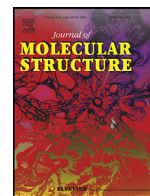




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# The identification of novel inhibitors of human angiotensin-converting enzyme 2 and main protease of Sars-Cov-2: A combination of *in silico* methods for treatment of COVID-19



Vahid Zarezade<sup>a,b,\*</sup>, Hamzeh Rezaei<sup>c</sup>, Ghodratollah Shakerinezhad<sup>d</sup>, Arman Safavi<sup>e</sup>, Zahra Nazeri<sup>b</sup>, Ali Veisi<sup>a</sup>, Omid Azadbakht<sup>a</sup>, Mahdi Hatami<sup>b</sup>, Mohamad Sabaghan<sup>a</sup>, Zeinab Shajirat<sup>a</sup>

<sup>a</sup> Behbahan Faculty of Medical Sciences, Behbahan, Iran

<sup>b</sup> Department of Clinical Biochemistry, School of Medicine, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

<sup>c</sup> Department of Clinical Biochemistry, School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran

<sup>d</sup> Health Education Research Department, ACECR\_Khuzestan, Ahvaz, Iran

<sup>e</sup> Department of Cell and Molecular Biology and Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran

## ARTICLE INFO

### Article history:

Received 26 January 2021

Revised 29 March 2021

Accepted 31 March 2021

Available online 6 April 2021

### Keywords:

COVID-19

SARS-CoV-2

Main protease

ACE2

Molecular dynamics simulation

## ABSTRACT

The angiotensin-converting enzyme 2 (ACE2) and main protease (MPro), are the putative drug candidates for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In this study, we performed 3D-QSAR pharmacophore modeling and screened 1,264,479 ligands gathered from Pubchem and Zinc databases. Following the calculation of the ADMET properties, molecular docking was carried out. Moreover, the *de novo* ligand design was performed. MD simulation was then applied to survey the behavior of the ligand-protein complexes. Furthermore, MMPBSA has utilized to re-estimate the binding affinities. Then, a free energy landscape was used to find the most stable conformation of the complexes. Finally, the hybrid QM-MM method was carried out for the precise calculation of the energies.

The Hypo1 pharmacophore model was selected as the best model. Our docking results indicate that the compounds ZINC12562757 and 112,260,215 were the best potential inhibitors of the ACE2 and MPro, respectively. Furthermore, the Evo\_1 compound enjoys the highest docking energy among the designed *de novo* ligands. Results of RMSD, RMSF, H-bond, and DSSP analyses have demonstrated that the lead compounds preserve the stability of the complexes' conformation during the MD simulation. MMPBSA data confirmed the molecular docking results. The results of QM-MM showed that Evo\_1 has a stronger potential for specific inhibition of MPro, as compared to the 112,260,215 compound.

© 2021 Elsevier B.V. All rights reserved.

In conclusion, our results showed that the *de novo* designed Evo\_1 compound has the potential to be used as a drug for the treatment of COVID-19; however, further *in vitro* and *in vivo* validations are required.

## 1. Introduction

In late December 2019, a new coronavirus (CoV) called coronavirus 2019 (2019-nCoV) or severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), began to spread pneumonia from Wuhan to all over China, and then formed a large global pandemic,

reaching almost worldwide by July 7, 2020. According to global statistics, which are being updated instantly, the virus death rate is 3.4% [1]. Early symptoms of coronavirus disease 2019 (COVID-19) include pneumonia, fever, muscle aches, and fatigue, and no specific or effective drug has been introduced to date. However, many researchers around the world are in the process of designing a vaccine or drug for this devastating disease [2].

The 2019-nCoV is a non-segmented, enveloped virus with a positive-sense single-stranded RNA of animal origin, belonging to the family of *Coronaviridae*, and the order of *Nidovirales*. Its genomic size is 26–36 kbps and it is one of the largest RNA viruses. The 2019-nCoV is a beta-CoV, similar to the virus causing severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome (MERS) [3].

\* Corresponding authors at: Deputy of Education and Research, Behbahan Faculty of Medical Sciences, Shahid Zibaei street, Behbahan, Khuzestan province, Iran.

E-mail address: [v.zarezade@behums.ac.ir](mailto:v.zarezade@behums.ac.ir) (V. Zarezade).

The angiotensin-converting enzyme 2 (ACE2) is a carboxypeptidase and a type I integral membrane protein of 805 amino acids, which converts angiotensin 1 to angiotensin 1 – 9 or angiotensin 2 to angiotensin 1–7 [4]. ACE2 functions as a counter-regulator of the renin-angiotensin system, which is a key regulator of cardiovascular homeostasis [5]. Moreover, it has been reported that it plays the role of a receptor for the 2019-nCoV, just like its function as the SARS-CoV receptor, except that it has approximately 20 times stronger binding affinity than that of the SARS [6, 7].

Spike glycoprotein responsible for CoV crown-like presentation, by creating spikes on its surface, supports the passage of 2019-nCoV into the host cells. The 2019-nCoV binds to the ACE2 catalytic domain through the SARS-CoV Spike Protein 1, which is accompanied by conformational changes in this protein, increasing its sensitivity to proteolytic digestion at the boundary between S1 and S2 subunits of spike glycoprotein [8]. This proteolysis is believed to be achieved in two ways, one through acid-dependent proteolysis, by cathepsin L, and the other through transmembrane protease serine 2 (TMPRSS2), and human airway trypsin-like protease (HAT) [9]. Two mechanisms for virus entry into the cell have been suggested as follows: In the first mechanism, the cleavage of ACE2 by disintegrin and metallopeptidase domain 17 (ADAM17)/tumor necrosis factor  $\alpha$ -converting enzyme (TACE), increases the release of ACE2 into the extracellular space, enhances and facilitates the virus entry into the cell [10]. The second proposed mechanism is TMPRSS2/HAT [11], which is mainly involved in virus entry into the cell rather than in the activation of spike protein 1 [10]. ACE2 chooses whether an animal or a cell can be infected and it can be used for beneficial treatment to administer complications caused by coronaviruses [7, 12].

Besides the ACE2, spike protein S1, and TMPRSS2 proteins, which are all known as viral treatment targets, the main protease (MPro) is also one of the most important drug targets for CoVs. Non-structural polyproteins catalyzed by MPro and some papain-like proteases, ultimately produce RNA-dependent RNA polymerase (RdRp) and virus-encoded RNA helicase and increase viral replication [13]. MPro is an attractive goal that can be used as antiviral drugs due to the practical significance of MPro in the viral life cycle, and because it does not have similar homologs in humans [14, 15]. SARS-CoV-2 MPro has a Cys-His catalytic dyad, and a substrate-binding site, which is situated in a cleft between domain I and domain II [14]. Thus far, various inhibitors have been introduced to control SARS, MERS, and other viruses, but they have not been effective or sufficiently specific for COVID-19.

Since the process of designing new drugs is very time-consuming, especially in the case of pandemic conditions of a relatively unknown disease, possible treatment options are particularly important. Therefore, In the event of an emergency, the first line of the fight against COVID-19 is the screening of existing drugs and examining their impact on the treatment process of the disease [2]. Recent drug screenings, which focus on the MPro as the target, have shown that nelfinavir, pitavastatin, perampanel, and praziquantel had relative antiviral activity against the 2019-nCoV [16]. Other antivirals, especially those approved by the FDA, including penciclovir, nitrazine, nalfamusta, and chloroquine, are also other treatment options for COVID-19 [17, 18].

Overall, the treatment of patients with COVID-19 infection is generally mostly symptomatic and supportive therapy and based on the previous experience of SARS and MERS. There have been limited studies in the field of potential drug design for COVID-19, in which various aspects of this subject have not been considered [19–21].

Since in these pandemic circumstances, designing potential drugs with high specificity and high potency is extremely vital and desired by the international community, and considering the importance of ACE2 for the virus entry and emphasizing on the cru-

cial role of MPro in virus activity, therefore in this study, we aimed to screen a vast database and recruit the powerful drug design techniques in order to find a medication that is effective for inhibiting human ACE2 and MPro, using computer-aided drug design techniques.

## 2. Materials and methods

### 2.1. Selection of the compounds

The first step in developing a robust pharmacophore model is to select known, biologically active inhibitors that give the model the screening power of numerous compounds from different databases. In this study, a dataset was created, including a training set and test set, composing 31 compounds found in the literature [22–24]. To arrange a training set, certain guidelines must be followed, as it is recruited to develop the pharmacophore model. The most active compounds should be included in the training package, as it is utilized to offer the inhibitory features. In addition, a potent training set should include at least 16 structurally distinct compounds [25], demonstrating a spectrum of 4 – 5 orders, and conveying structural and functional characterization of the small compounds. In this study, the training set is composed of 18 compounds (Fig. 1), which are precisely arranged and categorized into groups, showing distinct structures and activities, ranging from 0.013 to 920 nmol/L. Compounds with an activity range of less than 1 nmol/L, between 1 and 300 nmol/L, and greater than 300 nmol/L are considered as highly active, moderately active, and inactive compounds, respectively. Accordingly, the bindingdb database [26] was used to retrieve the 3D structures, which was prepared, using Discovery Studio (DS) 2016.

### 2.2. 3D QSAR pharmacophore model generation

The DS *Feature Mapping* protocol was started to critically examine the main chemical features of the training set compounds for producing the most potent pharmacophore models. The above-mentioned information was used for the generation of the pharmacophore. Using the *Catalyst HypoGen* algorithm, the *3D QSAR Pharmacophore Generation* module implemented in the DS was recruited to generate the pharmacophore. In addition, in order to create the conformations with the best coverage of conformational space at an uncertainty value of 3, with an inter-feature distance of 2.97 at 95% confidence level, the *Best* algorithm under the *Conformation Generation* section was appointed. Moreover, Hydrogen Bond Acceptor (HBA), Hydrogen Bond Donor (HBD), Hydrophobic (HYA), Negative Ionizable (NGI), and Ring Aromatic (RAR) features were checked with a minimum of 0 and maximum of 5 features, and remaining parameters were set as default. According to Debnath's analysis, an ideal pharmacophore should essentially show a high correlation coefficient, the least cost value, and the lowest Root Mean Square Deviation (RMSD). Accordingly, based on the Debnath analysis [27], the best pharmacophore from the produced pharmacophores was selected.

### 2.3. Validation of the pharmacophore

A valid 3D QSAR pharmacophore hypothesis must be able to predict the bioactivity of the training set inhibitors, in the same pattern as the actual experimental values. In addition, it is expected to behave in the same way for test set inhibitors, so it projects a similar biological effect as the training set. Moreover, the chance factor should not play a part in creating the hypothesis [28]. Cost analysis, Fischer's randomization, and test set prediction evaluation methods were therefore used to assess and validate the top ten hypotheses.

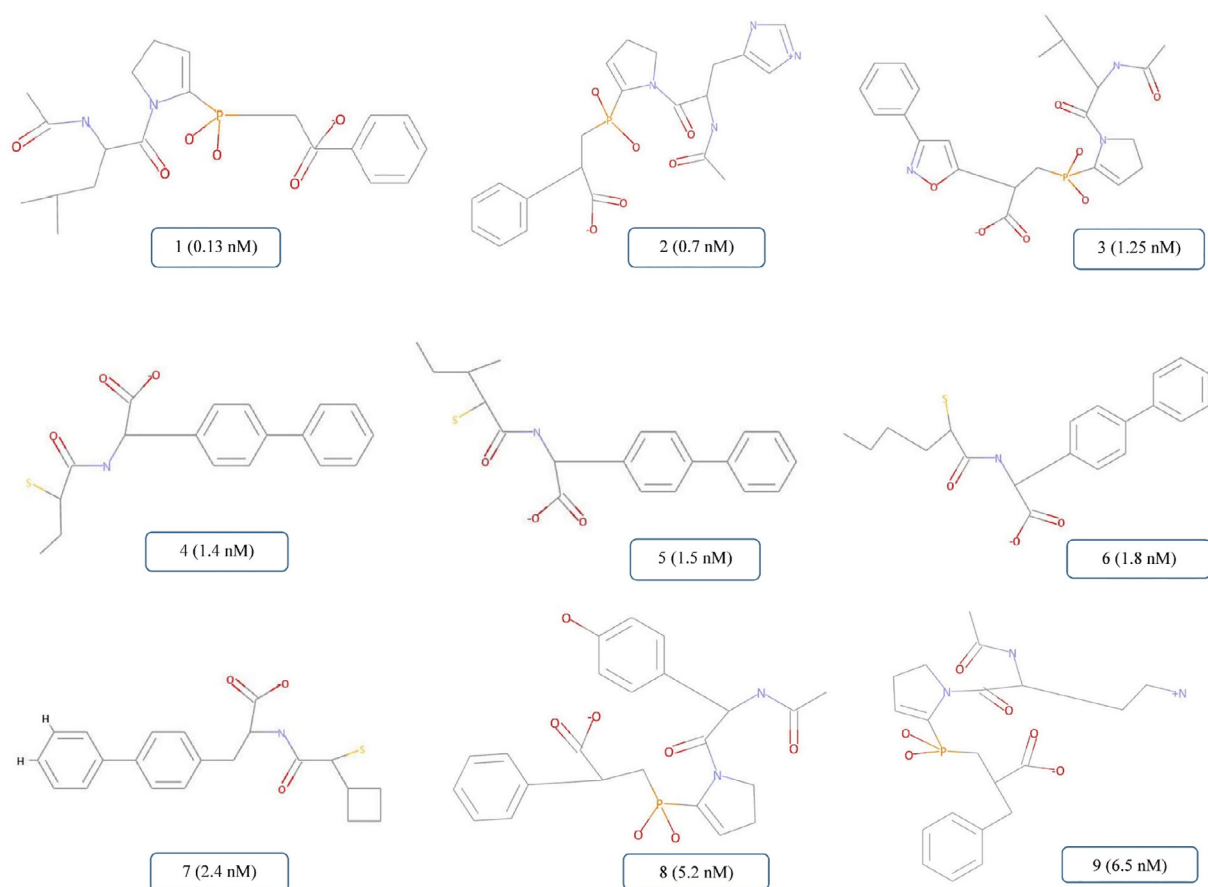


Fig. 1. The 2D structures of ACE2 inhibitors in the training set together with their biological activity data (Ki value, nM).

As a determining rule, the choice of the best model in cost analysis is based on the difference between null cost and total cost. The fixed cost implies that the best model makes a perfect representation of all data, whilst the null cost indicates that the worst model suits no feature. If the distance is more than 60 bits, the model is excellent in fitting all the data. A 40–60 bits' distance that infers the model 75–90%, is likely to demonstrate a real correlation in the data. If the distance is less than 40, then the model does not match all the data [29].

From the chosen hypothesis, the statistical significance was determined at a 95% confidence level, using Fischer's randomization method. This technique was employed to further assure that the pharmacophore is not randomly generated [30]. In order to assess the aptitude of the pharmacophore model, the test set was used to determine the compounds, with the same extent of experimental function, other than the training set. The test set validation was carried out, using the *Ligand Pharmacophore Mapping* protocol. In addition, based on the cost analysis, the optimal pharmacophore model should harbor the highest cost difference ( $\Delta\text{cost} = \text{null cost} - \text{total cost}$ ), and a high correlation coefficient for test and training sets, matching crucial pharmacophore features.

#### 2.4. Virtual screening and drug-likeness evaluation

Among the most refined methods configured in contemporary drug discovery, the virtual small molecule database screening is used to choose the possible drug leads for the related diseases. In this study, pharmacophore-based virtual screening was carried out, given that the verified model Hypo1 possesses the 3D features. About 1264,479 compounds, collected from Pubchem [31] and ZINC [32] databases, were used for screening, and further

retrieval of the candidate molecules. After the preparation of the ligands, the *Ligand Pharmacophore Mapping* package in the DS was used, selecting the *rigid* fitting method, and using the other options as default. Each compound that successfully matched the chosen pharmacophore model was considered as a potential drug, and the other compounds were excluded. In order to meet the drug-like criteria, the filtered ligands were submitted to ADMET and Lipinski's rule of five (RO5) [33] characteristics. The FAF-Drugs4 web server [34] was used to calculate the ADMET properties. Based on RO5, a well-absorbed compound possesses less than 5 hydrogen bonds and logP values of fewer than 5. The potential drug candidate must have a molecular weight of below 500 Da, and hold lower than 10 rotatable bonds and hydrogen bond acceptors. Furthermore, the ligands with a fit value of less than 13 were excluded. Compounds, meeting the criteria were then subjected to molecular docking calculations.

#### 2.5. Molecular docking studies

For effective assessment and identification of the lead candidates as potential drugs, molecular docking has been applied. Furthermore, this technique may be used to derive a detailed set of information on intramolecular events, such as interacting residues of the target active site, interaction types, the molecular pose of the hits at the active site, the energy level of interaction, etc. In this study, the CDOCKER protocol accessible in the DS was used and implemented concurrently, along with the CHARMm force field. The binding between the target and ligand is more favorable when CDOCKER energies are higher [28, 35]. In this study, the inhibitor-bound human ACE2 (PDB ID: 1R4L) is the first target with a resolution of 3 Å. A region of 20 Å, covering the crystalized inhibitor

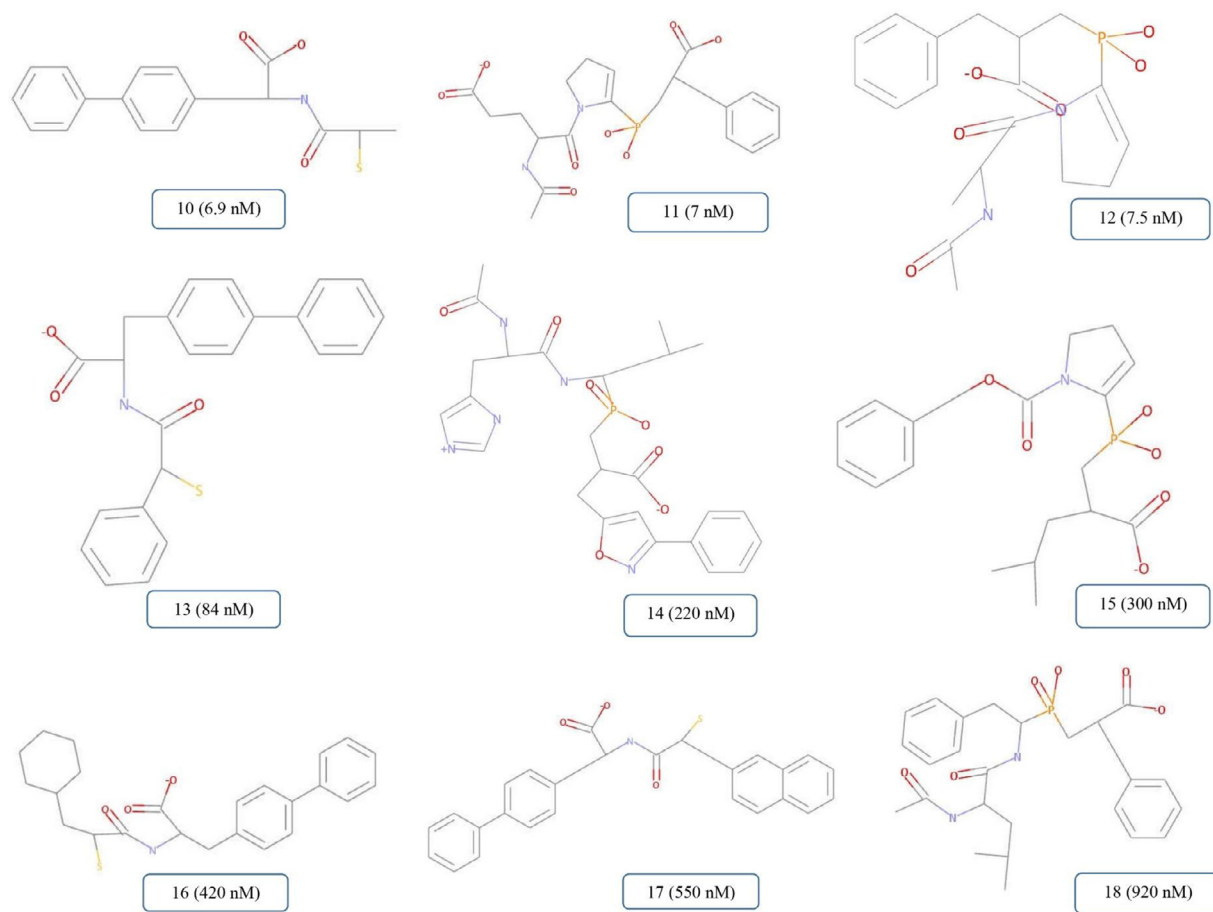


Fig. 1. Continued

was identified as the active site. To prepare the targets, the *Prepare Protein* module implemented in DS was recruited; except that the structural  $\text{Zn}^{2+}$  located at the target active site, all water molecules and unnecessary hetero atoms were discarded. The lead candidates in conjunction with the reference compound (the compound with the lowest  $K_i$  from the training set) were submitted to the docking protocol. Subsequently, the compounds obtained from the former step, together with the MPro reference inhibitors (Cinanserin and Ebselen) [14] were docked into the active site of the crystal structure of MPro from COVID-19 virus, in complex with an inhibitor N3 (PDB ID: 6LU7), as the second target. In addition, according to the interactions between the key residues and the ligand, and the highest dock score, the best-docked poses were assessed. To figure out the stability of the docking output complexes and to determine the reliability of the acquired data, the selected complexes were subjected to molecular dynamic simulations, using GROMACS v2018.4 [36].

## 2.6. De novo evolution protocol

One of the most effective approaches in developing higher efficacy drugs for the treatment of diseases is to strengthen the inhibitory activity of ligands, using the fragment-based design method. The basis of this method is to examine a library of fragments that can complement the active site of the target, and then manipulate the backbone of the selected inhibitors to create an improved structure with new substituents. The *Ludi* algorithm implemented in the DS *De Novo Evolution* module is recruited for this purpose [35]. In this protocol, the full evolution selecting fragments classified by the *Ludi* energy estimate score was used, and

the maximum RMSD for fitting was 0.3. Based on the highest *Ludi* scores from the *de novo* evolution, the new compounds were further docked back to the MPro active site for assessing the binding poses and interaction energies. Moreover, the pharmacokinetic and pharmacodynamics properties of the compounds were assessed, using the FAF-Drugs4 web server [34].

## 2.7. Molecular dynamics (MD) simulation studies

In order to assess the conformational alterations and the stability of the complexes and to verify the accuracy of the docking, the molecular dynamics (MD) simulation was carried out, concerning the top docking poses as initial inputs. GROMACS v2018.4 [36] was used for MD simulation, adapting the AMBER03 force field. The ligands were parametrized, using an antechamber package [37] of the Amber program, utilizing the GAFF force field [38]. The AM1-BCC charge model was applied to assign the partial charges of the ligands [39]. A cubic box was created and solvated with the TIP3P water molecules. Then, the opposite ions were included to counterbalance the system. Periodic boundary conditions (PBC) were applied in all three coordinate axes. Energy minimization of the system was conducted by recruiting the steepest descent algorithm and further equilibrated, using NVT and NPT, respectively. By applying a coupling constant of 0.1 for 100 ps, at a constant temperature of 300 K, the initial equilibration was carried out at constant steady volume (NVT), using the Berendsen thermostat algorithm. Afterward, the secondary equilibration was conducted with a coupling constant of 5.0 ps for 1 ns at constant pressure (NPT) of 1 bar, retained by Parrinello-Rahman barostat [40]. The hydrogen bonds involving atoms and the molecular geometry of water



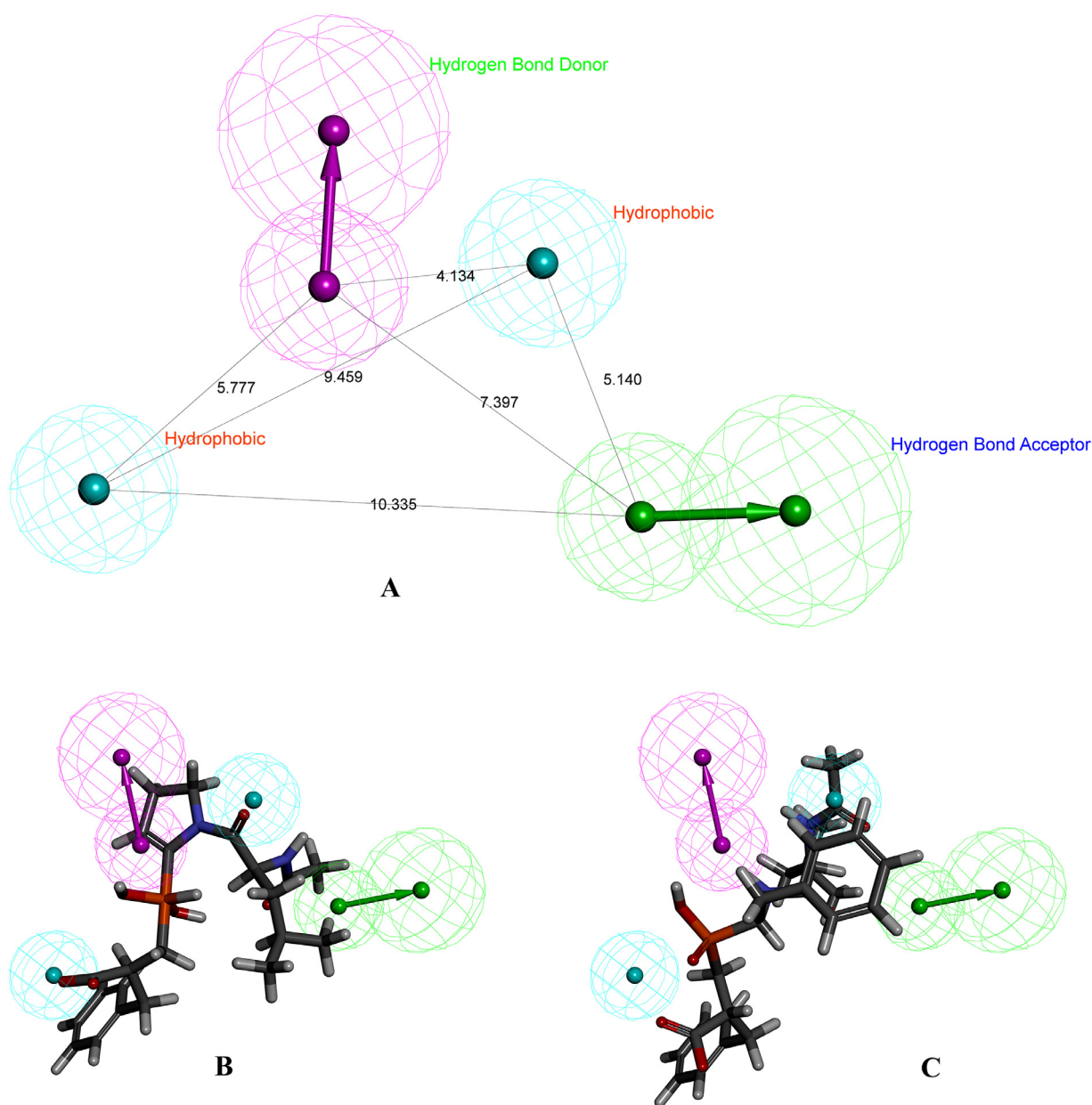
**Table 1**  
Ten pharmacophore models generated by the HypoGen for ACE2 inhibitors.

Hypothesis	Total Cost	Cost difference <sup>a</sup>	RMSD <sup>b</sup>	Correlation(R2)	Features <sup>c</sup>	Maximum Fit
1	85.2263	320.4527	1.16399	0.9839	HBA,HBD,HYA,HYA	13.9322
2	121.457	284.222	2.18194	0.9409	HBD,HBD,HYA	11.3805
3	145.216	260.463	2.91149	0.8894	HBA,HBD,HYA,HYA	12.8171
4	146.651	259.028	3.03014	0.879	HBA,HBD,HYA,HYA,NGI	13.7241
5	162.823	242.856	3.29537	0.8551	HBA,HBD,HBD,HYA	11.4649
6	163.411	242.268	3.23884	0.8609	HBA,HBA,HYA,RAR	12.8398
7	165.087	240.592	3.33661	0.8511	HBA,HBD,HBD,HYA	11.3807
8	170.229	235.45	3.40528	0.8445	HBA,HBD,HYA,HYA	11.778
9	174.893	230.786	3.51799	0.8327	HBA,HBA,HBD,HYA	10.7187
10	177.423	228.256	3.43974	0.8417	HBA,HBD,HBD,HYA	13.3921

<sup>a</sup> Cost difference between the null and the total cost. The null cost, the fixed cost, and the configuration cost are 405.679, 60.9745, and 16.2464, respectively. All costs are in units of bits.

<sup>b</sup> RMSD: The deviation of the log (estimated activities) from the log (measured activities) normalized by the log (uncertainties).

<sup>c</sup> Abbreviation used for features: HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; HYA, hydrophobic; NGI, Negative Ionizable; RAR, Ring Aromatic.



**Fig. 2.** (A) Best HypoGen pharmacophore model Hypo1 chemical features (B) Hypo1 mapping with the most active ACE2 inhibitor ( $K_i = 0.13$  nM) (C) Hypo1 mapping with the most inactive ACE2 inhibitor ( $K_i = 920$  nM).

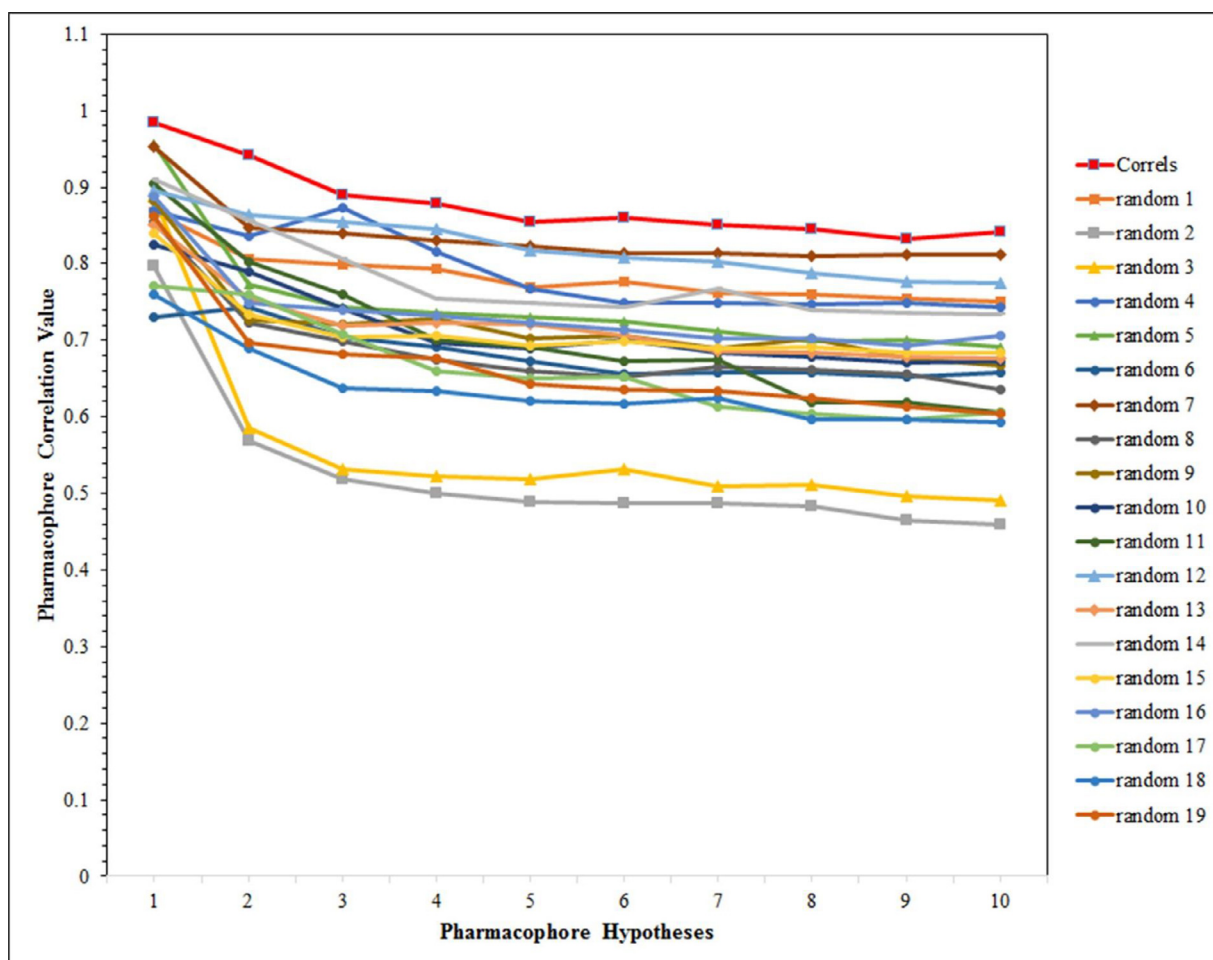


Fig. 3. The difference in correlation values of hypotheses between a Hypo1 spreadsheet and 19 random spreadsheets on the 95% confidence level.

were restrained, using SETTLE [41] and LINCS [42]. Particle Mesh Ewald (PME) [43] was used to estimate the electrostatic interactions in the long term with a cut-off of 1.2 nm. The short-ranged, non-bonded interactions were assessed within a cut-off of 1.2 nm. Furthermore, a cut-off distance of 12 Å was assigned to van der Waals and Coulombic interactions. By saving the coordinates for every 2 fs, each system was run for 50 ns. Accordingly, the resulting frames were assessed and visualized, using xmgrace, DS, and Visual Molecular Dynamics (VMD).

## 2.8. Free energy calculations

Calculating the binding free energy in a manner computed, using the widely applicable Poisson-Boltzmann Surface Area Mechanics (MMPBSA) method, is a key step in assessing the structure and function of complexes, following MD simulation [44]. In the present study, the binding free energy of the complexes obtained from MD outputs was evaluated, utilizing the GROMACS' *g\_mmpbsa* module [45, 46]. Using this method, assessing the binding free energy, includes three steps. First, the potential energy in a vacuum is computed. Further, polar and non-polar solvation energies were measured, respectively. Moreover, the solvent-accessible surface area (SASA) model was employed to calculate the non-polar solvation energy. The last 500 frames of each complex were chosen for all computations in order to estimate the binding free energy.

## 2.9. Free energy landscape (FEL)

Since in MD calculations, the selection of the conformation enjoying the lowest free energy can provide a more accurate basis for subsequent calculations, therefore in this study, FEL analysis was carried out, in order to determine the required input structure for Quantum mechanics-molecular mechanics (QM-MM) calculations [47]. Mathematica 11.3 has constructed two-dimensional and three-dimensional maps of FEL, using RMSF, the radius of gyration, and Gibbs free energy extracted from MD trajectories. To extract the representative structure, the FEL calculation was rendered, using GROMACS utility *gmx sham*, and then *gmx trjconv*.

## 2.10. Quantum mechanics-molecular mechanics (QM-MM) calculations

In comparison with the molecular mechanics (MM) method, Quantum Mechanics (QM) more accurately assesses the interactions between the ligands and receptors. However, due to the high computational cost, the QM method is normally restricted to the study of small molecules rather than biomacromolecular systems [48]. Nevertheless, for large system simulations, the hybrid QM-MM technique is more adaptable. In this study, the atoms of ligands and residues located at binding sites were selected to be in the QM site, while other atoms were arranged in the MM region. In addition, a zero charge was appointed to all QM regions. The full optimization of the compounds has been performed, using the OPLS\_2005 force field and DFT method, with M06/B3LYP for

**Table 2**  
Experimental and estimate activity of the training set compounds based on pharmacophore model Hypo1.

Compound No.	Fit value <sup>a</sup>	Experimental Ki nM	Estimate Ki nM	Error <sup>b</sup>	Experimental scale <sup>c</sup>	Estimated scale
ZINC14976187	13.28	0.13	0.36	+2.8	+++	+++
ZINC14976260	12.87	0.7	0.93	+1.3	+++	+++
ZINC14976297	12.51	1.25	2.1	+1.7	++	++
ZINC29128447	12.83	1.4	1	-1.3	++	+++
ZINC29038457	12.56	1.5	1.9	+1.3	++	++
ZINC29128445	12.50	1.8	2.2	+1.2	++	++
ZINC29128362	12.65	2.4	1.6	-1.5	++	++
ZINC14976224	12.24	5.2	4	-1.3	++	++
ZINC14976200	11.89	6.5	8.9	+1.4	++	++
ZINC29128449	12.20	6.9	4.4	-1.6	++	++
ZINC29052073	12.22	7	4.1	-1.7	++	++
ZINC14976175	12.02	7.5	6.6	-1.1	++	++
ZINC29129065	10.90	84	88	+1	++	++
ZINC14976323	10.13	220	510	+2.3	++	++
ZINC14976066	10.30	300	350	+1.2	++	++
ZINC29129064	10.38	420	290	-1.5	+	++
ZINC29129030	10.31	550	340	-1.6	+	++
ZINC14976273	10.16	920	480	-1.9	+	++

<sup>a</sup> Fit value indicates how well the features in Hypo1 overlap the chemical features in the training set compounds.

<sup>b</sup> Positive value indicates that the estimated Ki is higher than the experimental Ki; a negative value indicates that the estimated Ki is lower than the experimental Ki, in nM.

<sup>c</sup> Activity scale: Ki < 1 nM = +++ (highly active); 1 nM ≤ Ki < 300 nM = ++ (moderately active); Ki ≥ 300 nM = + (low active or inactive).

MM and QM regions, respectively, by employing Schrödinger Qsite. Other parameters remained as default.

### 3. Results and discussion

#### 3.1. 3D-QSAR pharmacophore models for ACE2 and hypotheses validation

Table 1 provides detailed results of the top 10 hypotheses of all the models made by the Catalyst HypoGen algorithm. As noted in Table 1, the HYA feature is common to all ten created hypotheses, indicating the significance of this feature. Hypo1 enjoys the top correlation coefficient (0.983), the lowest RMSD (1.16), the minimum total cost (85.22) and is therefore considered as the best pharmacophore model for further evaluations. As shown in Table 1, the Hypo1 scaffold consists of HBA, HBD, and two HYA features. Since these pharmacophore models were constructed based on the structure-activity relationship, therefore, in the comparison between the most active (Ki=0.13 nM) and the inactive (Ki=920 nM) training set ligands, which are shown in their alignment with Hypo1 (Fig. 2), it can be seen that the former covers all the features of the model, while the latter is not fully capable of doing so. Therefore, it can be concluded that Hypo1 could successfully distinguish between active and inactive compounds [35]. Furthermore, the *Ligand Pharmacophore Mapping* module was used to estimate the power of the Hypo1 in predicting the activity of the training set ligands. The results demonstrated the predictive potential of Hypo1, as indicated in Table 2.

In the present study, the validity of the Hypo1 was evaluated, using cost analyses, Fischer's randomization, and test set prediction methods. As presented in Table 1, the total cost and the null cost for Hypo1 are 85.22 and 405.679, respectively. The biggest difference between null cost and total cost (320.45) is in Hypo1, which is more than 60 bits and implies that Hypo1 is capable of representing a correct correlation to the data [49]. A configuration cost value, indicating the complexity of the pharmacophore model below 17, is considered acceptable for an authenticated hypothesis. The mentioned value for Hypo1 is 16.24, which is less than 17.

As a crucial issue in the validation of the pharmacophore hypotheses, the possibility of the construction of a hypothesis by chance has to be ruled out. Fischer's randomization test was re-

quired for this aim to randomly create 19 models at a significance level of 95%, using the Cat-Scramble method. As shown in Fig. 3, the correlation coefficient value of all 10 pharmacophore models is higher than that of all 19 randomly produced models. This means that none of the top ten hypotheses were created by chance [30].

A test set consisting of 13 ACE2 inhibitors (Figure S1) was assessed by Hypo1 to examine whether the selected pharmacophore model enjoys the ability to predict the activity values of distinct ACE2 inhibitors from the training set, which is appropriate to their experimental activity. The test set ligands were categorized as follows in three separate scales, according to their activity: highly active, Ki < 1.7 nM; moderately active, 1.7 nM ≤ Ki < 900 nM; inactive, Ki ≥ 900 nM. As enumerated in Table S1, the results indicate that Hypo1, as for the training set ligands, enjoys an acceptable ability to predict the activity of the test set compounds within their experimental range, except for 4 cases. Two of the most active compounds were deemed moderately active and two moderately active compounds were detected as inactive. As demonstrated in Fig. 4, the correlation coefficient of experimental activity and the estimated activity for the training set and the test set compounds is 0.98 and 0.83, respectively. It can be seen that Hypo1 has not only been able to predict the activity of the training set compounds in a very desirable manner but has also been so effective in estimating the activity of the test set ligands [29].

Overall, based on these findings, it can be concluded that Hypo1 has the necessary competence to properly predict the bioactivity of the study compounds, and was therefore chosen as the best hypothesis for the screening of a large database of possible inhibitors. In reviewing the literature, it was found that several studies were conducted to find effective drugs against COVID-19, and almost none of them were performed based on 3D-QSAR pharmacophore modeling [50–52]. From this point of view, our study of screened potential drugs based on pharmacophoric principles has strength, compared to other investigations.

#### 3.2. Virtual screening and drug-likeness evaluation

In this study, in order to increase the possibility of the discovery of effective medicines, numerous ligands were collected from Pubchem [31] and Zinc [32] databases. The ligands were prepared by the *Prepare Ligands* module implemented in DS, and then



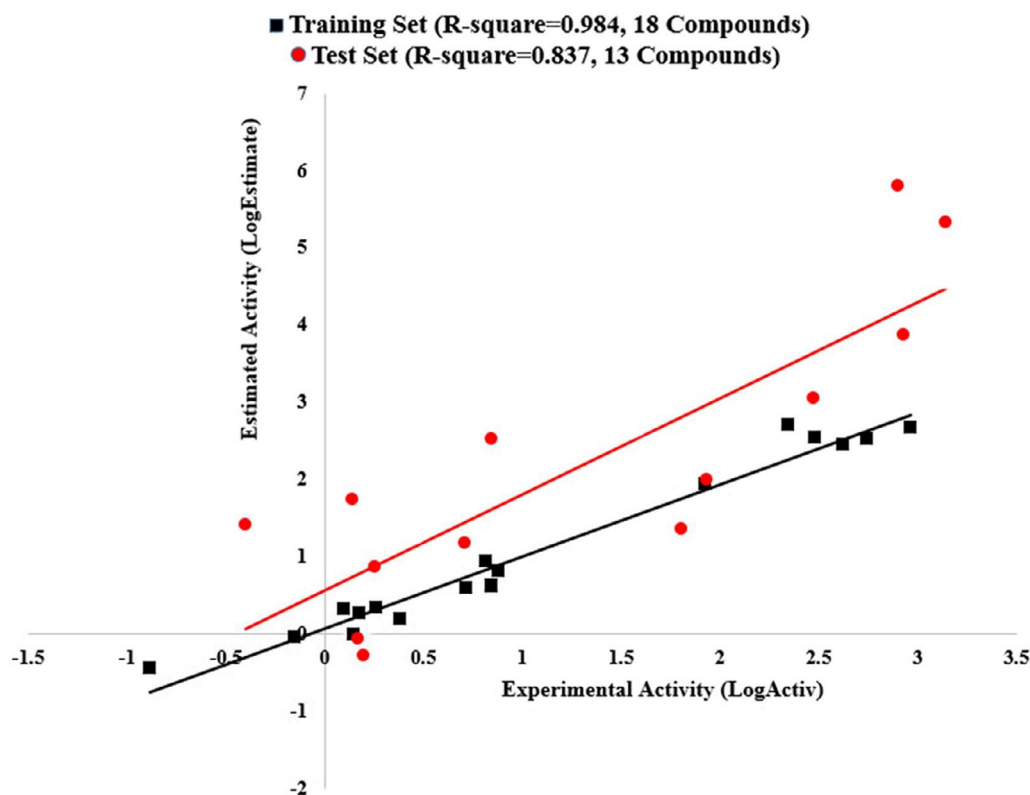


Fig. 4. Correlation graph between experimental and estimated activities in logarithmic scale for training and test set compounds based on Hypo1.

1,264,479 ligands were selected for screening. During ligand preparation, the ionization and 3D ligand coordination states were analyzed and the inappropriate ligands were removed. Hypo1 was recruited for ligand screening. Given that Hypo1 is a reliable model and has been recognized as the best pharmacophore model by the extensive evaluations mentioned in the previous section, any compound that can be passed through the Hypo1-based screening is expected to be capable of acting as a potential inhibitor of ACE2, due to the commonalities associated with the chemical features of Hypo1. Thus, in this step, 492,697 ligands passed the selected pharmacophore successfully and candidates for the next screening step. Subsequently, the filtered ligands were subjected to Lipinski and Veber Rules filter in DS, and 414,631 ligands passed the criteria. To determine efficiently the bioavailable drugs, Lipinski's rule of 5 can be used with specific cut-off values designated to physicochemical specifications (number of H-bond donors or acceptors, lipophilicity, and molecular weight). *In silico* models can evaluate the drug bioavailability as a major aspect [30]. Compounds with a fit value greater than 13 were selected and implemented for ADMET analysis, using the FAF-Drugs4 web server [34]. A group of descriptors, such as drug-like properties of compounds and ADMET with potential therapeutic characteristics can be predicted by this server (<http://fafdrugs3.mti.univ-paris-diderot.fr/descriptors.html>). Moreover, the server applies some limited filters, describing the descriptors' reference scope, for taking into account the potency of the drug [46]. The reference range and descriptors supplied in this experiment are defined in this link: <http://fafdrugs3.mti.univ-paris-diderot.fr/filters.html>. Of the 245 input structures, 81 ligands successfully met the criteria and were therefore subjected to the molecular docking process (Table S2).

### 3.3. Molecular docking analysis

The docking protocol was validated, using the pose selection method [53]. The co-crystal ligand of the ACE2 with PDB ID of 1R4L was extracted and re-docked into the ACE2 active site, by implementing the parameters described in the method section. Protocols are considered to have successful performance if they can return an RMSD value below 1.5 or 2 Å, based on the ligand size [53]. The acceptable RMSD value of 0.95 Å between the co-crystal ligand and its docked pose was obtained, indicating that the docking protocol is reliable (Figure S2).

In this study, two targets, ACE2 and MPro, were selected due to the particular importance of successfully restraining the SARS-CoV-2. The enzymes ACE2 and MPro act for the virus entry, and as the key operator of the virus, respectively, in which MPro requires more focus due to the need to inhibit the virus as specifically as possible. Based on this strategy, the filtered ligands from the previous steps were first docked into the ACE2 active site and, after selecting the best-docked ligands, they were docked with the MPro to obtain common inhibitors of both ACE2 and MPro enzymes.

The 81 output ligands from the virtual screening stage have been docked into the active site of ACE2. At the same time, the training set's most active ligand (ZINC14976187) as the reference drug was also docked into the ACE2 active site. The results were ranked according to CDOCKER\_ENERGY. According to the results, seventy-three compounds have more inhibitory strength than ZINC14976187. The top 10 results are outlined in Table 3. In the next step, seventy-three potential ACE2 inhibitors, along with Cinanserin and Ebselen compounds, identified in the Zhenming et al. [14] study as strong MPro inhibitors (as standard drugs), were docked to the active site of the enzyme MPro. It was found that 29 compounds could be stronger for inhibiting MPro than Cinanserin and Ebselen, the top 10 of which were listed in Table 4. As a result,

**Table 3**  
Molecular docking results of ACE2 inhibitors (top 10).

	Compound	CDOCKER_ENERGY	CDOCKER_INTERACTION_ENERGY	Estimated Ki (nM)	Interacting Residues
1	ZINC12562757 (112,249,612)	63.6671	67.6092	0.452847	Asn149, Arg273, Cys344, His345, Pro346, Thr347, Cys361, Lys363, Asp368, His374, Glu375, His378, His505, Tyr515
2	1,520,626	62.8862	79.2558	0.377976	Glu145, Asn149, Arg273, Phe274, Cys344, Pro346, Lys363, Glu375, Glu402, Tyr510
3	1,804,858	62.4542	66.8731	0.508249	Arg273, Phe274, His345, Pro346, Thr371, Glu375, His505, Tyr510, Tyr515, Arg518
4	112,252,955	61.2618	72.098	0.619392	Asn149, Asp269, Trp271, Arg273, Phe274, His345, Pro346, Ala348, His374, His378, Glu402, His505, Tyr510
5	112,339,522	61.0032	63.6881	0.444996	Arg273, Phe274, His345, Pro346, Thr347, His374, Glu402, Phe504, His505, Tyr510, Arg518
6	2,303,440	60.5106	60.756	0.653971	Arg273, His345, Pro346, His374, Glu406, Arg518
7	112,054,543	60.0011	70.0661	0.399127	Asn149, Trp271, Arg273, Phe274, His345, Pro346, Glu402, Tyr510, Tyr515
8	112,260,215	59.837	62.8354	0.639773	Arg273, Phe274, His345, Pro346, Leu370, Glu375, His378, Phe504, His505, Tyr510, Tyr515, Arg518
9	112,139,218	59.1964	76.4905	0.495676	Ala153, Arg273, Phe274, His345, Pro346, Lys363, Asp367, Thr371, His374, Glu402, Phe504, His505, Tyr510
10	1,811,184	59.1121	63.816	0.657419	His345, Pro346, Thr347, Asp367, Thr371, His374, Glu375, His378, Phe504, His505, Tyr510, Tyr515, Arg518
11	ZINC14976187 (Training Set Lowest Ki) <sup>a</sup>	16.0648	88.4878		Arg273, Phe274, His345, Pro346, Leu370, Thr371, His374, Glu402, His505, Arg518

<sup>a</sup> Represents the control compound.**Table 4**  
Molecular docking results of MPro inhibitors (top 10).

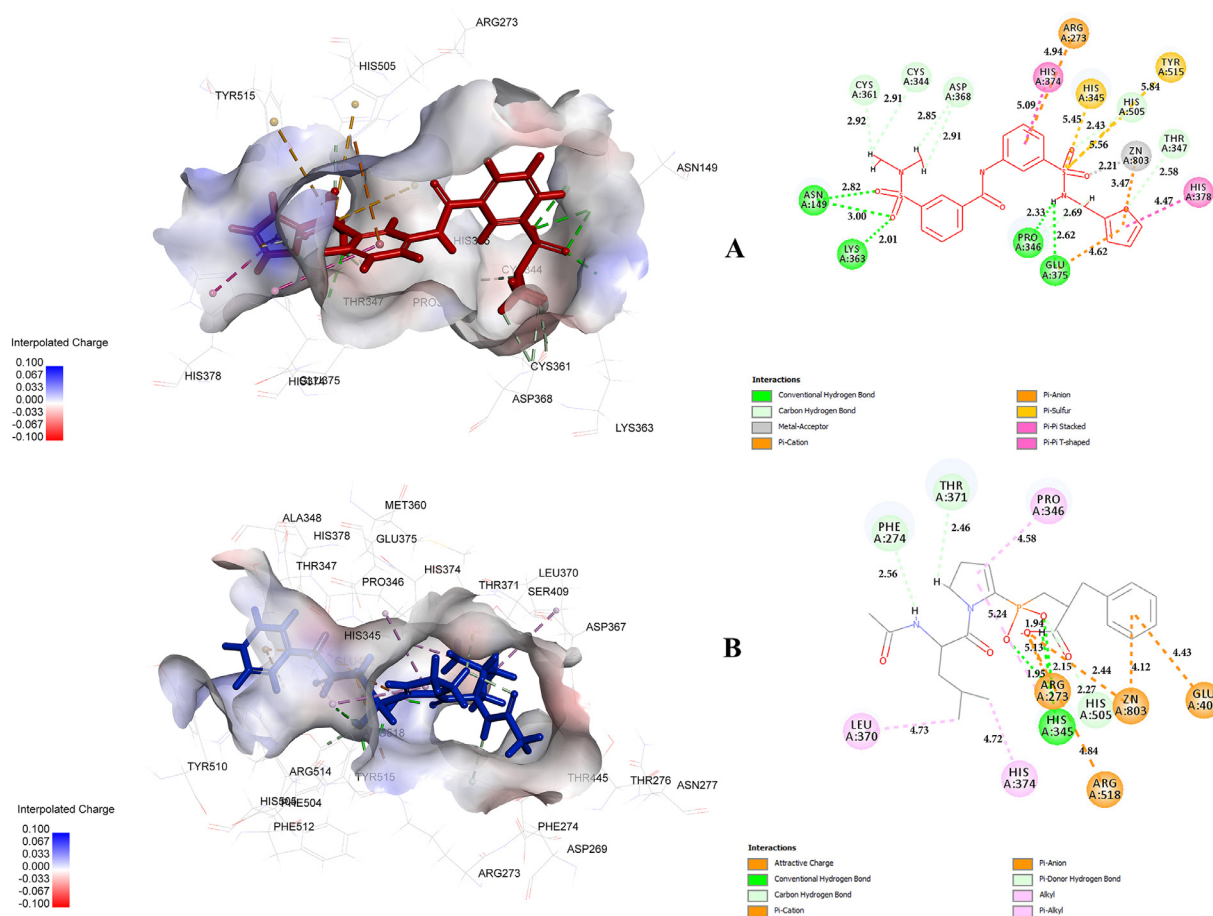
	Compound	CDOCKER_ENERGY	CDOCKER_INTERACTION_ENERGY	Estimated Ki (nM)	Interacting Residues
1	112,260,215	46.7127	45.9017	0.639773	Leu27, His41, Cys145, His164, Met165, Glu166, Pro168, Gln189
2	112,339,522	46.3533	47.0252	0.444996	Leu27, His41, Met49, Cys145, His164, Met165, Glu166, Leu167, Gln189
3	2,085,094	46.2895	53.9721	0.589431	His41, Ser144, Cys145, His164, Met165, Glu166, Arg188
4	2,087,017	43.7821	42.1409	0.689245	His41, Met49, Phe140, Met165, Glu166, Gln189
5	2,119,581	42.6101	58.1075	0.272343	Asn142, Cys145, His163, Met165, Glu166, Gln189, Thr190
6	2,117,728	42.1512	50.8951	0.614062	His41, Met49, Phe140, Leu141, Asn142, Cys145, Met165, Glu166
7	112,126,248	41.3936	41.5074	0.597168	His41, Met49, Leu141, Cys145, Met165, Glu166
8	112,212,289	41.0756	40.8123	0.593469	Gly143, Ser144, Cys145, Arg188
9	112,249,612	40.1564	45.1435	0.452847	Thr26, Met49, Asn142, Gly143, Cys145, Met165, Glu166
10	112,252,955	38.9913	50.5727	0.619392	Leu27, His41, Met165, Glu166, Gln189, Thr190
11	Cinanserin <sup>a</sup>	31.5818	41.3993		Cys145, Met165, Asp187, Gln189
12	Ebselen <sup>a</sup>	3.03982	25.1226		Met49, Met165, Glu166

<sup>a</sup> Represents the control compound.

these 29 compounds have been identified as common and potential inhibitors of dual-power inhibition of ACE2 and MPro.

ZINC12562757 was identified as the strongest potential ACE2 inhibitor with CDOCKER\_ENERGY of 63.66. In comparison to other inhibitors, as shown in Table 3, ZINC12562757 has formed the highest number of bonds with the key amino acid residues of the ACE2 active site and interacted with the greatest number of key amino acid residues. Fig. 5A illustrates the two- and three-dimensional docked poses and the interactions' distances of this compound. As shown in this figure, ZINC12562757 was able to establish two pi-pi interactions with His374 and His378 with distances of 5.09 Å and 4.47 Å, one pi-anion with Glu375 (4.62 Å), one pi-cation with Zn<sup>2+</sup> (3.47 Å), three pi-sulfurs with His345, His505 and Tyr515 with distances of 5.45 Å, 5.56 Å and 5.84 Å, seven carbon-hydrogen bonds with Cys344, Pro346, Thr347, Cys361, Asp368 and His505 with distances of 2.91 Å, 2.69 Å, 2.58 Å, 2.92 Å, 2.85 Å, 2.91 Å and 2.43 Å, one metal acceptor bond with Zn<sup>2+</sup> (2.21 Å) and five conventional hydrogen bonds with Asn149, Pro346, Lys363, and Glu375 with distances of 2.82 Å, 3 Å, 2.33 Å, 2.01 Å and 2.62 Å, respectively. On the other hand, ZINC14976187 with docking energy of 16.06 was able to create one pi-anion bond with Glu402 (4.43 Å), one pi-cation bond with Zn<sup>2+</sup> (4.12 Å), one pi-donor hydrogen bond with Phe274 (2.56 Å), two carbon-hydrogen bonds with Thr371 and His505 (2.46 Å and 2.27 Å), three conventional hydrogen bonds with Arg273 and His345 with distances of 1.94 Å, 1.95 Å and 2.15 Å, two alkyl

bonds with Pro346 and Leu370 (4.58 Å and 4.73 Å), and one pi-alkyl bond with His374 (4.72 Å), respectively (Fig. 5B). All hydrogen bond geometries that fall within the 90° to 180° cut-off were accepted. Hydrogen bonds with donor-acceptor distances of 2.2 – 2.5 are classified as "strong, mostly covalent," 2.5 – 3.2 as "moderate, mostly electrostatic," and 3.2 – 4.0 as "weak, electrostatic," according to Jeffrey [54]. As per this classification, all hydrogen bonds formed between ZINC12562757 and ACE2's binding site residues are strong to moderate. Thus, establishing more potent hydrogen bonds could be one of the probable reasons for ZINC12562757's highest docking energy. Moreover, Towler et al. crystallized the structure of ACE 2 and studied the structure carefully, in particular the enzyme active site [55]. Based on their report, the side chains of Arg273, His505, and His345 are involved in hydrogen bonding with the MLN-4760, the ACE2's co-crystallized potent inhibitor. Moreover, Pro346, Thr371, Glu375, and Tyr515 are forming H-bonds or being within H-bonding distance of the functional groups of MLN-4760. Thus, interaction with these residues plays a vital role in the potential inhibition of ACE2. As shown in Fig. 5, ZINC12562757 interacted with all of the key ACE2 residues mentioned above, except for Thr371. On the other hand, ZINC14976187 could not establish any interactions with Glu375 and Tyr515. In stabilizing certain specific transitional states in the ACE2 reaction mechanism, Tyr515 enjoys a specific function, so that interaction with that probably plays an important role in inhibiting the enzyme [55]. As illustrated in Fig. 5, ZINC12562757 has formed a pi-



**Fig. 5.** The schematic representations (2D and 3D) of the binding interactions, between the ACE2 active site and (A) ZINC12562757, (B) ZINC14976187 (Training Set Lowest Ki).

sulfur interaction with Tyr515 and this could be other likely causes of ZINC12562757's stronger docking energy.

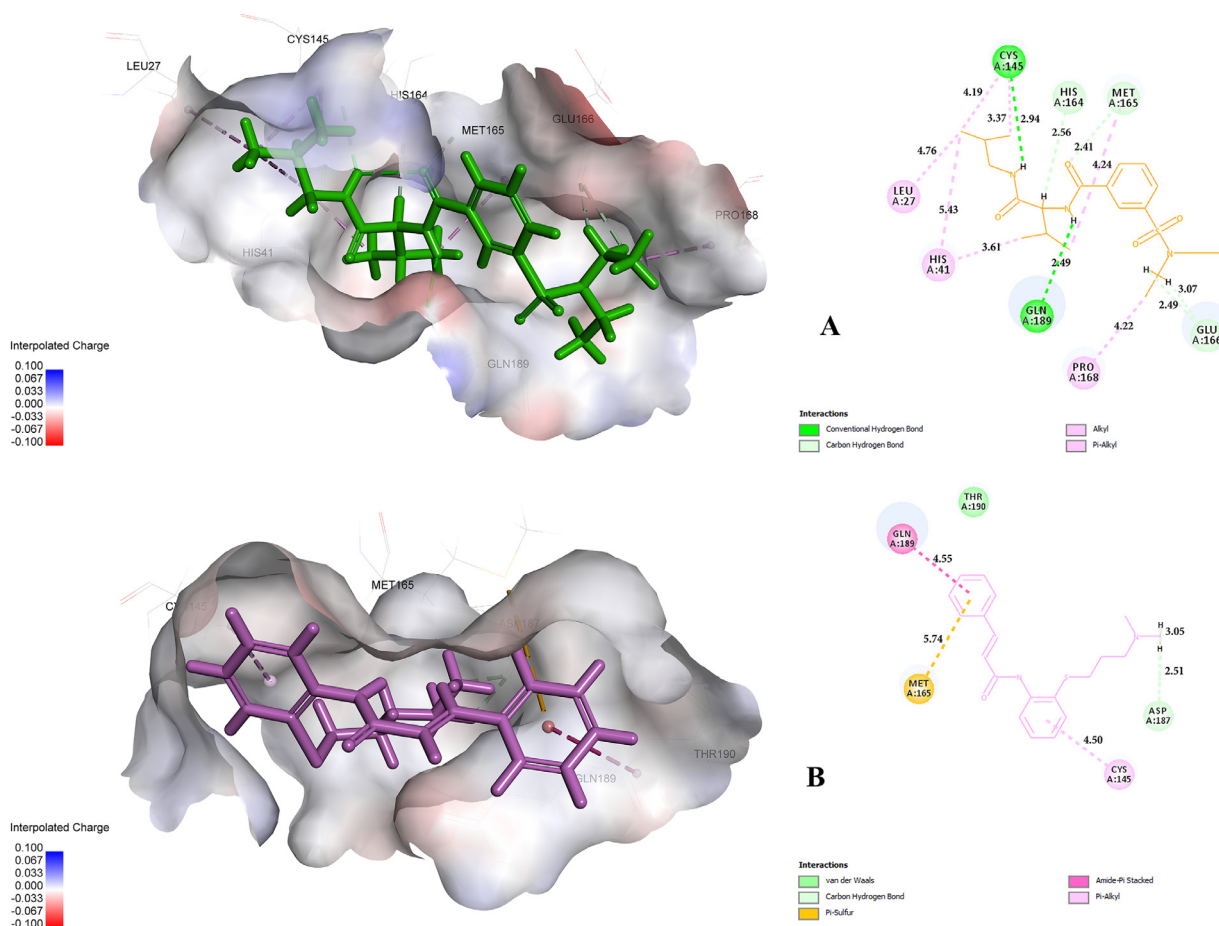
The compound 112,260,215, with 46.71 docking energy as the strongest inhibitor of MPro, was able to establish four carbon-hydrogen bonds with His164, Met165, and Glu166 amino acids with distances of 2.56 Å, 2.41 Å, 2.49 Å, and 3.07 Å, two conventional hydrogen bonds with Cys145 and Gln189 (2.94 Å and 2.49 Å), five alkyl bonds with Leu27, Cys145, Met165, and Pro168 with distances of 4.76 Å, 3.37 Å, 4.19 Å, 4.24 Å and 4.22 Å, and two pi-alkyl bonds with His41 (3.61 Å and 5.43 Å), respectively (Fig. 6A). While Cinanserin established, one amide-pi stacked with Gln189 (4.55 Å), one pi-sulfur extension with Met165 (5.74 Å), one pi-alkyl bond with Cys145 (4.50 Å), and one van der Waals bond and carbon-hydrogen bond with Asp187 (2.51 Å and 3.05 Å), respectively (Fig. 6B). The 90° to 180° cut-off was used to accept all hydrogen bond geometries. It can be seen that compound 112,260,215, in comparison to Cinanserin, created many more interactions especially potent hydrogen bonds with MPro active site residues. Zhenming Jin et al. reported in a detailed study that the residues of His41, Met49, Tyr54, Phe140, Asn142, Cys145, His163, Met165, Glu166, Leu167, His172, Phe185, Asp187, and Gln192 are highly conserved across 12 different species of MPro enzymes and serve as the ligand-binding pocket [14]. Of the 8 amino acids that interacted with 112,260,215 compounds, four of them, His41, Cys145, Met165, and Glu166, are the key residues of the MPro active site; however, Cinanserin interacted with 4 amino acids in total, including Cys145, Met165, Asp187, and Gln189. Figure S3 shows the two-dimensional (2D) structure of the best inhibitors of ACE2 and MPro, along with the standard drugs.

A survey of ligand interactions with active site residues of the ACE2 reveals that residues Arg273, Phe274, His345, Pro346, His505, and Tyr510 play a role in almost all interactions, helping to bind potential inhibitors to the active site of the ACE2, which may indicate the pivotal importance of these residues in the catalytic function of the ACE2. The study of interactions of inhibitory compounds with MPro also indicates the key role of Cys145, Met165, and Glu166 residues in the enzyme inhibition by docked ligands. Our docking results are consistent with those of previous studies [50, 52, 56, 57].

In total, the compounds ZINC12562757 and 112,260,215 were selected as the strongest ACE2 and MPro inhibitors based on the CDOCKER\_ENERGY, binding mode, interacted residues, and potency of hydrogen bond based on distances and geometries according to Jeffrey's category [54]. However, finding a potentially specific drug for direct inhibition of SARS-CoV-2 is a vital requirement; therefore, due to the direct role of MPro and the indirect role of ACE2 in virus activity, the compound 112,260,215 was selected, as the best potential inhibitor of MPro, as the backbone for the design of new, stronger ligands that inhibit MPro, using the *de novo* evolution method.

#### 3.4. De novo evolution analysis

Recruiting the backbone of the most powerful MPro inhibitor (112,260,215), using the *Ludi* algorithm, ten *de novo* compounds have been developed and classified according to the *Ludi* energy. The 2D structure of these compounds is shown in Fig. 7. These compounds have been docked back with MPro, the results of



**Fig. 6.** The schematic representations (2D and 3D) of the binding interactions, between the MPro active site and (A) 112,260,215, (B) Cinanserin (MPro standard inhibitor).

**Table 5**

LUDI energy and molecular docking results of *de novo* Evolution designed ligands with MPro.

Rank	Name	CDOCKER_ENERGY	CDOCKER_INTERACTION_ENERGY	Estimated Ki (nM)	Fit Value	LUDI 3	Fragments	Interacting Residues
1	Evo_1	50.2016	60.9334	17.811	11.5904	621	S12 SM6	Met49, Tyr54, Asn142, His163, Met165, Glu166, Leu167, Arg188, Thr190
2	Evo_4	46.7459	62.1499	2020.67	9.5356	608	S12 M68	His41, Gly143, Cys145, His163, His164, Met165, Glu166, Pro168, Gln189, Ala191
3	Evo_10	45.5299	59.8881	1141.27	9.78371	598	S12 SN9	His41, Met49, Leu141, Asn142, Gly143, Cys145, Met165, Glu166, Leu167, Pro168, Gln189, Thr190
4	Evo_2	44.7561	59.0297	1944.75	9.55224	610	S12 MJ5	His41, Met49, Asn142, Cys145, His164, Met165, Glu166, Gln189, Thr190
5	Evo_7	43.6307	58.8671	2146.59	9.50935	605	S12 SJ4	Leu27, His41, Met49, Gly143, Cys145, His164, Glu166, Leu167, Pro168, Arg188, Gln189
6	Evo_5	40.5554	61.8641	6094.69	9.05615	607	S12 SU3	Thr26, His41, Met49, Asn142, Gly143, Cys145, His163, His164, Met165, Glu166, Pro168, Asp187, Arg188
7	Evo_9	34.8059	59.5539	341.732	10.3074	598	S12 SC8	His41, Met49, Cys145, His163, Met165, Glu166, Pro168, Gln189, Thr190, Ala191
8	Evo_8	34.6544	56.3502	89.9031	10.8873	599	S12 D41	His41, Asn142, Met165, Glu166, Leu167, Pro168, Thr190
9	Evo_6	33.8748	51.6242	6169.65	9.05084	607	S12 C34	Met49, Phe140, Cys145, Met165, Glu166, Leu167, Pro168, Gln189, Thr190
10	Evo_3	17.8873	58.7464	594.035	10.0673	608	S12 D07	His41, Phe140, Cys145, His163, Met165, Glu166, Pro168, Gln189

which are shown in Table 5. The compounds Evo\_1 and Evo\_4 had *Ludi* energy of 621 and 608, respectively, and CDOCKER\_ENERGY of 50.20 and 46.74, indicating their higher binding power, compared to all inhibitory MPro compounds listed in Table 4. The results of the ADMET analysis demonstrated the success of the Evo\_1, Evo\_6, and Evo\_7 compounds in meeting the criteria, while the rest of the

compounds were rejected (Table S3). The Evo\_1 had not only the highest docking and *Ludi* energy, but also the lowest estimated Ki and the highest fit value as a result of the pharmacophore mapping of *de novo* compounds with Hypo1, and indicating that Evo\_1 mapped all Hypo1 features. Additionally, Evo\_1 passed the ADMET criteria successfully, while Evo\_4 did not. Evo\_1 was therefore cho-



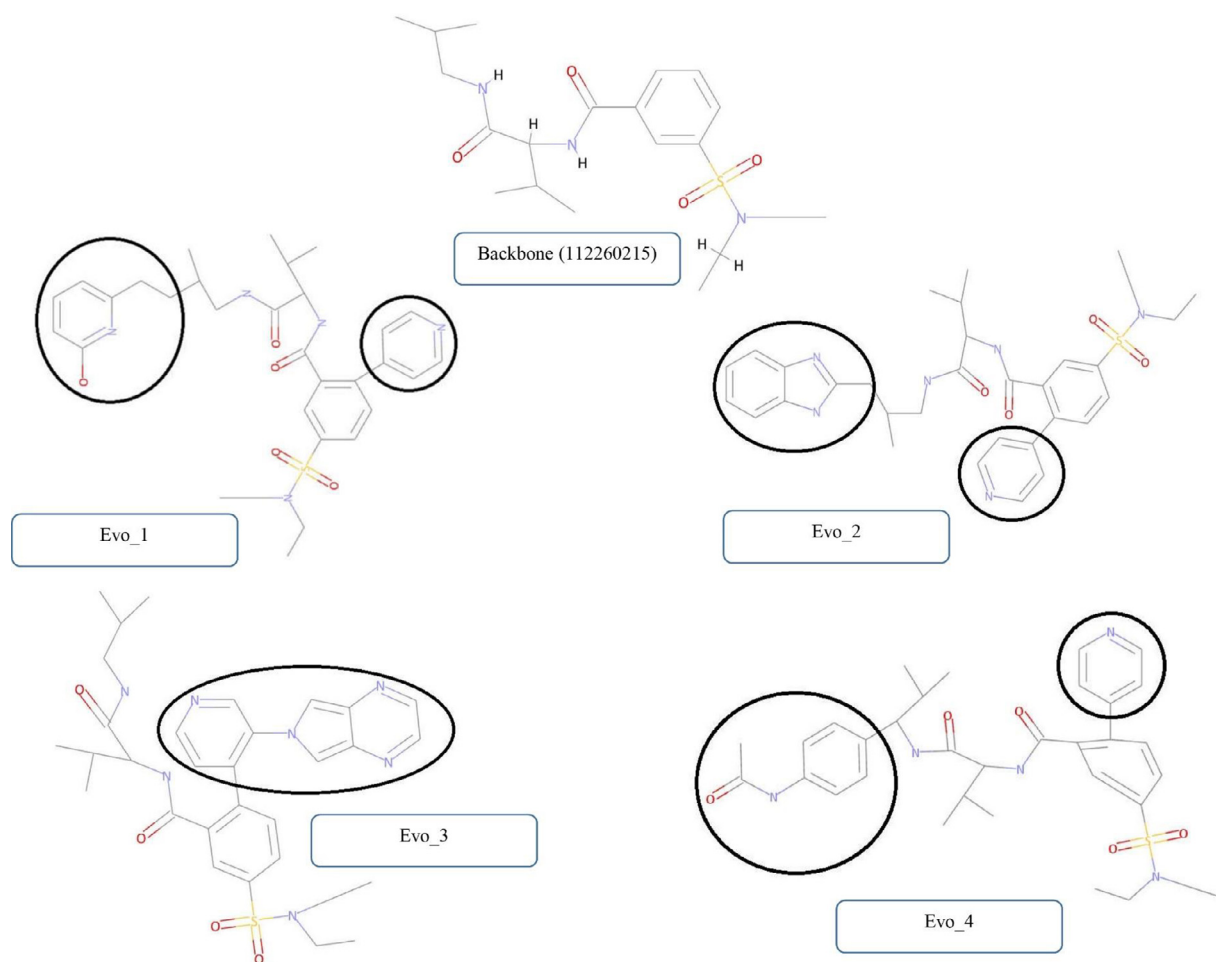


Fig. 7. The structures of the derivatives selected after *de novo* Evolution. The functional groups added in *de novo* Evolution are circled in black.

sen as the strongest inhibitor of the MPro among the *de novo* compounds based on CDOCKER\_ENERGY, ADMET properties, estimated Ki, fit value, and Ludi energy scales.

As shown in Fig. 8, 2D and 3D docking pose of Evo\_1 to MPro active site reveal that Evo\_1 was able to interact with active site residues through two alkyl bonds with Met165 and Leu167 with distances of 5.46 Å and 5.41 Å, three pi-alkyl bonds with Met49, His163, and Met165 with distances of 4.80 Å, 4.11 Å and 4.12 Å, four carbon-hydrogen bond with Glu166, Arg188 and Thr190 (2.48 Å, 2.49 Å, 2.71 Å and 3.01 Å) and two conventional hydrogen bonds with Tyr54 and Asn142 (3.07 and 2.45 Å), respectively. All hydrogen bond geometries falling between 90° and 180° have been accepted. Examining the interactions of Evo\_1 and 112,260,215 with MPro active site residues, it can be seen that although the 112,260,215 compound was able to interact with 4 key amino acids of the active site, Evo\_1 was able to establish 7 interactions with conserved key residues of the active site of the MPro and that is why it seems to have higher CDOCKER\_ENERGY, inhibiting MPro more effective than the 112,260,215 compound. The compound 112,260,215 formed two pi-alkyl bonds with His41, while Evo\_1 did not interact with His41. One pi-alkyl bond was established between Evo\_1 and Met49, while 112,260,215 was not able to establish any interaction with. Evo\_1 formed a conventional hydrogen bond with tyr54, but compound 112,260,215 was not able to do so. Asn142 is another key residue that unlike 112,260,215, Evo\_1 forms a conventional hydrogen bond with. Although in the

case of Cys145, the 112,260,215 compound showed no common ground with Evo\_1 by establishing 1 conventional hydrogen bond and 2 alkyl bonds with Cys145. On the other hand, only Evo\_1 was able to form 1 pi-alkyl bond with His163 and 1 alkyl bond with Leu167. Both could have interacted with Met165 and Glu166. A comparison between Evo\_1 and potential MPro inhibitors' interactions with the MPro active site residues shows that residues Met165 and Glu166 play a significant role in the binding of compounds with the MPro, which could be of particular importance in the synthesis of new strong MPro inhibitors.

An examination of the functional groups added to the backbone of the 112,260,215 compound in the *de novo* evolution method, as demonstrated in Fig. 7, shows that two aza-phenol and aza-benzene substitutions were added to the backbone in Evo\_1. Aza-benzene substitution has been added in all other *de novo* compounds, so it indicates that addition of these two functional groups, particularly aza-benzene, to the backbone has been able to increase the compound's binding inhibitory power. Overall, an effective strategy that considers the two key amino acids, Met165 and Glu166, and the recruitment of the aza-benzene functional group and its derivatives, could be critical in the synthesis of robust MPro inhibitors and subsequent successful inhibition of SARS-CoV-2. To the best of our knowledge, this is the first study designing new potential drugs based on the *de novo* evolution method [58, 59]. In order to check the stability and behavior of the enzyme-inhibitor complex, the best inhibitors of

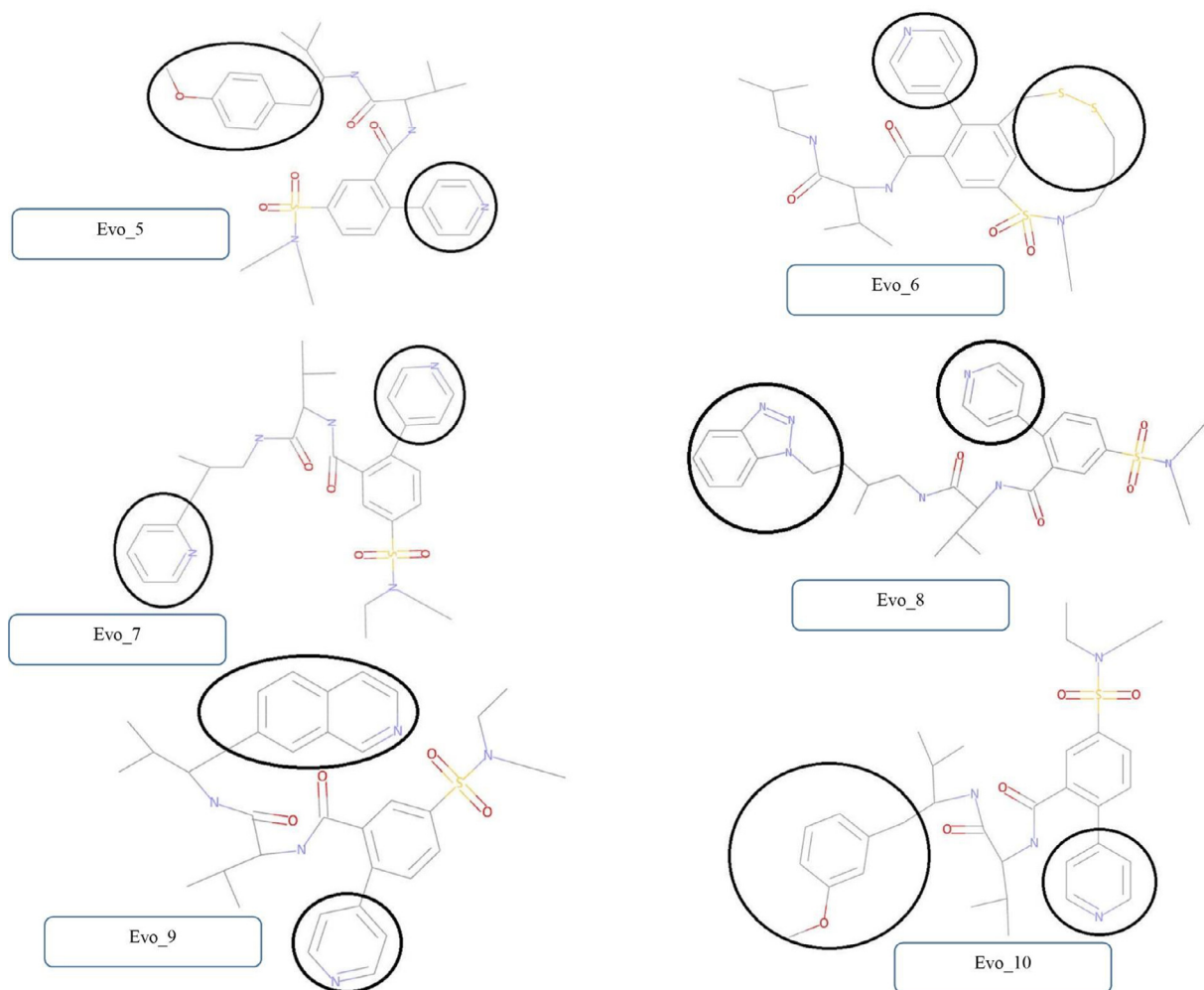
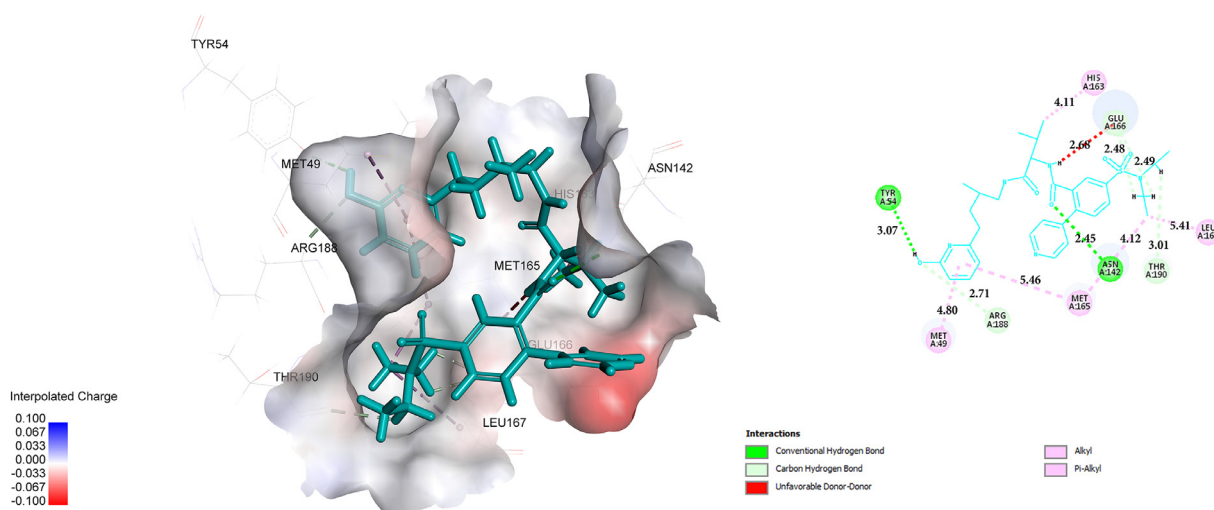


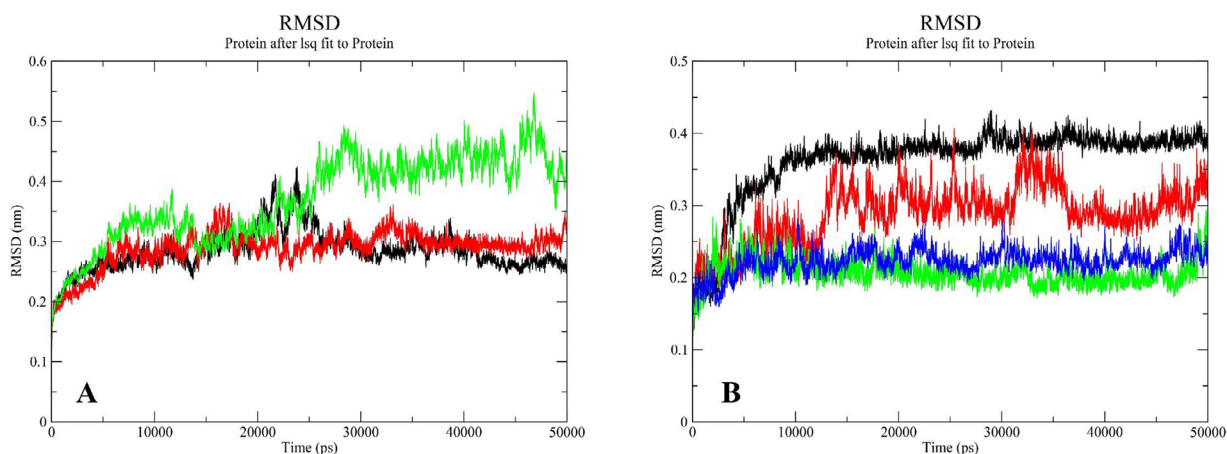
Fig. 7. Continued

Fig. 8. The schematic representations (2D and 3D) of the binding interactions, between the MPro active site and Evo\_1 *de novo* compound.

ACE2 and MPro with their standard inhibitors, which include ACE2-ZINC12562757, ACE2-ZINC14976187, MPro-112,260,215, MPro-Cinanserin, and MPro-Evo\_1 complexes (Figure S3), along with free ACE2 and free MPro, are implemented in the process of molecular dynamics simulation.

### 3.5. Molecular dynamics simulation analysis

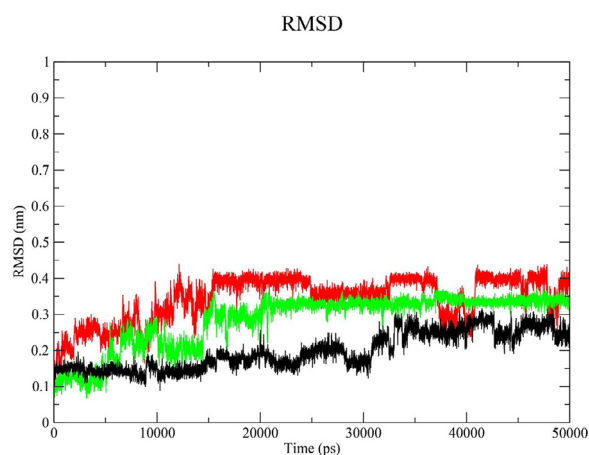
Molecular dynamics simulation is one of the most powerful *in silico* tools, which enables researchers to study behavior and characteristics, resulting from the influence of internal and external forces on the protein structure. In this study, the



**Fig. 9.** The root mean square deviation (RMSD) values of (A) free ACE2 (green), ACE2- ZINC14976187 (Training Set Lowest Ki) (black) and ACE2-ZINC12562757 (red), (B) free MPro (blue), MPro-112,260,215 (black), MPro-Cinanserin (red) and MPro-Evo\_1 (green) complexes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ACE2-ZINC12562757, ACE2-ZINC14976187, MPro-112,260,215, MPro-Cinanserin, and MPro-Evo\_1 complexes along with the free-form of the enzymes were simulated for 50 nanoseconds to investigate the macroscopic properties of the system, including RMSD, RMSF, H-bond, and DSSP.

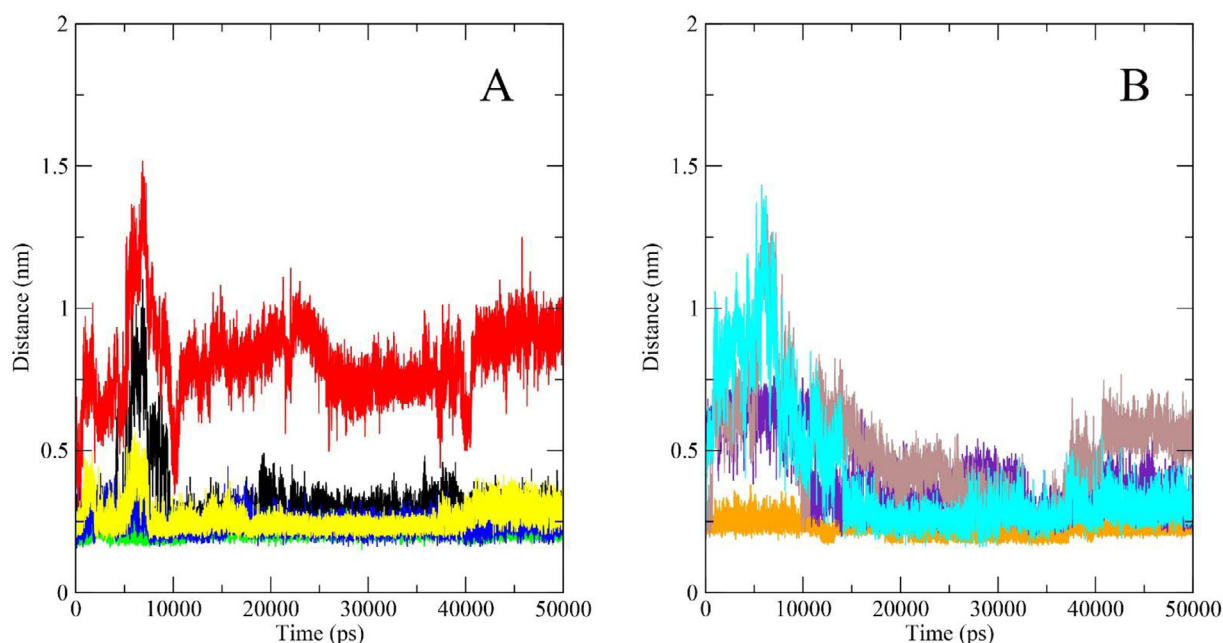
RMSD is an examination that explores the deviations of the atoms and residues of the system from their initial position in the coordinate space during the simulation process, the results of which represent the stability or instability of the system, the conformational changes of the protein domains, and other system components. By calculating the RMSD analysis, as shown in Fig. 9A, free ACE2 is affected by forces from the beginning of the simulation time and the RMSD level begins to rise, gaining relative stability in 15 to 20 nanoseconds and again, by the end of the simulation time, the RMSD has increased again, which could indicate a lack of proper stability in the free ACE2 structure throughout the simulation. On the other hand, the ACE2-ZINC14976187 complex had acceptable stability over the simulation period, except for a range of 20 to 25 nanoseconds, in which an increase in its RMSD was observed. Moreover, the ACE2-ZINC12562757 complex, which is the best potential ACE2 inhibitor, has had good stability throughout the simulation. From this analysis, it can be concluded that ligand binding to the active site of the ACE2 plays a role in the structural stability of the enzyme and ZINC12562757 has been more effective than the standard inhibitor in this field. Furthermore, Fig. 9B shows that free MPro enjoys notable stability during the simulation period, with a slight fluctuation range in RMSD. In the presence of Cinanserin as a standard inhibitor, MPro did not have good stability. At the beginning of the simulation time, the RMSD level in the MPro-112,260,215 complex increased significantly, but the enzyme preserves its stability during the simulation period. Additionally, over the period, the Evo\_1 compound has been able to inhibit and stabilize the MPro conformational stability in a more specific way than other compounds. Moreover, Kumar et al. [60] reported that the MPro structure preserves its stability during MD simulation, which agrees with our findings. In order to take a closer look at the effect of ligands on MPro, the RMSD of ligands was also examined. As it is shown in Fig. 10, in the range of about 1 Å, the 112,260,215 compound had a slight increase in RMSD over the simulation time and was reasonably stable. Cinanserin's deviation from its initial position is gradually increasing from the beginning to the time of 15 ns, and then maintaining its stability with deviation in a range of 3 Å by the end of the simulation time. Evo\_1 had an RMSD increase for up to 15 ns and was quite stable until the simulation period ended. These results could indicate the stability of the lig-



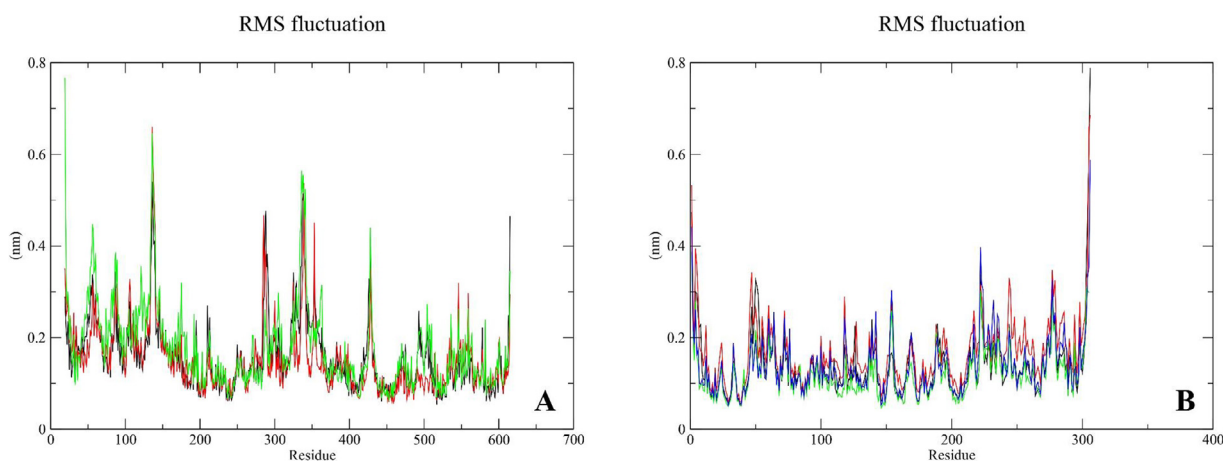
**Fig. 10.** RMSD values of 112,260,215 (black), Cinanserin (red) and Evo\_1 (green) ligands. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and and their interactions with key residues of the MPro enzyme active site and further confirmed the docking findings of MPro inhibitors, especially Evo\_1. Moreover, the distance between Evo\_1 and interacting residues in docking was analyzed during simulation time and depicted in Fig. 11. As illustrated in Fig. 11A, the distance between Evo\_1 and Met49 increased to about 10 ns, while remained quite stable until the end of the simulation period. In the case of Tyr54, although the distance from Evo\_1 in 5 to 10 ns of the simulation time was violated from the defined cut-off distance of 12 Å, its stability was preserved within a range of about 5 Å till the end of the simulation time. The distance between Evo\_1 and Asn142, His163, and Met165 was observed fairly constant and stable during MD simulation. Moreover, as shown in Fig. 11B, the distance between Evo\_1 and Glu166 remained constant during the simulation time. In the case of Leu167 and Arg188, the distance preserved in a range of about 5 Å, but the distance between Evo\_1 and Thr190 increased to about 15 ns, while remaining quite stable until the end of the simulation period. These results strongly support the findings of Evo\_1 RMSD and docking analysis.

RMSF analysis examines residual fluctuations and flexibility during the simulation. The higher the fluctuations the lower the stability will be. Fig. 12A shows that most residues have an RMSF value of less than 0.3 nm, and ZINC12562757 has the greatest impact on the stability of the ACE2 structure, which can be de-



**Fig. 11.** Depiction of the distances between Evo\_1 and (A) Met49 (black), Tyr54 (red), Asn142 (green), His163 (blue) and Met165 (yellow), (B) Glu166 (orange), Leu167 (violet), Arg188 (brown) and Thr190 (cyan). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



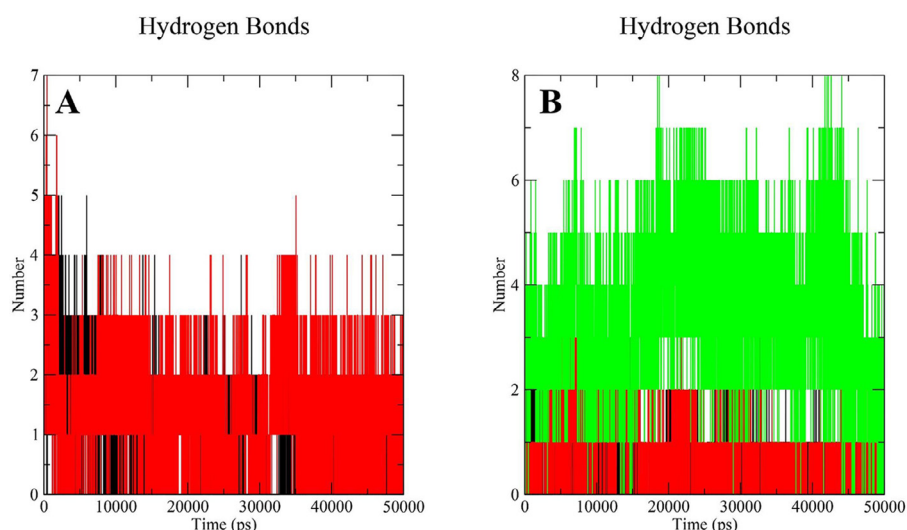
**Fig. 12.** The root mean square fluctuation (RMSF) values of (A) free ACE2 (green), ACE2- ZINC14976187 (Training Set Lowest Ki) (black) and ACE2-ZINC12562757 (red), (B) free MPro (blue), MPro-112,260,215 (black), MPro-Cinanserin (red) and MPro-Evo\_1 (green) complexes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

duced from the lower RMSF over the simulation period. Fig. 12B, on the other hand, shows that Cinanserin causes more fluctuations in residues of MPro than other inhibitors and that the Evo\_1 compound results in less flexibility and greater stability in the MPro structure. The results of the RMSF analysis of the two enzymes confirm the data of the RMSD analysis. This result supports the findings of the previous study [61].

The H-bond analysis was carried out to determine the intermolecular interactions and binding strength of the ligands to the active sites of enzymes ACE2 and MPro [58]. Fig. 13A shows that ZINC12562757 establishes 2 to 7 hydrogen bonds with the residues in the active site of the ACE2, whereas ZINC14976187 creates a maximum of 5 hydrogen bonds throughout the simulation. This might indicate a stronger binding power on the ZINC12562757 than the ZINC14976187. Moreover, as shown in Fig. 13B, the Cinanserin and 112,260,215 compounds were able to establish a maximum of 3 hydrogen bonds with the active site residues of the MPro, while the Evo\_1 compound generated 7 bonds, and a

maximum of 8 hydrogen bonds with the active site of the MPro in most simulation times, indicating the outstanding power of Evo\_1 in binding process and inhibition of the MPro. The capability of hydrogen bond formation plays a major role in ligand bioactivity [62]. As shown in Table 6, ACE2's Arg273, Glu375, and Arg518 residues had the highest percentage of H-bond occupancy during simulation time when forming hydrogen bonds with ZINC14976187. His345, Pro346, and Arg514 also hold the largest share of H-bond occupancy in hydrogen bond formation with ZINC12562757. ZINC12562757 had the highest percentage of H-bond occupancies with key residues Pro346 and His345, while Arg518 and Glu375 played the most roles for ZINC14976187 residues, which may indicate the reason for ZINC12562757's higher strength than ZINC14976187 in ACE2 inhibition. Furthermore, in the formation of hydrogen bonds with MPro active site, residues Glu166 and Gln189, in bonding with Cinanserin and residues Glu166 and Asp187 in connection with 112,260,215, showed the highest percentage of H-bond occupancy, respectively, which Cinanserin and 112,260,215





**Fig. 13.** H-bond values of (A) ACE2-ZINC14976187 (Training Set Lowest Ki) (black) and ACE2-ZINC12562757 (red), (B) MPro-112,260,215 (black), MPro-Cinanserin (red) and MPro-Evo\_1 (green) complexes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 6**

Hydrogen bond occupancy between residues of ACE2, MPro and ligands (ZINC14976187, ZINC12562757, Cinanserin, 112,260,215, Evo\_1) complexes.

Name	Hydrogen bond occupancies
ACE2-ZINC14976187	Arg273 (10.6%), Glu375 (20.4%), Arg518 (34.9%)
ACE2-ZINC12562757	His345 (21.9%), Pro346 (57%), Arg514 (18.8%)
MPro-Cinanserin	Glu166 (41.7%), Gln189 (12%)
MPro-112,260,215	Glu166 (39.4%), Asp187 (22.1%)
MPro-Evo_1	His41 (62.1%), Asn142 (19.7%), Gly143 (39%), Ser144 (24.2%), Cys145 (58.6%), Glu166 (36.4%), Gln189 (76.1%)

established these hydrogen bonds with one and two key residues of the MPro active site, respectively; thus this could be one of the reasons why the compound 112,260,215 has more potency in inhibiting MPro than Cinanserin. On the other hand, H-bond occupancies demonstrated that the residues His41, Asn142, Gly143, Ser144, Cys145, Glu166, and Gln189 were the major contributors to the formation of H-bond with Evo\_1, most of which were the key residues of the active site of the MPro. This result strongly supports the Evo\_1 greater inhibitory activity relative to the other MPro inhibitors. The results of the H-bond occupancy analysis are in good agreement with the docking findings; however, the number and type of hydrogen bonding residues are slightly different in the two methods, which may be due to the ligand and protein dynamics and conformation changes over time of MD simulation.

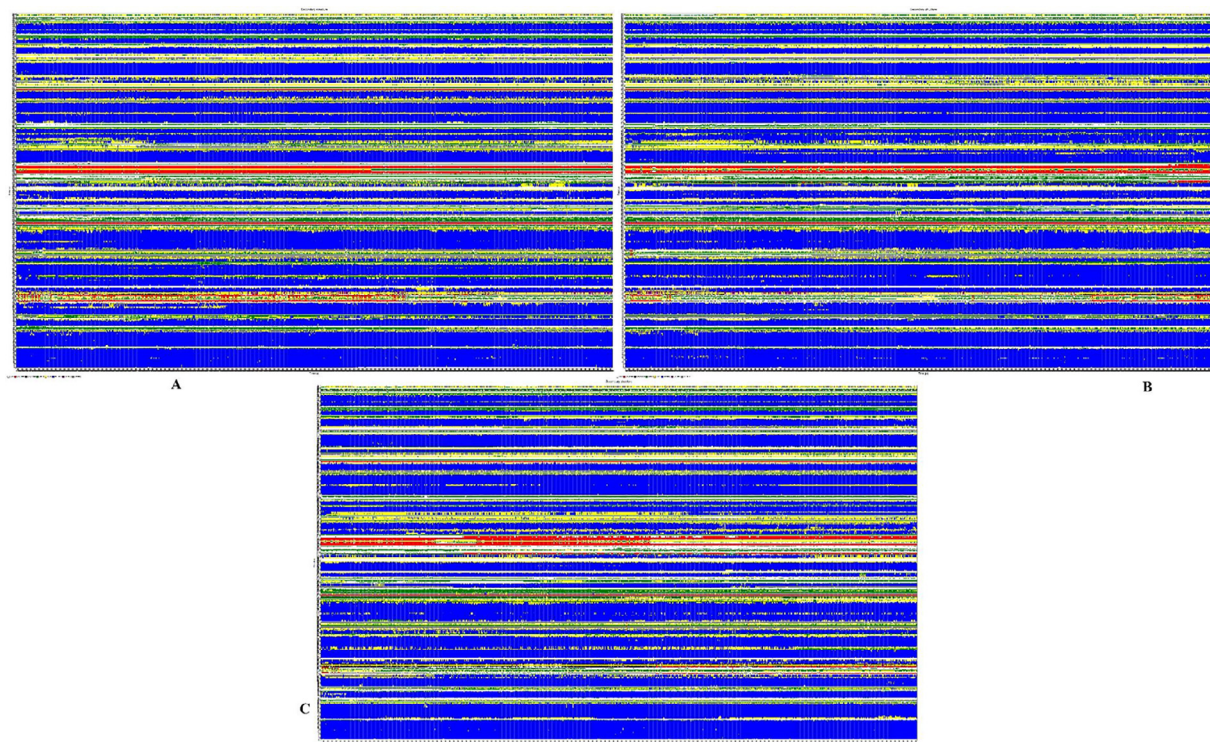
DSSP analysis was carried out in the next step to investigate the alterations in the enzyme's secondary structure. Fig. 14 shows that by comparing the structure of free ACE2, ACE2-ZINC14976187, and ACE2-ZINC12562757, the changes in the secondary structures are not very noticeable and the secondary structure of the proteins is preserved during simulation time, except in limited areas. By exploring the details, it can be said that the key residues of the enzyme ACE2 active site, including Arg273 and Phe274 in free ACE2 and ACE2-ZINC14976187, were constantly changing from the coil to bend during simulation, while those changes were rare in ACE2-ZINC12562757. Residues of His345 and Pro346 in free ACE2 and ACE2-ZINC14976187 have been changed from the coil to bend at different times, while there is no structural change in these residues in ACE2-ZINC12562757. The His505 and Tyr510 residues remained unchanged in all three ACE2 modes over the simulation time. These results confirm the results of RMSD and RMSF and indicate that ACE2 is more stable in binding to ZINC12562757. The amino acids Met165 and Glu166 jointly played a key role in binding to Cinanserin, 112,260,215, and Evo\_1 at the active site of the MPro. In Fig. 15, by examining these amino acids, it can be seen

that the secondary structure of these 2 amino acids fluctuates between coil and  $\beta$ -sheet structures at different times of simulation in free MPro, MPro-112,260,215, and MPro-Cinanserin, most of which are dedicated to the coil. However, these amino acids are often  $\beta$ -sheet and less coil-modified, in MPro-Evo\_1. The results of this analysis confirm the results of RMSD and RMSF, and indicate the higher strength of Evo\_1 in inhibiting and stabilizing the enzyme structure than other MPro inhibitors.

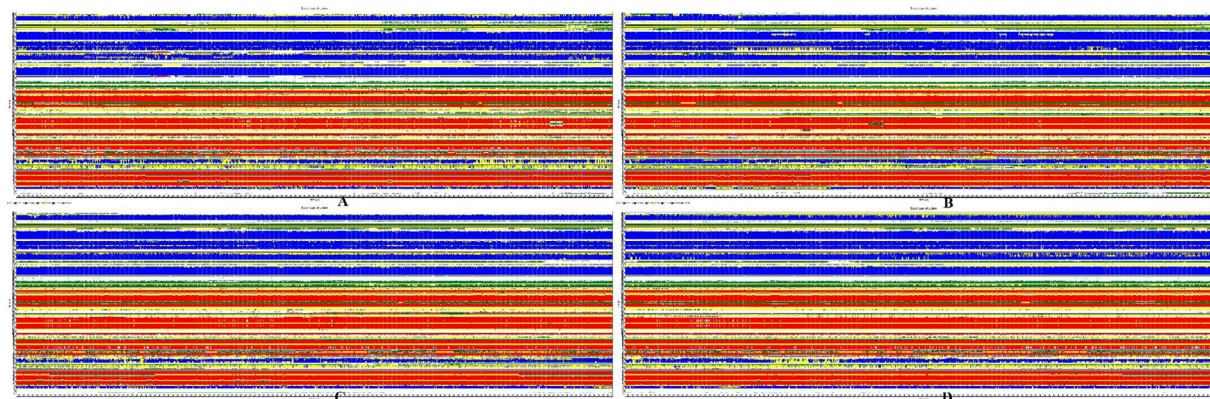
### 3.6. Gibbs free energy estimation

Although the calculation of the binding energies of protein-ligand complexes enjoys considerable accuracy, using the CDOCKER molecular docking algorithm, more accurate calculations may be useful in achieving more reliable results due to the limitations of this technique [63]. Furthermore, since the MD induces dynamic and conformational changes in the structure of the study complexes, the measurement of protein-ligand binding energies at the end of the MD and its comparison with the results of molecular docking form the basis for comparison, as accurately as possible on the assessment of the findings [64]. Hence, 500 final frames of each MD complex were used in this study, using the MMPBSA method, to calculate the binding free energy, as shown in Table 7, the binding energy for the ACE2-ZINC14976187 complex was  $-8.202 \pm 26.78$ , while the ZINC12562757 with a binding energy of  $-63.002 \pm 27.25$  has interacted with the active site of ACE2.

The Cinanserin compound had binding energy of  $-79.65 \pm 8.75$  for binding to MPro, while it was  $-92.66 \pm 13.02$  and  $-159.83 \pm 96.42$ , respectively, for compounds 112,260,215 and Evo\_1. The results confirm the docking data for ACE2, and indicate a stronger binding strength of Evo\_1 to MPro than other compounds concerning the MPro, as seen in docking.



**Fig. 14.** The secondary structure as a function of the simulation time for (A) free ACE2, (B) ACE2- ZINC14976187 (Training Set Lowest Ki), and (C) ACE2-ZINC12562757 complexes.



**Fig. 15.** The secondary structure as a function of the simulation time for (A) free MPro, (B) MPro-Cinanserin, (C) MPro-112,260,215, and (D) MPro-Evo\_1 complexes.

**Table 7**

Binding free energy results of ACE2 and MPro complexes using MM-PBSA calculations. All energies in kJ mol<sup>-1</sup>. SASA Solvent accessible surface area.

	ACE2-ZINC14976187	ACE2-ZINC12562757	MPro-Cinanserin	MPro-112,260,215	MPro-Evo_1
van der Waal energy	-0.001 ± 0.000 KJ/mol	-178.430 ± 11.073	-129.295 ± 8.287	-216.432 ± 13.238	-140.280 ± 74.119
Electrostatic energy	-0.510 ± 0.248	-57.183 ± 7.739	-18.660 ± 5.469	-82.249 ± 13.746	-198.036 ± 90.239
Polar solvation energy	-8.034 ± 26.723	199.815 ± 27.036	83.137 ± 7.448	228.615 ± 11.642	193.749 ± 75.883
SASA energy	0.342 ± 1.939	-15.527 ± 0.958	-14.832 ± 0.813	-22.600 ± 0.849	-15.263 ± 8.151
Binding energy	-8.202 ± 26.784	-63.002 ± 27.256	-79.650 ± 8.757	-92.668 ± 13.025	-159.830 ± 96.427

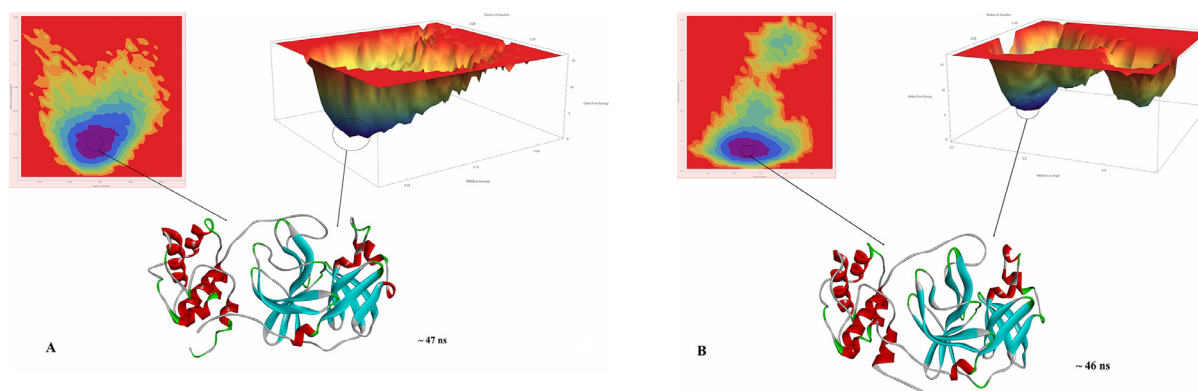
### 3.7. FEL analysis

Since the more targeted inhibition of SARS-CoV-2, by the prohibition of its specific enzyme (MPro), compared to other virus drug targeting, appears to be a logical approach, thus the MPro has been considered as the target for resumption of the computations. Also, as the MD and MMPBSA techniques confirmed the molecular docking results, and in addition to the fact that the two compounds 112,260,215 and Evo\_1 showed the best results in binding to MPro,

the two complexes MPro-112,260,215 and MPro-Evo\_1 were selected for FEL calculations and then QM-MM analysis.

As shown in Fig. 16, the dark blue sections in the 2D and 3D diagrams in Figures A and B are MPro conformations, which had the lowest Gibbs free energy levels during the simulation time. Since QM-MM calculations are very accurate, we decided to consider the most stable conformations as the inputs for QM-MM calculations for the two complexes. According to the FEL calculations, the frames 47,060 ps for Evo\_1 (Fig. 14A), and 46,366 ps for





**Fig. 16.** The 2D and 3D free energy landscape of (A) MPro-Evo\_1 and (B) MPro-112,260,215 complexes depicted as a function of the radius of gyration and RMSD, and representative structure with minimal energy.

**Table 8**

The HOMO (eV), LUMO (eV), Energy gap, and binding energies (kJ/mol) of 112,260,215 and Evo\_1 in complex with MPro.

	HOMO	LUMO	Energy Gap	Binding Energy
MPro-112,260,215	-0.1183	0.0541	0.1724	-1666.81
MPro-Evo_1	-0.1135	0.0419	0.1554	-2275.51

112,260,215 (Fig. 14B), respectively, show these two representative conformations. Subsequently, their PDB structures were extracted, using the GROMACS 2018.4 gmx trjconv tool and applied in the next step.

### 3.8. QM-MM studies

MD's most stable output snapshots, identified via the FEL analysis, were applied in QM-MM. The QM-MM hybrid method was used to more accurately examine the interactions between 112,260,215 and Evo\_1 ligands and the active site residues of the MPro [65]. Molecular frontier orbitals, including HOMO and LUMO, were calculated for each ligand. Also, the binding energy of protein-ligand was calculated after the complete optimization of the complexes. Computational studies have shown that the use of frontier molecular orbitals, including HOMO and LUMO, is useful for describing molecule's biological activity and structural properties [66, 67]. HOMO and LUMO play an important role in the transfer of charge in chemical reactions.

The higher the molecule HOMO energy, the stronger the nucleophiles will be, and the lower the LUMO energy, makes the molecule a stronger electrophile [68]. As shown in Table 8, the gap energy of 112,260,215 is 0.17, while the gap energy of Evo\_1 is 0.15. The smaller the energy gap, the more likely it is that the conformation will change to create more interactions, and thus the molecule will become more stable [69]. Moreover, it was observed that the binding energy of the MPro-112,260,215 complex is -1666.81, while that of the MPro-Evo\_1 is -2275.51. The results show therefore greater stability of the complex MPro-Evo\_1 and higher power of Evo\_1, compared to the compound 112,260,215, in stabilizing the conformation of the active site residues of MPro, and finally, more inhibitory power of Evo\_1.

## 4. Conclusion

The world's high rate of outbreaks and deaths caused by SARS-CoV-2 infection requires that an effective drug be produced as soon as possible, and embedded into treatment protocols around the world. In this study, we found a potential drug for inhibiting or reducing SARS-CoV-2 activity, using 3D-QSAR pharmacophore modeling, virtual screening, molecular docking, *de novo* evolution

design, MD simulation, MMPBSA, FEL, and QM-MM techniques. These *in silico* techniques are capable of identifying lead compounds that can inhibit human ACE2 as the virus entry pathway, and the virus's MPro as the specific enzyme of the SARS-CoV-2. Ten pharmacophore models were generated, using several well-known inhibitors of the ACE2. Hypo1, which was selected as the best model based on the criteria for evaluation and validation of pharmacophores, was used to screen over 1.2 million ligands. Following the virtual screening and calculation of the ADMET, Lipinski, and Veber descriptors, 81 compounds were identified and docked with the active site of ACE2. About 73 compounds with higher docking energy than the ACE2 standard inhibitor were then docked with the active site of MPro. Twenty-nine compounds were able to inhibit both targets together.

The compounds ZINC12562757 and 112,260,215 were selected as the best potential inhibitors of the ACE2 and MPro, respectively. Then, 10 *de novo* compounds were designed based on the structure of the best MPro inhibitor (112,260,215), and among them, Evo\_1 was selected with the highest docking energy of all MPro inhibitors. The ACE2-ZINC12562757, ACE2-ZINC14976187, MPro-112,260,215, MPro-Cinanserin, and MPro-Evo\_1 complexes were simulated for 50 ns. Results of RMSD, RMSF, H-bond, and DSSP analyzes have shown that among ACE2 and MPro inhibitors, ZINC12562757 and Evo\_1 have established the highest structural stability and hydrogen bond formation, and the least changes in the related enzyme active site key residues secondary structure, respectively. Furthermore, the results of MMPBSA indicate strong binding energies between ligands and enzymes further confirmed our docking results. Finally, the most stable conformation, resulting from the simulation of the two MPro-112,260,215 and MPro-Evo\_1 complexes, was extracted, using FEL analysis and submitted in QM-MM. The results from the QM-MM show a lower energy gap and binding energy than the Evo\_1, compared to the compound 112,260,215, which indicates that the Evo\_1 has more potential to specifically inhibit the MPro.

A significant number of *in silico* studies have found potential drugs to inhibit SARS-CoV-2, but our study has shown a novel insight into finding potential drugs with respect to the application of 3D-QSAR pharmacophore modeling, *de novo* evolution ligand design, free energy landscape, and QM-MM. The results of this study can be served as a useful strat-

egy for producing effective remedies for the treatment of COVID-19.

### Declaration of Competing Interest

None declared.

### CRediT authorship contribution statement

**Vahid Zarezade:** Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing, Project administration. **Hamzeh Rezaei:** Writing - original draft. **Ghodratollah Shakerinezhad:** Resources. **Arman Safavi:** Writing - original draft. **Zahra Nazari:** Writing - original draft. **Ali Veisi:** Conceptualization. **Omid Azadbakht:** Methodology, Software, Writing - original draft. **Mahdi Hatami:** Conceptualization. **Mohamad Sabaghan:** Resources. **Zeinab Shajirat:** Resources.

### Acknowledgement

We thank the Behbahan Faculty of Medical Sciences for financial support (grant number: 99003).

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.molstruc.2021.130409](https://doi.org/10.1016/j.molstruc.2021.130409).

### References

- N. Kandel, et al., Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries, *Lancet North Am. Ed.* 395 (10229) (2020) 1047–1053.
- F. Wu, et al., A new coronavirus associated with human respiratory disease in China, *Nature* 579 (7798) (2020) 265–269.
- Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: genome structure, replication, and pathogenesis, *J. Med. Virol.* 92 (4) (2020) 418–423.
- A.J. Turner, J.A. Hiscox, N.M. Hooper, ACE2: from vasopeptidase to SARS virus receptor, *Trends Pharmacol. Sci.* 25 (6) (2004) 291–294.
- W. Wang, et al., Angiotensin-Converting Enzyme 2 Metabolizes and Partially Inactivates Pyr-Apelin-13 and Apelin-17, *Hypertension* 68 (2) (2016) 365–377.
- D. Wrapp, et al., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, *Science* 367 (6483) (2020) 1260–1263.
- H. Zhang, et al., Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target, *Intensive Care Med* 46 (4) (2020) 586–590.
- F. Li, et al., Conformational States of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein Ectodomain, *J. Virol.* 80 (14) (2006) 6794–6800.
- I. Glowacka, et al., Evidence that TMPRSS2 Activates the Severe Acute Respiratory Syndrome Coronavirus Spike Protein for Membrane Fusion and Reduces Viral Control by the Humoral Immune Response, *J. Virol.* 85 (9) (2011) 4122–4134.
- A. Heurich, et al., TMPRSS2 and ADAM17 Cleave ACE2 Differentially and Only Proteolysis by TMPRSS2 Augments Entry Driven by the Severe Acute Respiratory Syndrome Coronavirus Spike Protein, *J. Virol.* 88 (2) (2014) 1293–1307.
- A. Shulla, et al., A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry, *J. Virol.* 85 (2) (2011) 873–882.
- S.-r. Li, et al., Searching therapeutic strategy of new coronavirus pneumonia from angiotensin-converting enzyme 2: the target of COVID-19 and SARS-CoV, *Eur. J. Clin. Microbiol. Infect. Dis.* 39 (6) (2020) 1021–1026.
- F. Gao, et al., Prediction of proteinase cleavage sites in polyproteins of coronaviruses and its applications in analyzing SARS-CoV genomes, *FEBS Lett.* 553 (3) (2003) 451–456.
- Z. Jin, et al., Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, *Nature* 582 (7811) (2020) 289–293.
- L. Zhang, et al., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved  $\alpha$ -ketoamide inhibitors, *Science* 368 (6489) (2020) 409–412.
- Z. Xu, et al., 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) 921627 2020.01.27.
- M. Wang, et al., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) *in vitro*, *Cell Res.* 30 (3) (2020) 269–271.
- H. Lu, Drug treatment options for the 2019-new coronavirus (2019-nCoV), *Biosci Trends* 14 (1) (2020) 69–71.
- A.-T. Ton, et al., Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds, *Mol Inform* (2021) n/a(n/a).
- M. Prajapat, et al., Drug targets for corona virus: a systematic review, *Indian J Pharmacol* 52 (1) (2020) 56–65.
- S.K. Panda, et al., ACE-2-Derived Biomimetic Peptides for the Inhibition of Spike Protein of SARS-CoV-2, *J. Proteome Res.* 20 (2) (2021) 1296–1303.
- N.A. Dales, et al., Substrate-Based Design of the First Class of Angiotensin-Converting Enzyme-Related Carboxypeptidase (ACE2) Inhibitors, *J. Am. Chem. Soc.* 124 (40) (2002) 11852–11853.
- D.N. Deaton, et al., Thiol-based angiotensin-converting enzyme 2 inhibitors: P1 modifications for the exploration of the S1 subsite, *Bioorg. Med. Chem. Lett.* 18 (2) (2008) 732–737.
- A. Mores, et al., Development of Potent and Selective Phosphinic Peptide Inhibitors of Angiotensin-Converting Enzyme 2, *J. Med. Chem.* 51 (7) (2008) 2216–2226.
- H. Gao, Predicting tyrosinase inhibition by 3D QSAR pharmacophore models and designing potential tyrosinase inhibitors from Traditional Chinese medicine database, *Phytomedicine* 38 (2018) 145–157.
- M.K. Gilson, et al., BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology, *Nucleic Acids Res.* 44 (D1) (2015) D1045–D1053.
- A.K. Debnath, Pharmacophore Mapping of a Series of 2,4-Diamino-5-deazapteridine Inhibitors of Mycobacterium avium Complex Dihydrofolate Reductase, *J. Med. Chem.* 45 (1) (2002) 41–53.
- S. Rampogu, et al., Targeting natural compounds against HER2 kinase domain as potential anticancer drugs applying pharmacophore based molecular modelling approaches, *Comput Biol Chem* 74 (2018) 327–338.
- S. Pal, et al., Ligand-based Pharmacophore Modeling, Virtual Screening and Molecular Docking Studies for Discovery of Potential Topoisomerase I Inhibitors, *Comput Struct Biotechnol J* 17 (2019) 291–310.
- C.L. Gupta, et al., Pharmacophore-based virtual screening approach for identification of potent natural modulatory compounds of human Toll-like receptor 7, *J. Biomol. Struct. Dyn.* 37 (18) (2019) 4721–4736.
- S. Kim, et al., PubChem 2019 update: improved access to chemical data, *Nucleic Acids Res.* 47 (D1) (2018) D1102–D1109.
- J.J. Irwin, B.K. Shoichet, ZINC – A Free Database of Commercially Available Compounds for Virtual Screening, *J Chem Inf Model* 45 (1) (2005) 177–182.
- C.A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution, *Drug Discovery Today: Technologies* 1 (4) (2004) 337–341.
- D. Lagorce, et al., FAF-Drugs3: a web server for compound property calculation and chemical library design, *Nucleic Acids Res.* 43 (W1) (2015) W200–W207.
- Y. Jiang, H. Gao, Pharmacophore-based drug design for the identification of novel butyrylcholinesterase inhibitors against Alzheimer's disease, *Phytomedicine* 54 (2019) 278–290.
- M.J. Abraham, et al., GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* (1–2) (2015) 19–25.
- J. Wang, et al., Automatic atom type and bond type perception in molecular mechanical calculations, *J. Mol. Graph. Model.* 25 (2) (2006) 247–260.
- J. Wang, et al., Development and testing of a general amber force field, *J Comput Chem* 25 (9) (2004) 1157–1174.
- A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *J Comput Chem* 23 (16) (2002) 1623–1641.
- M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, *J Appl Phys* 52 (12) (1981) 7182–7190.
- S. Miyamoto, P.A. Kollman, Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models, *J Comput Chem* 13 (8) (1992) 952–962.
- B. Hess, et al., LINCS: a linear constraint solver for molecular simulations, *J Comput Chem* 18 (12) (1997) 1463–1472.
- T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, *J Chem Phys* 98 (12) (1993) 10089–10092.
- S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin Drug Discov* 10 (5) (2015) 449–461.
- R. Kumari, R. Kumar, A. Lynn, g\_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations, *Journal of Chemical Information and Modeling*, 54 (7) (2014) 1951–1962.
- V. Zarezade, et al., In silico assessment of new progesterone receptor inhibitors using molecular dynamics: a new insight into breast cancer treatment, *J Mol Model* 24 (12) (2018) 337.
- A. Singh, P. Somvanshi, A. Grover, Pyrazinamide drug resistance in RpsA mutant ( $\Delta 438A$ ) of Mycobacterium tuberculosis: dynamics of essential motions and free-energy landscape analysis, *J. Cell. Biochem.* 120 (5) (2019) 7386–7402.
- S. Bharadwaj, et al., Structure-based screening and validation of bioactive compounds as Zika virus methyltransferase (MTase) inhibitors through first-principle density functional theory, classical molecular simulation and QM/MM affinity estimation, *J. Biomol. Struct. Dyn.* (2020) 1–14.
- Y. Ma, et al., 3D QSAR Pharmacophore Based Virtual Screening for Identification of Potential Inhibitors for CDC25B, *Comput Biol Chem* 73 (2018) 1–12.
- N. Baildya, N.N. Ghosh, A.P. Chattopadhyay, Inhibitory activity of hydroxychloroquine on COVID-19 main protease: an insight from MD-simulation studies, *J Mol Struct* 1219 (2020) 128595.
- R. Yu, et al., Computational screening of antagonists against the SARS-CoV-2 (COVID-19) coronavirus by molecular docking, *Int. J. Antimicrob. Agents* (2020) 106012.



- [52] M. Tahir ul Qamar, et al., Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants, *J Pharm Anal* (2020).
- [53] K.E. Hevener, et al., Validation of Molecular Docking Programs for Virtual Screening against Dihydropteroate Synthase, *J Chem Inf Model* 49 (2) (2009) 444–460.
- [54] G.A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, New York, 1997.
- [55] P. Towler, et al., ACE2 X-Ray Structures Reveal a Large Hinge-bending Motion Important for Inhibitor Binding and Catalysis, *J. Biol. Chem.* 279 (17) (2004) 17996–18007.
- [56] B. Havranek, S.M. Islam, An in silico approach for identification of novel inhibitors as potential therapeutics targeting COVID-19 main protease, *J. Biomol. Struct. Dyn.* (2020) 1–12.
- [57] S. Alagu Lakshmi, et al., Ethnomedicines of Indian origin for combating COVID-19 infection by hampering the viral replication: using structure-based drug discovery approach, *J. Biomol. Struct. Dyn.* (2020) 1–16.
- [58] A. Kumar, et al., Identification of phytochemical inhibitors against main protease of COVID-19 using molecular modeling approaches, *J. Biomol. Struct. Dyn.* (2020) 1–11.
- [59] L. Mittal, et al., Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach, *J. Biomol. Struct. Dyn.* (2020) 1–19.
- [60] D. Kumar, et al., Promising inhibitors of main protease of novel corona virus to prevent the spread of COVID-19 using docking and molecular dynamics simulation, *J. Biomol. Struct. Dyn.* (2020) 1–15.
- [61] D. Kumar, et al., Understanding the binding affinity of noscapienes with protease of SARS-CoV-2 for COVID-19 using MD simulations at different temperatures, *J. Biomol. Struct. Dyn.* (2020) 1–14.
- [62] Z.-Z. Wang, et al., Investigating the binding mechanism of (4-Cyanophenyl)glycine derivatives as reversible LSD1 by 3D-QSAR, molecular docking and molecular dynamics simulations, *J Mol Struct* 1175 (2019) 698–707.
- [63] E. Wang, et al., End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: strategies and Applications in Drug Design, *Chem. Rev.* 119 (16) (2019) 9478–9508.
- [64] X. Song, et al., Inhibitory mechanism of epicatechin gallate on tyrosinase: inhibitory interaction, conformational change and computational simulation, *Food Funct* 11 (6) (2020) 4892–4902.
- [65] K. Cui, J.R. Schmidt, Enabling Efficient and Accurate Computational Studies of MOF Reactivity via QM/MM and QM/QM Methods, *The Journal of Physical Chemistry C* 124 (19) (2020) 10550–10560.
- [66] M. Arooj, et al., 3D QSAR Pharmacophore Modeling, in Silico Screening, and Density Functional Theory (DFT) Approaches for Identification of Human Chymase Inhibitors, *Int J Mol Sci* 12 (12) (2011) 9236–9264.
- [67] S. Rajamanikandan, J. Jeyakanthan, P. Srinivasan, Discovery of potent inhibitors targeting *Vibrio harveyi* LuxR through shape and e-pharmacophore based virtual screening and its biological evaluation, *Microb. Pathog.* 103 (2017) 40–56.
- [68] N. Ardalan, et al., Novel mutant of *Escherichia coli* asparaginase II to reduction of the glutaminase activity in treatment of acute lymphocytic leukemia by molecular dynamics simulations and QM-MM studies, *Med. Hypotheses* 112 (2018) 7–17.
- [69] S.S. Mohseni, et al., Identification of novel inhibitor against endonuclease subunit of Influenza pH1N1 polymerase: a combined molecular docking, molecular dynamics, MMPBSA, QMMM and ADME studies to combat influenza A viruses, *Comput Biol Chem* 77 (2018) 279–290.