecular docking and binding free energy studies. In the second part, more detailed information was obtained by MD simulation.

### *Protein and ligand preparation*

The crystal structure of the 3CL<sup>pro</sup> (M<sup>pro</sup>) enzyme (PDB ID: 6LU7) with 306 amino acids at a resolution of 2.16 Å was selected (37). Chain A of the enzyme was selected for docking, and inhibitor N3 and water molecules were removed from the structure then, protein preparation was done

using the Protein Preparation Wizard module in Schrodinger suite 2015 (38). First, bond orders were assigned, disulfide bonds were created, and missing side chains were filled. Then, the minimization energy of protein was carried out using the OPLS-AA\_2005 force field.

The 2D structure of the designed ligands (Table 1) was plotted and optimized using the Ligprep module in the Schrodinger suite 2015 (39). In this part, the ionization state was considered for each input compound at the physiological pH ( $7.4 \pm 0.5$ ). Finally, ligand optimization was performed using the OPLS-AA\_2005 force field to obtain the best conformer for each ligand.

### *Pharmacokinetic and physicochemical properties prediction*

Pharmacokinetic properties are known as absorption, distribution, metabolism, excretion, and toxicity properties (ADMET) (40). In this stage, to achieve drug-like molecules, ADMET properties were calculated in the drug discovery process using the QikProp module in Schrodinger suite 2015. In addition to Lipinski rules, the main criteria were investigated such as membrane permeability, lipophilicity, human oral absorption, cardiotoxicity, and potential interaction with hERG channels (40).

### *Molecular docking*

At first, receptor grid generation was done to create a box for the docking process using the Glide module in Schrodinger suite 2015 (41). The center of the docking box was selected based on the cysteine-histidine catalytic dyad amino acids Cys145 and His41, and the Glu166 residue (42,43). The docking process was done using the Ligand Docking module in Schrodinger Suite 2015. In the docking procedure, the ligand was considered a flexible molecule for rotating all rotatable ligand bonds to obtain the best optimal ligand conformer within the active sites of the receptor. Finally, docking results were saved using the XP GlideScore scoring function.

**Table 1.** Chemical structure of the new remdesivir derivatives.

Entry	Structure of ligand	Entry	Structure of ligand
L1		L7	
L2		L8	
L3		L9	
L4		L10	
L5		L11	
L6		L12	

### **Ligand binding energy calculation using molecular mechanics/generalized born solvent area method**

The binding free energy is a dependable criterion to rank ligands in terms of their binding affinities than XP GlideScore (44). The binding free energy ( $\Delta G_{\text{bind}}$ ) between ligand-receptor is a combination of molecular mechanical energy and solvent effects (polar and non-polar solvent) (45). In this study, the binding free energy was calculated using the molecular mechanics/generalized born solvent area (MM-GBSA) technique in the Prime module of the Schrödinger suite 2015 (46). The best pose of the ligand-protein complex was chosen to calculate energy using the OPLS-2005 force field and GBSA. The binding free energy of each ligand was calculated by the following equation:

$$\Delta G_{\text{binding}} = G_{\text{Complex}} - (G_{\text{Ligand}} + G_{\text{Protein}})$$

### **MD simulations**

MD simulations were performed for the free 3CL<sup>pro</sup> and in complex with remdesivir and the best-designed compound to study the overall effects of this ligand on the enzyme. Information on the effects of enzyme fluctuation and rearrangement of active site residues was obtained. Therefore, two separate phases of MD simulation were performed in the absence and presence of docked ligands as previously described (47-49). For simulation, the crystallographic structure of the 3CL<sup>pro</sup> (the PDB code 6LU7) and the best complexes derived from MM-GBSA calculations were introduced to the GROMACS-2021.5 package (50). Amber99.sb force field was used to create the protein topology parameters and the ligand topology parameters were created using the AnteChamber Python Parser InterfacE (ACPYPE) (48,49,51).

The complex was solvent-coated in a dodecahedron box with a TIP3P water model (52). Na<sup>+</sup> or Cl<sup>-</sup> ions were randomly substituted with water molecules to neutralize the system. The linear constraint solver (LINCS) algorithm was used to limit the length of hydrogen-containing bonds (53). Electro-static interactions were calculated using the particle Mesh Ewald (PME) method (54). In the first step, the whole system was minimized using the steepest descent algorithm and then with the

conjugate gradient algorithm. In the next step or equilibration step, 500 ps of MD simulation at NVT ensemble and then 500 ps at NPT ensemble was performed by position restraint of the protein atoms.

To achieve equilibrium at 300 K and 1 bar pressure, the system was heated at low-temperature coupling ( $\tau = 0.1$  ps) and pressure coupling ( $\tau = 1$  ps). The Berendsen algorithm was selected for the thermostat and the barostat in the equilibrium phase (55). In the final step or production step, 100 ns of MD simulation was performed at 300 K with a time step of 2 fs without position restraint of protein atoms. The thermostat and barostat for the production step were the Nose-Hoover thermostat and the Parrinello-Rahman Barostat (55). VMD (56) and PyMol Tcl (57) were used for visualization of the results.

## **RESULTS**

### **Docking validation and designing remdesivir analogs**

To validate the docking process and investigate the main interactions between the crystallographic ligand (N3) and M<sup>pro</sup>, a molecular docking study was performed. According to the findings, the catalytic His163 and Cys145 residues and Gly143 and Glu166 amino acids are known as the main residues in the bottom and the edges pocket of M<sup>pro</sup>, respectively (42). The docking study analysis showed that all reported residues have formed hydrogen bonds with ligand N3. Also, the main hydrophobic amino acids located around ligand N3 such as His41, Phe140, Lue141, Asn142, Ser144, His164, Met165, His172, and Gln189 (Fig. 2).

To predict new remdesivir derivatives, the main interactions between remdesivir and M<sup>pro</sup> were investigated using a docking study. As shown in Fig. 3, the oxygen atoms in phosphoramidate and carbonyl groups played the main role in the hydrogen bond formation with Cys145, Gly143, and Glu166 residues. Thus, these parts of the remdesivir molecule were kept and other parts changed to reach the new remdesivir derivatives with the best interaction and binding energy than remdesivir. The commutable parts of the remdesivir were shown in Fig. 1 and new remdesivir derivatives were reported in Table 1.

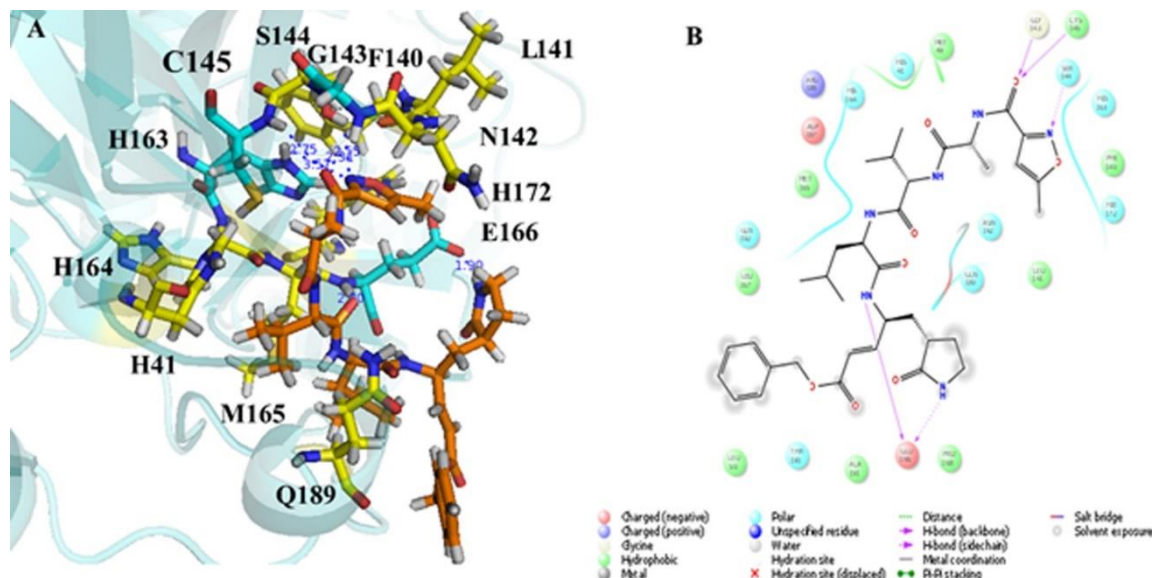


Fig. 2. The main interactions ligand N3 in M<sup>pro</sup> active site. (A) 3D-plot and (B) 2D-plot.

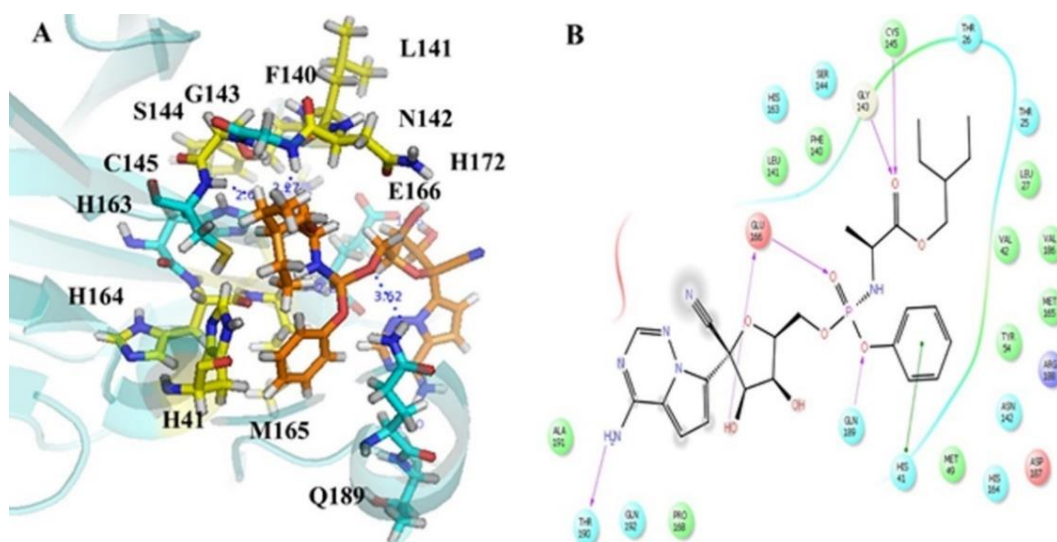


Fig. 3. The main interactions of remdesivir in the M<sup>pro</sup> active site. (A) 3D-plot and (B) 2D-plot.

### Pharmacokinetic and physicochemical properties prediction

The pharmacokinetic and physicochemical properties of the designed compounds were calculated by the QikProp package. Pharmacokinetic and physicochemical parameters are as follows: donor hydrogen bonds, acceptor hydrogen bonds, dipole moment, total solvent accessible surface area (SASA), The octanol/water partition coefficient (QlogP O/W), IC<sub>50</sub> of blocking hERG K<sup>+</sup> channels (Qplog HERG), cell permeability of Caco-2 (Qplog Caco) and MDCK (QPPMDCK) in nm/sec, brain/blood partition coefficient (Qplog BB), human serum albumin

binding (Qplog Khsa), and percent human oral absorption.

The obtained results are shown in Table 2. All designed compounds with appropriate pharmacokinetic properties passed the screening of ADME prediction.

### Molecular docking analysis

Molecular docking was performed to find the best position of ligands in the 3CL<sup>pro</sup> enzyme. Among the various conformations that were selected, one conformation with the best state and the highest interactions with the M<sup>pro</sup> active site, and the lowest binding energy (kcal/mol) was chosen.

**Table 2.** Pharmacokinetic and physicochemical parameters of designed compounds.

Compound ID	DHB	AHB	Di	SASA	Qplog O/W	Qplog HERG	Qplog Caco	Qplog BB	QPPM DCK	Qplog Khsa	%HOA
L1	2	10	3.43	715.12	3.5	-5.91	1225.17	-0.92	1062.06	-0.08	100
L2	2	10	4.19	724.74	3.77	-5.76	1230.67	-0.97	1015.30	0	100
L3	2	10	3.96	731.38	3.97	-5.76	1825.55	-0.79	1702.01	0.02	95.61
L4	3	11.5	6.60	846.77	2.41	-5.98	226.62	-2.06	174.70	-0.55	83.19
L5	2	11.5	7.02	886.16	3.08	-6.17	410.33	-1.82	305.31	-0.39	91.74
L6	2	11	5.16	992.82	4.94	-7.41	909.34	-1.64	570.56	0.18	95.84
L7	2	11.5	2.49	867.70	3.24	-5.70	883.05	-1.61	460.27	-0.39	100
L8	2	11.5	5.86	944.08	3.72	-6.08	411.17	-1.91	315.72	-0.15	82.57
L9	2	12.5	7.54	850.26	2.09	-4.67	57.90	-2.80	34.63	-0.54	44.79
L10	2	11.5	7.14	936.23	3.91	-6.09	411.42	-1.73	779.98	-0.19	83.68
L11	2	11.5	7.26	938.02	3.93	-6.07	390.87	-1.75	791.59	-0.18	83.41
L12	2	12.25	7.06	949.08	3.50	-6.07	411.01	-1.97	315.67	-0.30	81.28
References	0-6	2-20	1-12.5	300-1000	-2-6.5	< -5	< 25 poor > 500 great	3-1.2	< 25 poor > 500 great	-1.5-1.5	> 80% high < 25% poor

DHB, Donor hydrogen bond; AHB, acceptor hydrogen bond, Di, dipole moment; SASA, solvent accessible surface area; Qplog BB, brain/blood partition coefficient; Qplog Khsa, human serum albumin binding, HOA, human oral absorption.

Investigation of the docking results showed that all of the studied compounds sat in the M<sup>pro</sup> active site and the XP GScore of all of them was negative into 3CL<sup>pro</sup>. The lowest XP GScore and the involved catalytic residues in the interaction between ligands and M<sup>pro</sup>, the most important energy contributions, and the average distance between designed ligands and reported residues are mentioned in Table 3.

The docking results showed that the catalytic residues formed hydrogen bonds with all ligands, except ligands 5 and 6 (L5 and L6). In ligands 12 and 8, the values of XP GScore were obtained more than remdesivir, but other compounds had the XP GScore less than remdesivir.

### MM-GBSA results

The obtained binding affinities of docking calculations often are not trusty measuring criteria to rank compounds (58). In the ligand binding energy calculations, adding solvation energy and surface accessibility area can lead to more acceptable accuracy in prioritizing them (45). In this study, the ranking of the designed compounds was performed using MM-GBSA calculations. The  $\Delta G_{\text{binding}}$  values were reported in Table 4. The lowest value (L12) was chosen as the best compound and was introduced to the MD simulation study for more investigation.

### MD simulations

*In-silico* screening and molecular dynamics simulations could improve information in deciphering functional mechanisms of complex situations and also help to design new compounds to treat diseases. Recently MD simulations have been used to investigate the main interactions between protein-ligand to predict new compounds for COVID-19 treatment (59,60). In this study, a variety of characteristics was also employed such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), and radius of gyration (Rg), in the course of a 100-ns simulation period for examining flexibility and stability of the 3CL<sup>pro</sup>-ligand complex.

### RMSD analysis

To examine global changes in the 3CL<sup>pro</sup> conformation as a result of the presence of remdesivir in the binding pocket, the RMSD of the protein backbone relative to the initial structure as a function of MD simulation time was measured. Graphs of RMSD values are shown for the backbone 3CL<sup>pro</sup> enzyme in the absence and presence of remdesivir and L12 during 100 ns of MD simulation in Fig. 4. As can be seen in the RMSD diagrams, during the first 45 ns of the simulation, the protein was stable, but 0.3 nm fluctuation was seen in the last 55 ns. The RMSD values of backbone atoms of ligand L12 gradually increased during the first 45 ns and reached an equilibrium value with 0.1 nm fluctuation until the end of 100 ns.

**Table 3.** Energy-based interactions details of the studied ligands.

Ligands	XP GScore (kcal/mol)	Main catalytic residues	Main contribution of energy (kcal/mol)	Distance (Å)
<b>Remdesivir</b>	-7.014	Cys145 Gly143 Glu166	Vdw <sup>a</sup> (-2.058)	2.149 1.620 2.075
			Coul <sup>b</sup> (-0.573)	
			Vdw (-1.960)	
			Coul (-0.615)	
<b>L1</b>	-6.141	Gly143	Vdw <sup>a</sup> (-1.205)	2.050
			Coul <sup>b</sup> (-1.773)	
			Vdw <sup>a</sup> (-5.698)	
			Coul <sup>b</sup> (-2.582)	
<b>L2</b>	-5.994	Glu166	Vdw <sup>a</sup> (-4.604)	1.723
			Coul <sup>b</sup> (-1.958)	
<b>L3</b>	-5.426	Gly143 Glu166	Vdw (-0.864)	1.880 2749
			Coul (-1.090)	
			Vdw (-2.610)	
			Coul (-3.403)	
<b>L5</b>	-6.617	-	-	-
<b>L6</b>	-6.020	-	-	-
<b>L7</b>	-6.456	Gly143 Glu166	Vdw (-1.540)	1.888 2.249
			Coul (-0.970)	
			Vdw (-3.308)	
			Coul (-3046)	
<b>L8</b>	-7.040	Cys145 Glu166	Vdw <sup>a</sup> (-3.072)	2.246 1.946
			Coul <sup>b</sup> (-1.362)	
			Vdw (-3.367)	
			Coul (+2.129)	
<b>L9</b>	-6.298	Cys145 Glu166	Vdw <sup>a</sup> (-3.093)	2.332 2.198
			Coul <sup>b</sup> (-0.698)	
			Vdw (-3862)	
			Coul (-2.820)	
<b>L10</b>	-5.596	Glu166	Vdw <sup>a</sup> (-5.635)	1.992
			Coul <sup>b</sup> (-3.340)	
<b>L11</b>	-6.988	Cys145 Glu166	Vdw <sup>a</sup> (-3.059)	2.172 1.940
			Coul <sup>b</sup> (-0.989)	
			Vdw (-3.699)	
			Coul (+0.863)	
<b>L12</b>	-8.030	Cys145 Glu166	Vdw <sup>a</sup> (-3.160)	2.257 2.234
			Coul <sup>b</sup> (-1.330)	
			Vdw (-5.264)	
			Coul (+1.903)	

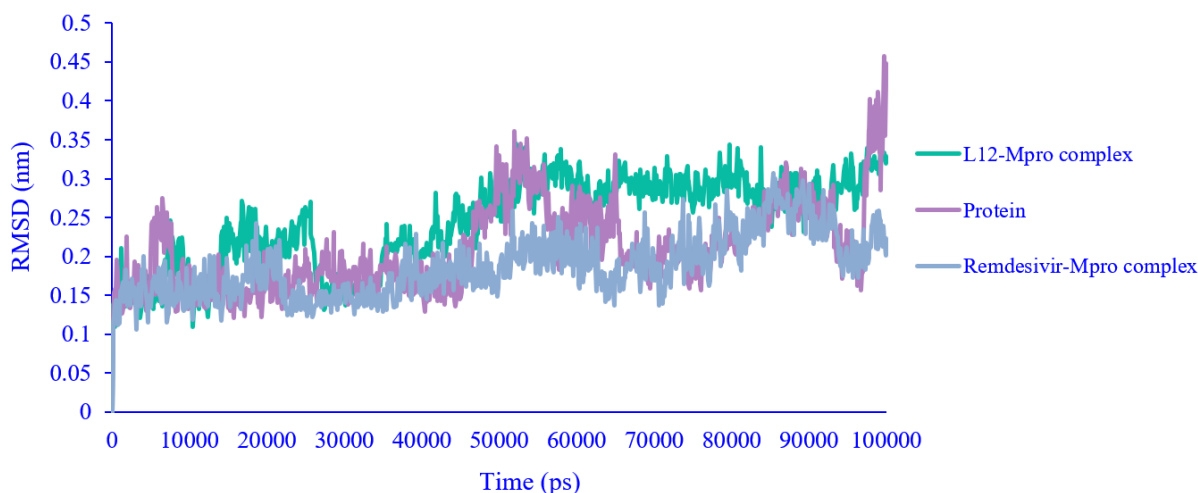
a, Van der Waals (a distance-dependent interaction between atoms or molecules) energy; b, Coulomb (the primary force determining the behavior of colliding atoms or molecules) energy.

**Table 4.** Binding energy results of all designed compounds using MM-GBSA calculations

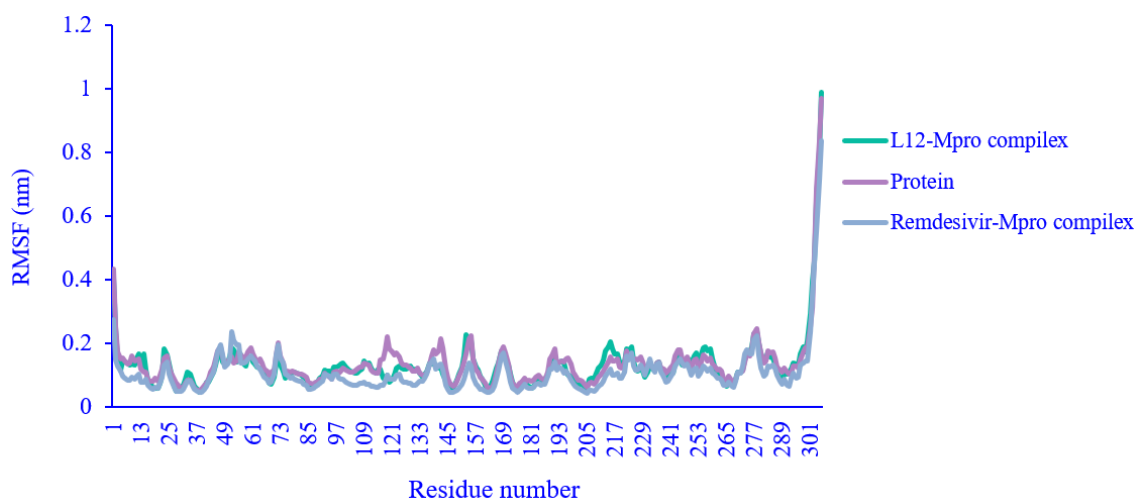
Entry	$\Delta G_{\text{Bind}}^a$	$\Delta G_{\text{Coul}}^b$	$\Delta G_{\text{Coval}}^c$	$\Delta G_{\text{Hbond}}^d$	$\Delta G_{\text{Lipo}}^e$	$\Delta G_{\text{SolvGB}}^f$	$\Delta G_{\text{vdw}}^g$
<b>Remdesivir</b>	-85.532	-18.808	3.163	-1.812	-38.710	25.818	-54.795
<b>L1</b>	-62.135	-9.352	5.190	-0.498	-37.075	28.069	-47.885
<b>L2</b>	-87.618	-17.805	3.452	-1.004	-42.048	25.458	-54.486
<b>L3</b>	-79.556	-14.475	6.842	-1.139	-37.473	21.204	-53.442
<b>L4</b>	-74.374	-15.258	8.002	-1.041	-43.612	27.100	-49.340
<b>L5</b>	-75.170	-9.808	5.141	-0.853	-36.074	19.903	-52.322
<b>L6</b>	-70.903	-6.600	16.955	-0.358	-44.574	21.627	-56.931
<b>L7</b>	-85.047	-25.867	11.698	-1.054	-44.654	32.111	-56.914
<b>L8</b>	-80.029	-10.499	0.360	-0.989	-43.306	31.059	-56.293
<b>L9</b>	-74.950	-0.804	7.314	-0.599	-36.899	15.122	-58.358
<b>L10</b>	-74.666	-13.821	12.100	-0.335	-44.000	33.102	-60.711
<b>L11</b>	-66.266	-5.695	3.317	-0.984	-37.128	30.325	-55.006
<b>L12</b>	-88.173	-11.474	1.833	-1.012	-42.594	23.876	-58.573

a, Binding free energy; b, Coulomb energy; c, covalent energy; d, hydrogen bonding energy; e, lipophilic energy; f, the generalized born electrostatic solvation energy; g, Van der Waals energy.





**Fig. 4.** RMSD for 3CL<sup>PRO</sup> backbone in the absence and presence of remdesivir and L12 during 100-ns molecular dynamics simulation. RMSD, Root mean square deviations.



**Fig. 5.** RMSF for 3CL<sup>PRO</sup> backbone in the absence and presence of remdesivir and L12 during 100-ns molecular dynamics simulation. RMSF, Root mean square fluctuation.

The RMSD values for the 3CL<sup>PRO</sup>-remdesivir complex were stable during the first 50 ns but 0.15 nm fluctuation was seen in the last 50 ns. The RMSD values of the protein backbone indicated that adding ligands led to more stability during MD simulation.

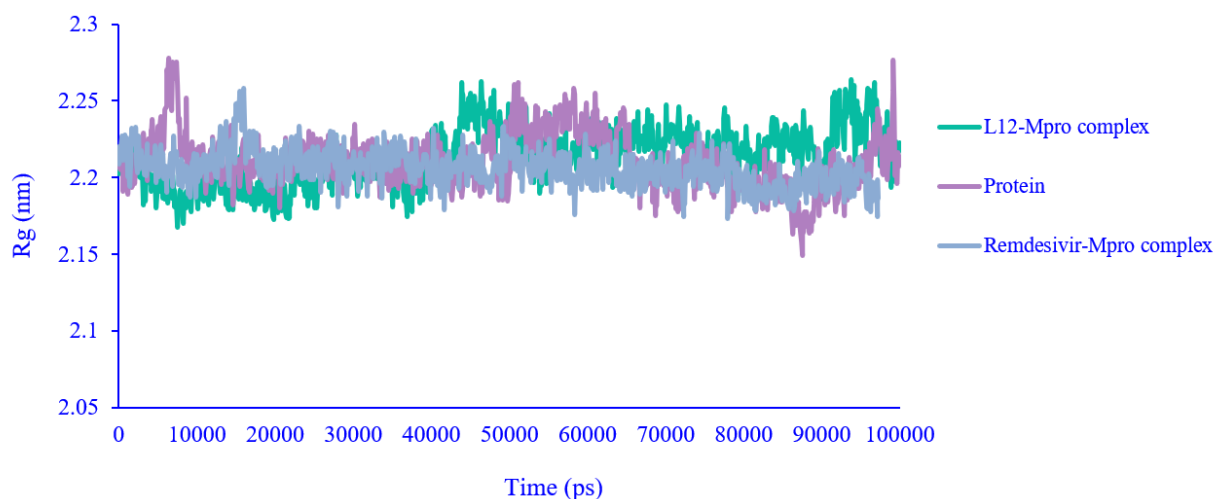
#### RMSF analysis

To further study the dynamic behavior or structural flexibility of the 3CL<sup>PRO</sup> structure after ligand binding, the difference in RMSF per residue was calculated for three systems (Fig. 5). Residues Gly143-Cys145, His163, His164, and His41 that bind with remdesivir showed a relatively small degree of flexibility, and residues positioned in the binding site seem to be more rigid as a result of binding to the ligand. Put differently, flexibility declined in the substrate-binding area, showing

the fact that the inhibitor slightly affects the residues situated at the substrate-binding pocket. Furthermore, we observed higher fluctuations in a few regions, including residues 44-49, 191-195, 215-222, and 275-279, but catalytic regions are shown more stable than other regions. These results showed that the flexibility of protein has been kept during MD simulation and the presence of ligand has caused stability in catalytic amino acids.

#### Rg analysis

The compaction of the three systems studied in the simulations was investigated by calculating the Rg value to determine changes in the enzyme because of ligand binding. This index represents the general dimension of protein during the simulation.



**Fig. 6.** Rg for 3CL<sup>pro</sup> backbone in the absence and presence of remdesivir and L12 during 100-ns molecular dynamics simulation. Rg, Radius of gyration.

Analyses have shown a minor decrease in the Rg value of 3CL<sup>pro</sup> through remdesivir and L12 binding, which refers to the minor unfolding of 3CL<sup>pro</sup> and a more compact structure following the ligand binding (Fig. 6). As shown by the molecular modeling outputs, ligands binding to 3CL<sup>pro</sup> largely happens into a certain active site and make specific changes in the protein micro-environment. Interaction with ligands led to the stability of 3CL<sup>pro</sup> conformation.

## DISCUSSION

In this study, remdesivir scaffold, as the first FDA-approved COVID-19 treatment, was chosen as the lead compound to predict the new M<sup>pro</sup> protein inhibitors. The critical residues of the active site of the M<sup>pro</sup> involved Cys145, Gly143, and Glu166, which create the hydrogen bond to the phosphoramidate and carbonyl groups in remdesivir. Thus, the phosphoramidate and carbonyl groups were kept, variable sections 1 and 2 changed (Fig. 1), and 12 new compounds were designed. Molecular docking results showed that, all 12 compounds made hydrogen bonds with catalytic amino acids, except for compounds L5 and L6. The compounds L12 and L8 showed the XP-GScore more than remdesivir and other compounds. The binding free energy was calculated for all 12 compounds and remdesivir as a reference compound. In compounds L1, L2, and L3, variable section 1 was changed to quinolone-4 amine, and variable section 2 was

changed to ethyl, propyl, and isopropyl groups, respectively. Among these compounds, L2 with the propyl group in variable section 2 showed the lowest binding free energy. In the second category of compounds, compounds L4-L12, in variable section 1, the propyl group was kept and the quinolone-4 amine part was replaced with N-Alkyl N-phenyl moiety. Compounds L4 (N-phenyl group), L5 (N-methyl N-phenyl group), L6 (N, N diphenyl group), and also the compounds with the electron-withdrawing substituent group on the phenyl ring (L9, L10, and L11) showed high binding free energy. Instead, the compounds with the electron-releasing substituent group on the phenyl ring demonstrated the lowest binding free energy (L8 and L12). Compound L12 with the lowest binding free energy was introduced to the MD simulation study.

In the molecular dynamic simulation, the investigation of RMSF results demonstrated that the 3CL<sup>pro</sup> backbone in the complex with L12 was more stable (with fluctuation lower than 0.1 nm) than the backbone in the complex with remdesivir and without ligand after 50 ns simulation. The RMSF results in L12 were the same as remdesivir, and the new ligand did not disrupt the protein backbone. Generally, the MD simulation results confirmed molecular docking and MM-GBSA results.

## CONCLUSION

Herein we have reported an *in-silico* study about the new compounds (12 compounds)

designed based on the remdesivir structure. The main interactions of remdesivir and designed compounds were investigated in the 3CL<sup>pro</sup> active site. Also, the pharmacokinetic and physicochemical properties of the compounds were calculated. The main interactions of all compounds were investigated through molecular docking then the binding free energy values between them were calculated by the MM-GBSA method and dynamic simulation was performed on the best combination among 12 compounds to further investigate the stability and interaction. Among the designed compounds, compounds 8 and 12 showed the highest binding affinity to M<sup>pro</sup> protein with XP GScore of -7.040 and -8.030 kcal/mol, respectively. Also, in both compounds, the main interactions were seen with Cys145 and Glu166 amino acids. The obtained binding free energy using the MM-GBSA method plays the main role in prioritizing the screened compounds. Compound L12 with -88.173 kcal/mol was chosen as the best compound and introduced to MD simulation. The simulation results of compound 12 were compared with the results of protein simulation without the presence of an inhibitor and in the presence of remdesivir. The Rg, RMSF, and RMSD results confirmed better stability of L12 compared with remdesivir. Supported by the theoretical results, compound 12 could have the potential to inhibit the 3CL<sup>pro</sup> enzyme, which requires further *in vitro* studies and enzyme inhibition must also be confirmed at protein levels.

#### Acknowledgments

The researcher(s) used the resources supported by Yasuj University of Medical Sciences, Yasuj, Iran and the Department of Pharmaceutical Chemistry, School of Pharmacy, Hormozgan University of Medical Sciences, Bandar Abbas, Iran to complete the computational work. We are grateful to Shahid Beheshti University of Medical Sciences, Tehran, Iran for financially supporting this research (Grants No. 24866. IR.SBMU.RETECH.REC.1399.642).

#### Conflict of interest statement

All authors declared no conflict of interest in this study.

#### Authors' contributions

A. Arefi Oskouie, M. Mansourian, S. Taheri, and M. Abbasi contributed to the study conception and design; M. Abbasi, A. Arefi Oskouie, and M. Mansourian collected the data; M. Abbasi, M. Mansourian, and K. Mahnam analyzed and interpreted the results; A. Arefi Oskouie, M. Abbasi, and M. Mansourian prepared the main draft of the manuscript. All authors reviewed the results and approved the final version of the manuscript.

#### REFERENCES

1. Kahn JS, McIntosh K. History and recent advances in coronavirus discovery. *Pediatr Infect Dis J*. 2005;24(11 Suppl):S223-S226. DOI: 10.1097/01.inf.0000188166.17324.60.
2. Jones DS. History in a crisis - lessons for Covid-19. *N Engl J Med*. 2020;382(18):1681-1683. DOI: 10.1056/NEJMp2004361.
3. World Health Organization, WHO Health Emergencies Programme, WHO COVID-19 dashboard. Available on: <https://covid19.who.int/>
4. Nadeem MS, Zamzami MA, Choudhry H, Murtaza BN, Kazmi I, Ahmad H, *et al*. Origin, potential therapeutic targets and treatment for coronavirus disease (COVID-19). *Pathogens*. 2020;9(4):307,1-13. DOI: 10.3390/pathogens9040307.
5. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. *J Virol*. 2020;94(7):e00127-20,1-25. DOI: 10.1128/JVI.00127-20.
6. Warren TK, Jordan R, Lo MK, Ray AS, Mackman RL, Soloveva V, *et al*. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature*. 2016;531(7594):381-385. DOI: 10.1038/nature17180.
7. Santos R, Ferreira AJ, Verano-Braga T, Bader M. Angiotensin-converting enzyme 2, angiotensin-(1-7) and Mas: new players of the renin-angiotensin system. *J Endocrinol*. 2013;216(2):R1-R17. DOI: 10.1530/joe-12-0341.
8. Zhang DH, Wu KL, Zhang X, Deng SQ, Peng B. *In silico* screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus. *J Integr Med*. 2020;18(2):152-158. DOI: 10.1016/j.joim.2020.02.005.
9. Freitas BT, Durie I A, Murray J, Longo J E, Miller HC, *et al*. Characterization and noncovalent inhibition of the deubiquitinase and deISGylase activity of SARS-CoV-2 papain-like protease. *ACS Infect Dis*. 2020;6(8):2099-2109. DOI: 10.1021/acsinfecdis.0c00168.
10. Naik VR, Munikumar M, Ramakrishna U, Srujana M, Goudar G, Naresh P, *et al*. Remdesivir (GS-5734) as a therapeutic option of 2019-nCoV main protease-*in*

- silico* approach. *J Biomol Struct Dyn.* 2021;39(13):4701-4714.  
DOI: 10.1080/07391102.2020.1781694.
11. Liu C, Zhou Q, Li Y, Garner LV, Watkins SP, Carter LJ, et al. Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Cent Sci.* 2020;6(3):315-331.  
DOI: 10.1021/acscentsci.0c00272.
  12. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerhering L, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved  $\alpha$ -ketoamide inhibitors. *Science.* 2020;368(6489):409-412.  
DOI: 10.1126/science.abb3405.
  13. Koulgi S, Jani V, Uppuladinne M, Sonavane U, Nath AK, Darbari H, et al. Drug repurposing studies targeting SARS-CoV-2: an ensemble docking approach on drug target 3C-like protease (3CL<sup>pro</sup>). *J Biomol Struct Dyn.* 2021;39(15):5735-5755.  
DOI: 10.1080/07391102.2020.1792344.
  14. Nguyen HL, Thai NQ, Truong DT, Li MS. Remdesivir strongly binds to both RNA-dependent RNA polymerase and main protease of SARS-CoV-2: evidence from molecular simulations. *J Phys Chem B.* 2020;124(50):11337-11348.  
DOI: 10.1021/acs.jpcc.0c07312.
  15. 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):289-293.  
DOI: 10.1038/s41586-020-2223-y.
  16. Xue X, Yu H, Yang H, Xue F, Wu Z, Shen W, et al. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J Virol.* 2008;82(5):2515-2527.  
DOI: 10.1128/JVI.02114-07.
  17. Wang F, Chen C, Tan W, Yang K, Yang H. Structure of main protease from human coronavirus NL63: insights for wide spectrum anti-coronavirus drug design. *Sci Rep.* 2016;6(1):22677,1-12.  
DOI: 10.1038/srep22677.
  18. Vardhan S, Sahoo SK. *In silico* ADMET and molecular docking study on searching potential inhibitors from limonoids and triterpenoids for COVID-19. *Comput Biol Med.* 2020;124:103936,1-12.  
DOI: 10.1016/j.combiomed.2020.103936.
  19. Muralidharan N, Sakthivel R, Velmurugan D, Gromiha MM. Computational studies of drug repurposing and synergism of lopinavir, oseltamivir and ritonavir binding with SARS-CoV-2 protease against COVID-19. *J Biomol Struct Dyn.* 2021;39(7):2673-2678.  
DOI: 10.1080/07391102.2020.1752802.
  20. Cao B, Wang Y, Wen D, Liu W, Wang J, Fan G, et al. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. *N Engl J Med.* 2020;382(19):1787-1799.  
DOI: 10.1056/NEJMoa2001282.
  21. Stower H. Lopinavir-ritonavir in severe COVID-19. *Nat Med.* 2020;26(4):465,1-2.  
DOI: 10.1038/s41591-020-0849-9.
  22. Akhtar MJ. COVID19 inhibitors: a prospective therapeutics. *Bioorg Chem.* 2020;101:104027,1-26.  
DOI: 10.1016/j.bioorg.2020.104027.
  23. Siordia JA, Bernaba M, Yoshino K, Ulhaque A, Kumar S, Bernaba M, et al. Systematic and statistical review of COVID19 treatment trials. *SN Compr Clin Med.* 2020;2(8):1120-1131.  
DOI: 10.1007/s42399-020-00399-6.
  24. Khalili JS, Zhu H, Mak NSA, Yan Y, Zhu Y. Novel coronavirus treatment with ribavirin: Groundwork for an evaluation concerning COVID-19. *J Med Virol.,* 2020 ;92(7):740-746.  
DOI: 10.1002/jmv.25798.
  25. Touret F, de Lamballerie X. Of chloroquine and COVID-19. *Antiviral Res.* 2020;177:104762,1-2.  
DOI: 10.1016/j.antiviral.2020.104762.
  26. Pagliano P, Piazza O, De Caro F, Ascione T, Filippelli A. Is hydroxychloroquine a possible post-exposure prophylaxis drug to limit the transmission to health care workers exposed to COVID19. *Clin Infect Dis.* 2020;71(15):887-888.  
DOI: 10.1093/cid/ciaa320.
  27. Uno Y. Camostat mesilate therapy for COVID-19. *Intern Emerg Med.* 2020;15(8):1577-1578.  
DOI: 10.1007/s11739-020-02345-9.
  28. Horby P, Lim W, Emberson J, Mafham M, Bell J, Linsell L, et al. Effect of dexamethasone in hospitalized patients with COVID-19. *N Engl J Med.* 2021;384(8):693-704.  
DOI: 10.1056/NEJMoa2021436.
  29. Wang Y, Anirudhan V, Du R, Cui Q, Rong L. RNA-dependent RNA polymerase of SARS-CoV-2 as a therapeutic target. *J Med Virol.* 2021;93(1):300-310.  
DOI: 10.1002/jmv.26264.
  30. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) *in vitro*. *Cell Res.* 2020;30(3):269-271.  
DOI: 10.1038/s41422-020-0282-0.
  31. FDA Approves First Treatment for COVID-19. Available on: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19>. 2020.
  32. Consortium WST. Repurposed antiviral drugs for COVID-19-interim WHO solidarity trial results. *N Engl J Med.* 2021;384(6):497-511.  
DOI: 10.1056/NEJMoa2023184.
  33. Kabinger F, Stiller C, Schmitzová J, Dienemann C, Kokic G, Hillen HS, et al. Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nat Struct Mol Biol.* 2021;28(9):740-746.  
DOI: 10.1038/s41594-021-00651-0.
  34. Khan SA, Zia K, Ashraf S, Uddin R, Ul-Haq Z. Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 *via* integrated computational approach. *J Biomol Struct Dyn.* 2021;39(7):2607-2616.  
DOI: 10.1080/07391102.2020.1751298.
  35. Mishra A, Rathore AS. RNA dependent RNA polymerase (RdRp) as a drug target for SARS-CoV2. *J Biomol Struct Dyn.* 2022;40(13):6039-6051.  
DOI: 10.1080/07391102.2021.1875886.

36. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, *et al.* Structure of M<sup>pro</sup> from SARS-CoV-2 and discovery of its inhibitors. *Nature*. 2020;582(7811):289-293. DOI: 10.1038/s41586-020-2223-y.
37. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013;27(3):221-234. DOI: 10.1007/s10822-013-9644-8.
38. Schrödinger Release 2023-2: LigPrep, Schrödinger, LLC, New York, NY, 2021.
39. Kramer C, Ting A, Zheng H, Hert J, Schindler T, Stahl M, *et al.* Learning medicinal chemistry absorption, distribution, metabolism, excretion, and toxicity (ADMET) rules from cross-company matched molecular pairs analysis (MMPA). *J Med Chem*. 2018;61(8):3277-3292. DOI: 10.1021/acs.jmedchem.7b00935.
40. Lagorce D, Douguet D, Miteva M, Villoutreix B. Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. *Sci Rep*. 2017;7:46277,1-15. DOI: 10.1038/srep46277.
41. Schrodinger L. Glide, Version 6.6. New York (NY). 2015.
42. Abbasi M, Sadeghi-Aliabadi H. An *in-silico* screening strategy to the prediction of new inhibitors of COVID-19 M<sup>pro</sup> protein. *Iran J Pharm Res*. 2021;20(4):125-136. DOI: 10.22037/ijpr.2021.114997.15146.
43. Lokhande KB, Doiphode S, Vyas R, Swamy KV. Molecular docking and simulation studies on SARS-CoV-2 M<sup>pro</sup> reveals mitoxantrone, leucovorin, birinapant, and dynasore as potent drugs against COVID-19. *J Biomol Struct Dyn*. 2021;39(18):7294-7305. DOI: 10.1080/07391102.2020.1805019.
44. Pearlman DA, Charifson PS. Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. *J Med Chem*. 2001;44(21):3417-3423. DOI: 10.1021/jm0100279.
45. Huang N, Kalyanaraman C, Irwin JJ, Jacobson MP. Physics-based scoring of protein-ligand complexes: Enrichment of known inhibitors in large-scale virtual screening. *J Chem Inf Model*. 2006;46(1):243-253. DOI: 10.1021/ci0502855.
46. Mulakala C, Viswanadhan V. Could MM-GBSA be accurate enough for calculation of absolute protein/ligand binding free energies? *J Mol Graph Model*. 2013;46:41-51. DOI: 10.1016/j.jmgm.2013.09.005.
47. Mansourian M, Mahnam K, Rajabi HR, Roushani M, Doustimotlagh AH. Exploring the binding mechanism of saccharin and sodium saccharin to promoter of human p53 gene by theoretical and experimental methods. *J Biomol Struct Dyn*. 2020;38(2):548-564. DOI: 10.1080/07391102.2019.1582438.
48. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 2008;4(3):435-447. DOI: 10.1021/ct700301q.
49. Nasab RR, Mansourian M, Hassanzadeh F, Shahlaei M. Exploring the interaction between epidermal growth factor receptor tyrosine kinase and some of the synthesized inhibitors using combination of *in-silico* and *in-vitro* cytotoxicity methods. *Res Pharm Sci*. 2018;13(6):509-522. DOI: 10.4103/1735-5362.245963.
50. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1:19-25. DOI: 10.1016/j.softx.2015.06.001.
51. da Silva AWS, Vranken WF. ACPYPE-antechamber python parser interface. *BMC Res Notes*. 2012;5:367,1-8. DOI: 10.1186/1756-0500-5-367.
52. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983;79(2):926-935. DOI: 10.1063/1.445869.
53. Hess B, Bekker H, Berendsen HJ, Fraaije JG. LINCS: a linear constraint solver for molecular simulations. *J Comput Chem*. 1997;18(12):1463-72. DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.
54. Darden T, York D, Pedersen L. Particle mesh Ewald: an N-log (N) method for Ewald sums in large systems. *J Chem Phys*. 1993;98(12):10089-10092. DOI: 10.1063/1.464397.
55. Berendsen HJ, Postma JV, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81(8):3684-3690. DOI: 10.1063/1.448118.
56. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38. DOI: 10.1016/0263-7855(96)00018-5.
57. Makarewicz T, Kazmierkiewicz R. Molecular dynamics simulation by GROMACS using GUI plugin for PyMOL. *J Chem Inf Model*. 2013;53(5):1229-1234. DOI: 10.1021/ci400071x.
58. Sirous H, Chemi G, Campiani G, Brogi S. An integrated *in silico* screening strategy for identifying promising disruptors of p53-MDM2 interaction. *Comput Biol Chem*. 2019;83:107105,1-61. DOI: 10.1016/j.compbiolchem.2019.107105.
59. Razzaghi-Asl N, Mirzayi S, Mahnam K, Adhami V, Sepehri S. *In silico* screening and molecular dynamics simulations toward new human papillomavirus 16 type inhibitors. *Res Pharm Sci*. 2022;17(2):189-208. DOI: 10.4103/1735-5362.335177.
60. Yazdani M, Khezri J, Hadizadeh N, Zamani Amir Zakaria J, Naderi M, Mahmoodian S, *et al.* Depinar, a drug that potentially inhibits the binding and entry of COVID-19 into host cells based on computer-aided studies. *Res Pharm Sci*. 2021;16(3):315-325. DOI: 10.4103/1735-5362.314830.