

# Virtual Screening of Heterocyclic Molecules to Identify Potential SARS-COV2 virus M<sup>pro</sup> Protease Inhibitors for Further Medicinal Chemistry design

**Mansour Sobeh**

Mohammed VI Polytechnic University

**Reda Ben Mrid**

Mohammed VI Polytechnic University

**Abdelaziz Yasri** (✉ [aziz.yasri@um6p.ma](mailto:aziz.yasri@um6p.ma))

Mohammed VI Polytechnic University

---

## Research Article

**Keywords:** SARS-COV2 virus, Mpro protease, Heterocyclic Inhibitors, Molecular Docking

**Posted Date:** June 30th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-37557/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

In this work, we aimed to identify potential SARS-COV2 virus M<sup>Pro</sup> protease inhibitors using molecular docking. As a validation step of our docking approach, we used 7 known M<sup>Pro</sup> inhibitors and predicted their binding energies. We found a very good correlation ( $R = -0.86$ ) between the binding energies and the pIC<sub>50</sub> (-LogIC<sub>50</sub>) values. We then undertook virtual screening of 50753 heterocyclic molecules from the PubChem Database. The screened molecules were first filtered out using Lipinski, Veber and Ghose rules resulting in 21 142 which was submitted to the docking phase. The docked compounds were ranked according to their predicted binding energy to the protease and a threshold of -9.0 Kcal/mol was applied resulting in a set of 2711 hits. These hits were split into different groups according to their chemical class using a rule-based classification approach. The best compound from each of the most populated classes were subjected to binding mode analysis which led to ligand-receptor interaction maps and suggested some medicinal chemistry-oriented modifications to further optimize the potency of the obtained hits. Predicted IC<sub>50</sub> of the hit molecules ranged from 0.85 to 0.43  $\mu$ M against SARS-COV2 virus M<sup>Pro</sup> protease.

## 1. Introduction

Since the onset at the end of 2019 of infection with the SARS-COV2 virus to the present day (May 26<sup>th</sup>, 2020) [1, 2], the pandemic situation has resulted in 5,370,375 confirmed cases and 344,454 overall deaths in 195 countries and territories. In Europe, 2,025,176 confirmed cases have been reported including 173,454 death. The countries most affected by the pandemic are the United States, Brazil, Spain, Italy, United Kingdom, India and Germany [3]. The USA remains the country most affected by the coronavirus with nearly a third of Covid-19 contaminations and almost a quarter of the dead (99,805 deaths).

The aggressive nature of this viral infection has pushed the scientific community into a race against the clock to find remedies and solutions to control its spread around the world. One of the solutions implemented very quickly by some countries and belatedly by others is the containment of the population combined with the wearing of protective masks. These solutions have made it possible to reduce the number of new cases significantly in certain countries, but the risk of a new wave of contamination remained real. In addition to the health consequences, the pandemic has caused a slowdown in economic activity and stunted growth judged by economic indicators with an unprecedented decrease in industrial production. This impact could exceed \$4.1 Trillion [4].

Faced with this global health crisis, pharmaceutical players were working hard to develop therapeutic solutions (vaccines or drugs). Chloroquine was the first drug to be clinically tested against SARS-COV-2 [5]. The pioneering team in this therapeutic fight against the SARS-COV2 virus is that of a French researcher at Aix-Marseille Hospital in France. Their study showed a synergistic effect in vitro between chloroquine (primarily used to prevent and treat malaria) and azythromycin [6] (an antibiotic used for the treatment of a number of bacterial infections) against the SARS-COV2 virus in concentrations compatible with those obtained in the human lung [7]. Other existing drugs were tested against the virus such as Remdesivir and proved efficacy in treating patients carrying COVID-19 disease [8].

Target based drug discovery approach combined to molecular modeling and medicinal chemistry proved to be a highly efficient strategy to quickly develop therapeutics against human diseases [9–11]. With the genomic characterization of the SARS-COV2 virus and the elucidation of the 3D structures of certain of its proteins such as the main protease [12, 13], a door opened for the use of the target-based approach drug discovery to identify therapeutic agents against this virus [14]. Recent works using computational approaches such as structure-based drug design, molecular modeling and virtual screening assisted in the identification of potential inhibitors from different sources (existing drugs, natural compounds, chemical synthesis) against some key targets of SARS-COV2 virus [15, 16].

In this work, following target-based drug discovery approach, we undertook the virtual screening of heterocyclic compounds against the main protease of the SARS-COV2 virus. The potential inhibitors were grouped into chemical classes and each class was characterized by its mode of binding in the active site of the protease. This work opens interesting perspectives for medicinal chemistry combined with the structured-based drug design to optimize the activity of the selected hits from this study and to synthesize and evaluate the activity of new anti-SARS-COV2 main protease inhibitors.

## 2. Materials And Methods

*2.1. Chemical drawing and classification.* All molecular and Markush structures were designed by MarvinSketch [35]. The chemical classification of the molecules was performed using ClassyFire engine [36].

*2.2 Ligand Data preparation.* Data on the inhibition of main protease M<sup>Pro</sup> of SARS-COV2 virus of 7 molecules was obtained from of the work of Jin, Z. et al. (2019) [13]. 50753 2D structures of heterocyclic molecules were download from the PubChem Database [37]. The salts forms of the molecules as well as the mixtures were processed and the largest fragment in each molecule was kept. The strong acids were deprotonated, the strong bases were protonated, and explicit hydrogen were added to the structures. The resulting set of molecules was subjected to sequential filtering steps including the Lipinski rules [38] (MW < = 500, logP < = 5, HBA < = 10; HBD < = 5), Veber's rules [39] (Rotatable Bonds ≤ 10; Polar surface Area ≤ 140 Å<sup>2</sup>) and Ghose rules [40] (40 < = Molar refractivity < = 130; 20 < = Atom Count < = 70). The filtered set of heterocyclic molecules as well as the known inhibitors were converted to 3D structure (when keeping the original stereochemistry information) with an energy minimization step using openBabel Software [41].

*2.3. Molecular Docking process.* The 3D structure coordinates of the SARS-COV2 main protases M<sup>Pro</sup> (PDB code: 5RGI) was obtained from the Protein Data Bank [42, 43]. The molecular docking was performed by AutoDock Vina package [44]. AutoDock MGTTOOLS program was used to prepare the active site of the receptor [44]. All the water molecules as well the ligand structure co-crystallized with the protein were removed. Nonpolar hydrogens were added to the protein and a grid box was constructed around the protease active with a size of 20 Å, 16 Å et 20 Å in X, Y and Z directions with a central point defined at 9, 0,1 and 21 three-dimensional coordinates. The 3D structure of the protease was then saved into a PDBQT

file format. All the ligand 3D structures were prepared for docking using a python script provided with AuDock Vina package and which produced the PDBQT files needed for the docking step. All the other AutoDock Vina parameters were kept at their default values.

*2.4. Visualization.* This step was performed using PyMol software [45]. AutoDock Vina generates 9 poses for each docked ligand. An inhouse python script was used to extract the pose with the best score which expresses the binding energy of the ligand in Kcal/mol as unit. The generation and visualization of 2D ligand-receptor interactions maps was performed by Discovery Studio Visualizer [46].

## 3. Results

### 3.1. Prediction of binding affinities of known inhibitors

As a validation step of the docking approach, known inhibitors were docked into the active site of M<sup>Pro</sup> protease. This step concerned 7 inhibitors taken from the work of Jin, Z. et al. (2019) [13] (Table 1). Carmofur, anti-cancer agent, is a pyrimidine derivatives obtained from fluorouracil (chemotherapy agent) that can be administered orally [17]. Ebselen, a benzoisoselenazolone derivative, is a non-toxic seleno-organic drug with antiinflammatory, antiatherosclerotic and cytoprotective properties [18, 19]. Tideglusib, a thiodiazole derivative, is a potent, selective and irreversible inhibitor of glycogen synthase kinase 3 (GSK-3) [20]. TDZD-8, another thiodiazole derivative is also an inhibitor of GSK3 kinase [21]. PX-12, a sulfides derivative, is an irreversible inhibitor of the redox protein thioredoxin [22], which has been associated with cancer and tumor growth. Disulfiram is a carbamate derivative used as an alcohol deterrent and inhibitor of aldehyde dehydrogenase [23]. Shikonin, a quinone derivative, is an apoptotic agent active in a variety of cancer cell lines but not in normal cell lines [24].

**Due to technical limitations, Table 1 is provided in the Supplementary Files section.**

Molecular Docking of these known inhibitors showed docking scores ranging from -7.9 Kcal/mol to -1.7 Kcal/mol. When comparing the pIC<sub>50</sub> (-Log IC<sub>50</sub>) values and the docking scores, we obtained a good anti-correlation coefficient (R) of -0.86 and a regression determination coefficient R<sup>2</sup> of 0.74 (Figure 1).

### 3.2. Drug likeness of ligand molecules

Before molecular docking step, the set of heterocyclic compounds were subjected to a filtering step using Lipinski's rule of five. This first filtering kept 34783 out of 50753 compounds. The application of the Veber rules resulted into 32669 compounds and the Ghose rules reduced the set of compounds to 21 142. This final set of compounds was then submitted to the docking phase into the M<sup>Pro</sup> protease active site.

### 3.3. Virtual screening of heterocyclic compounds

The filtered set of heterocyclic compounds were virtually screened within the M<sup>Pro</sup> protease active site prepared using AutoDock Vina. The program predicts the best poses that a ligand may have and its binding energy (docking score) to the receptor. The docked compounds were then ranked according to their score of binding energies. In order to look for compounds potentially more active than Ebselen (Score = -7.9 Kcal/mol, pC<sub>50</sub> = 6.17) and Tideglusib (Score = -7.9 Kcal/mol, pIC<sub>50</sub> = 5.81), we set a threshold for the docking binding energy at -9.0 Kcal/Mol. Using this threshold, we selected 2711 compounds whose binding energy was less or equal to -9.0 kcal/mol (Supplementary Table S1). This set is referred to as HitSet.

The distributions of the four Lipinski's rules with the selected set of compounds were quite symmetrical except for the hydrogen bonds donor property where the distribution was skewed at the left. In fact, 80% of the molecules have less or equal to two hydrogen bond donors (Figure 2).

### 3.3.1. Compounds classification

The compounds of the HitSet were submitted to the classification engine Classyfire. This web application uses a rule-based classification approach that relies on a comprehensible, comprehensive, and computable chemical taxonomy. As a result, we obtained 162 different chemical classes (Supplementary Materials Table S1). Nine chemical classes had more than 70 compounds: Benzene and substituted derivatives, quinolines and derivatives, indoles and derivatives, benzopyrans, diazanaphthalenes, diazinanes, carboxylic acids and derivatives, azoles and organo-oxygens. We have analyzed the members of each chemical classes and selected few representatives with the best docking scores for analyzing their binding mode inside the M<sup>Pro</sup> protease active site. We have also analyzed some exceptional natural molecules that were not within the HitSet due to the noncompliance with Ghose rules but presented interesting predicted binding energies

## 4. Discussion

### 4.1. Prediction of binding energy of known inhibitors

In the first step of the docking procedure, we have docked the 7 known inhibitors into the M<sup>Pro</sup> active site protease. The existence of the good correlation ( $R = -0,86$ ) between the docking scores and the IC<sub>50</sub> values of the known inhibitors suggested that our docking could be used to predict the binding mode of potential inhibitors and their IC<sub>50</sub> for M<sup>Pro</sup> protease. The visual inspection of the most active inhibitor ebselen (IC<sub>50</sub> = 0,26  $\mu$ M) docked into the protease active site revealed that this compound established  $\pi$ -anion interaction with Glu-166,  $\pi$ -sigma interaction with Met-165, van der Waals (VDW) contacts with His-41 and His-163 and weak hydrogen bond (HB) contact with His-163 (Figure 3).

### 4.2. Prediction of binding energy of Heterocyclic compounds

From each highly populated chemical class obtained from the HiTSet, we analyzed the lead compound showing the highest docking score. For these compounds, we have elucidated and discussed their binding mode into M<sup>Pro</sup> protease active site and predicted their binding energy as well as their IC<sub>50</sub> (Table 2).

## 4.2.1 Benzene and substituted derivatives

This chemical class is the largest in terms of number of compounds. The most active compound holds a fluorobenzene moiety and presented a docking score of  $-10,6$  Kcal/mol which corresponded to a predicted IC<sub>50</sub> of  $0.6$   $\mu$ M derived from the simple linear regression we have obtained before. This compound, referred to as HBF-0259, was reported as an inhibitor of hepatitis B virus surface antigen (HBsAg) secretion [25]. The visual inspection of its binding mode to M<sup>Pro</sup> protease revealed multiple contacts to the residues of the active site of types, HB (Leu-141, Gly-143, Ser-144, Cys-145), VDW (Met-49, His-163), and  $\pi$ -sigma (Leu-27, His-41, Met-165) (Figure 4-a). Further optimization of the potency of this compound could be carried out through diverse substitutions on the phenyl groups attached to the pyrimidine core. The substitutions could target new interactions with other amino acids in the active site for enhanced affinity to the protease.

## 4.2.2 Quinolines and derivatives

The best compound from this class, a triazine derivative, resulted in a docking score of  $-10.8$  Kcal/mol which corresponds to a predicted IC<sub>50</sub> of  $0.57$   $\mu$ M. This compound, named Capmatinib, is a potent inhibitor of c-Met kinase with antineoplastic activity [26]. It showed dense network of contacts within the M<sup>Pro</sup> protease active site including HB (Ser-144, Thr-25), VDW (Cys-44, Gln-189, Ser-46, His-172, His-163, Leu-141, Leu-27),  $\pi$ -Alkyl (Met-49, Cys-145, Thr-26) and  $\pi$ - $\pi$  T-shaped to His-41 (Figure 4-b). Different potency optimization possibilities are available on the quinoline moiety as well as on the Fluro-phenyl moiety.

**Table 2.** Predicted most active compounds from the HitSet chemical classes.

Compound	Chemical Class	Direct Parent Class	Predicted Binding Energy (Kcal/mol)	Stdev Binding Energy (Kcal/mol)	Predicted IC <sub>50</sub> (μM)
HBF-0259	Benzene and substituted derivatives	Fluorobenzenes	-10,60	1,10	0.60
Capmatinib	Quinolines and derivatives	Quinolines and derivatives	-10.80	1.17	0.57
Arcyroxocin B	Indoles and derivatives	Hydroxyindoles	-10.80	0.72	0.57
DR396	Benzopyrans	Xanthenes	-11.30	0.52	0.51
Pazinaclone	Carboxylic acids and derivatives	Beta amino acids and derivatives	-11.50	1.03	0.49
Fumiquinazoline C	Diazanaphthalenes	Quinazolines	-11.60	0.95	0.48
57	Diazinanes	N-arylpiperazines	-10.30	0.67	0.64
SR-3029	Azoles	Phenylimidazoles	-11.20	0.76	0.52
AMG-579	Organooxygen compounds	Aryl-phenylketones	-10.50	0.49	0.61
Silydianin	Flavonoids	Flavones	-11.1	0.69	0.53

### 4.2.3. Indoles and derivatives

In this chemical class, the best two compounds (Score = -10.9 Kcal/mol) are straurosporine like compounds which are potent non-specific inhibitors of kinases but present high toxicity. For this reason, we choose to analyse the binding mode of the third best compound presenting a docking score of -10.8 Kcal/mol (predicted IC<sub>50</sub> = 0.57 μM) and which is referred to as Arcyroxocin B. This alkaloid was isolated from two fungal species *Arcyria denudata* and *Arcyria obvelata* and showed cytotoxicity against Jurkat cells [27], an immortalized line of human T lymphocyte cells that are used to study acute T cell leukemia. When docked into the active site of M<sup>PrO</sup> protease, Arcyroxocin B atoms established multiple HB contacts (Cys-145, Gly-143, Ser-144, Leu-141, His-163, Glu-166), VDW interactions (His-41, Phe-140, Met-165), and π-Alkyl interaction with Cys-145 (Figure 4-c). The potency of this compound can be further optimized through substitutions on the phenyl groups of the two indole moieties.

### 4.2.4. Benzopyrans

Two compounds from this chemical class, fluorescein 5-maleimide and DR396 a benzoic acid derivative compound, presented docking scores of  $-11.3$  Kcal/mol corresponding to predicted  $IC_{50}$  of  $0.5$   $\mu$ M. We choose to study the binding mode of DR396 since it has a reported biological activity as an apoptotic DNase  $\gamma$  inhibitor [28]. In fact, DR396 showed several interactions with the protein amino acids of type HB (Thr-190, Gln-192), Halogen (Glu-166, Phe-140, Gly-143, Asn-142), VDW (Arg-188, Leu-167, Ser-144, Cys-145, His-41, Met-49, Gln-189), and  $\pi$ -interactions (Met-165) (Figure 4-d). Chemical diversifications for optimizing the potency of this compounds could take place at different positions of the molecule on the benzopyran core as well as on the phenyl group of the benzoic acid.

## 4.2.5. Carboxylic acids and derivatives

In this chemical class of selected hits, the best compound showed a docking score of  $-11.5$  Kcal/mol and a predicted  $IC_{50}$  of  $0,48$   $\mu$ M. This compound referred to as Pazinaclone (DN-2327), is an anxiolytic drug [29]. The evaluation of the binding mode of pazinaclone in the  $M^{pro}$  protease active site showed that it is interacting with several protein amino acids through HB (Gly-143, Cys-145, Ser-144, Asn-145, Glu-166), VDW (Leu-27, Leu-141, His-163, His-164), Alkyl interaction with Met-165, and  $\pi$ -interactions with His-41 (Figure 4-e). For optimization purposes, chemistry could take place with diverse substitutions on the naphthyridine group.

## 4.2.6. Diazanaphthalenes

Two compounds from this chemical class, fumiquinazoline C and H, presented a docking score of  $-11,6$  Kcal/mol corresponding to predicted  $IC_{50}$  of  $0.47$   $\mu$ M. These two molecules are produced by *Aspergillus fumigatus*, a marine fungus that makes a series of fumiquinazoline peptidyl alkaloids [30]. As the compounds are very similar, we reported only the binding of fumiquinazoline C whose derivatives were reported to have antitumor activity [38]. Fumiquinazoline C predicted binding mode revealed that the atoms of this compounds are in close contact with different protein residues through HB (His-41, His-163, Glu-166, Cys-145), VDW interactions (Leu-141, His-172, His-163, His-164, Met-165), and  $\pi$ -interactions ( $\pi$ -Alkyl with Pro-168 and Cys-145,  $\pi$ - $\pi$  T-Shaped with His-41) (Figure 4-f).

## 4.2.7. Diazinanes

The best compound in terms of docking score from this class shows predicted values for binding energy of  $-10.3$  Kcal/mol and an  $IC_{50}$  of  $0.63$   $\mu$ M. It was reported to have an antagonist activity against 5-HT receptor [31]. Compound 57 presented a high density of interactions within the  $M^{pro}$  protease active site. In fact, it was able to make HB contacts to four amino acids (Gly-143, His-163, Leu-141, Ser-144), VDW interactions four amino acids (Met-49, Phe-140, His-164, Met-165), and  $\pi$ -interactions to three amino acids (Glu-166, His-41, Cys-145) (Figure 4-g). In terms of possibilities of chemistry in order to

optimize its potency, substitutions are possible in the piperazine group, as well as the naphthalene group and the indazole core.

## 4.2.8. Azoles

The azoles chemical class showed compound SR-3029 with the highest docking score of  $-11.2$  Kcal/mol and a predicted  $IC_{50}$  of  $0.52 \mu\text{M}$ . SR-3029 was reported to be highly selective casein kinase 1 Delta /1 Epsilon inhibitor with potent antiproliferative properties [32]. The docked pose of SR-3029 in the  $M^{\text{pro}}$  protease active site showed some interactions of types fluorin-HB (Cys-44, Glu-166, Phe-140), multiples VDW interactions (Met-49, His-41, Leu-27, Thr-25, Thr-45, His-172), and  $\pi$ -interactions (Cys-145, Gly-143, Leu-141, Ser-144) (Figure 4-h). Possible structural diversifications could be done on fluoro-phenyl group or the non-substituted positions on the benzodiazole core.

## 4.2.9. Organo-oxygens

The best compound of this chemical class is AMG-579 which showed a docking score of  $-10.5$  Kcal/mol and a predicted  $IC_{50}$  of  $0.6 \mu\text{M}$ . AMG-579 is a potent, selective, and efficacious inhibitor of phosphodiesterase 10A [33]. AMG-579 pose in the  $M^{\text{pro}}$  active site showed different interactions with the protein amino acids: HB (Thr-26, Ser-144), VDW (Leu-27, His-163, Leu-141, Asn-142),  $\pi$ -interactions ( $\pi$ -Alkyl to Pro-168 and Cys-145), and alkyl interaction with Cys-145 (Figure 4-i). This molecule can be further optimized for its potency by chemical modifications on the amid moiety, the phenyl group from the benzodiazole and the central phenyl group.

## 4.2.10. Exceptional chemical class

Few compounds showed good docking scores but did not obey the different filtering rules such as Lipinski's rules. We choose to select the best of these compounds to study its binding mode to  $M^{\text{pro}}$  protease. One of these compounds is silydianin which showed a docking score of  $-11.1$  Kcal/mol and a predicted  $IC_{50}$  of  $0.53 \mu\text{M}$ . Silydianin, an isomer of Silymarin extracted from the plant *Silybium marimum*, can control the main symptoms of asthma through modulating immune system responses [34].

Silydianin pose in the  $M^{\text{pro}}$  active site showed different interactions with the protein amino acids: HB (Thr-190, Gln-192, His-164, Ser-144, Leu-141, Phe-140, Glu-166), VDW (Met-49, His-163, Cys-145, His-41, Arg-188),  $\pi$ -interactions ( $\pi$ -Sigma to Gln-189,  $\pi$ -Alkyl to Pro-168), and alkyl interaction with Cys-168 (Figure 4-j).

Overall, and for all the analyzed hits from the different chemical classes, it appeared that some amino acids within the  $M^{\text{pro}}$  active site were the key residues entering in close interactions with the different ligands. In fact, four amino acids were involved in HB contacts with ligands very frequently. These amino acids concerned Ser-144, Cys-145, Glu-166, Gly-143. They were involved in HB interactions with at

least 4 ligands. Four amino acids established VDW interaction with at least 4 analyzed ligands (His-163, Leu-141, Leu-27, Met-49). These VDW interactions are the predominant type of interactions that the majority of ligands have established with the protein. Three amino acids showed  $\pi$  interactions with at least 3 ligands (Cys-145, His-41, Met-165). These amino acids could be targeted during the design a new of optimized M<sup>Pro</sup> inhibitors.

As we have considered in our binding mode analysis only one representative from each chemical, other compounds from the same class could investigated also for experimental inhibition screening against the M<sup>Pro</sup> protease. This will provide interesting data to generate structure-activity relationships maps to be used for optimizing the activity of the molecules of interest.

## 5. Conclusions

In this work, a virtual screening strategy was successfully used to predict the binding energies and the IC<sub>50</sub> of known M<sup>Pro</sup> protease inhibitors. It was also applied to find potential inhibitors from heterocyclic compounds to this viral protease and to identify their interaction maps with the protease active site. To the best of our knowledge, this is the first report using this approach combined to chemical analysis of the ligands to identify key interactions between the ligand and the M<sup>Pro</sup> protease and deliver guidelines of potential medicinal chemistry-oriented modifications of the compounds in order to optimize their potency.

The different hits from this study were categorized in chemical classes and the most active compounds from these classes showed predicted IC<sub>50</sub> ranging from 0.85 to 0.43  $\mu$ M. This set of potential M<sup>Pro</sup> protease inhibitors came from different sources either natural compounds such capmatinib, arcyroxocin B, and fumiquinazoline C or from organic chemistry synthesis such as pazinaclone, compound 57, SR-3029, AMG-579, HBF-0259, DR396. Most of these identified compounds had various reported biological activities. This set of compounds and their homologous ones in the chemical series will be good candidates for experimental screening against SARS-COV2 M<sup>Pro</sup> main protease. The screening of the analogues of the hit compounds will certainly provide data to generate structure-activity relationships maps to be used for medicinal chemistry investigations.

## Declarations

*Availability of data and materials:* The following are available online at [https://1drv.ms/x/s!AnTvRYBsySEg2Vlq4T\\_77P9\\_CR0?e=8qqVyt](https://1drv.ms/x/s!AnTvRYBsySEg2Vlq4T_77P9_CR0?e=8qqVyt), *Table S1:* List of selected compounds from molecular docking with their PubChem ID, chemical class, direct parent class, predicted binding energy and its standard deviation and predicted IC<sub>50</sub> for SARS-COV2 virus M<sup>Pro</sup> protease.

*Competing Interest:* The authors declare no competing interest.

*Funding:* This work did not receive any funding.

*Author Contributions:* Conceptualization, A. Y.; methodology, M. S., R. B., A. Y.; validation, M. S., A. Y.; writing—original draft preparation, A. Y., R. B., M. S.; writing—review and editing, M. S., R. B., A. Y.; visualization. All authors have read and agreed to the published version of the manuscript.

## References

1. Zhou P, Yang X-L, Wang X-G, et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579:270–273. <https://doi.org/10.1038/s41586-020-2012-7>
2. Wu F, Zhao S, Yu B, et al (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579:265–269. <https://doi.org/10.1038/s41586-020-2008-3>
3. (2020) Coronavirus (COVID-19). <https://covid19.who.int/>
4. Coronavirus Pandemic Could Cost the World \$4.1 Trillion. <https://time.com/5814933/coronavirus-pandemic-cost-4-trillion/>
5. Cortegiani A, Ingoglia G, Ippolito M, et al (2020) A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19. *J Crit Care*. <https://doi.org/10.1016/j.jcrc.2020.03.005>
6. Greenwood, David (2008) Antimicrobial drugs: chronicle of a twentieth century medical triumph (1. publ. ed.). In: Oxford: Oxford University Press. p 239
7. Andreani J, Le Bideau M, Duflot I, et al (2020) In vitro testing of combined hydroxychloroquine and azithromycin on SARS-CoV-2 shows synergistic effect. *Microb Pathog* 145:104228. <https://doi.org/10.1016/j.micpath.2020.104228>
8. Wang Y, Zhou F, Zhang D, et al (2020) Evaluation of the efficacy and safety of intravenous remdesivir in adult patients with severe COVID-19: study protocol for a phase 3 randomized, double-blind, placebo-controlled, multicentre trial. *Trials* 21:422. <https://doi.org/10.1186/s13063-020-04352-9>
9. Croston GE (2017) The utility of target-based discovery. *Expert Opinion on Drug Discovery* 12:427–429. <https://doi.org/10.1080/17460441.2017.1308351>
10. Cheve G, Dayde-cazals B, Fauvel B, Bories C, Yasri A. Azaindole derivatives as inhibitors of protein kinases. US Patent 9,493,455
11. Lakhilili W, Chevé G, Yasri A, Ibrahim A (2015) Determination and validation of mTOR kinase-domain 3D structure by homology modeling. *Onco Targets Ther* 8:1923–1930. <https://doi.org/10.2147/OTT.S84200>
12. Anand K, Ziebuhr J, Wadhwani P, et al (2003) Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* 300:1763–1767. <https://doi.org/10.1126/science.1085658>
13. Jin Z, Du X, Xu Y, et al (2020) Structure of M(pro) from COVID-19 virus and discovery of its inhibitors. *Nature*. <https://doi.org/10.1038/s41586-020-2223-y>
14. Zhang L, Lin D, Kusov Y, et al (2020) alpha-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J Med Chem*. <https://doi.org/10.1021/acs.jmedchem.9b01828>

15. Macchiagodena M, Pagliai M, Procacci P (2020) Identification of Potential Binders of the Main Protease 3CL(pro) of the COVID-19 via Structure-Based Ligand Design and Molecular Modeling. *Chem Phys Lett* 750:137489. <https://doi.org/10.1016/j.cplett.2020.137489>
16. Quiros Roldan E, Biasiotto G, Magro P, Zanella I (2020) The possible mechanisms of action of 4-aminoquinolines (chloroquine/hydroxychloroquine) against Sars-Cov-2 infection (COVID-19): A role for iron homeostasis? *Pharmacol Res* 158:104904. <https://doi.org/10.1016/j.phrs.2020.104904>
17. Noda T, Kosakai H, Tsujimura K, et al (1983) [Clinical trial of carmofur (HCFU) in the treatment of malignant ovarian cancer–The first report: combination therapy with endoxan, mitomycin C, and toyomycin]. *Gan To Kagaku Ryoho* 10:1972–1979
18. Schewe T (1995) Molecular actions of ebselen—an antiinflammatory antioxidant. *Gen Pharmacol* 26:1153–1169. [https://doi.org/10.1016/0306-3623\(95\)00003-j](https://doi.org/10.1016/0306-3623(95)00003-j)
19. Muller A, Cadenas E, Graf P, Sies H (1984) A novel biologically active seleno-organic compound—I. Glutathione peroxidase-like activity in vitro and antioxidant capacity of PZ 51 (Ebselen). *Biochem Pharmacol* 33:3235–3239. [https://doi.org/10.1016/0006-2952\(84\)90083-2](https://doi.org/10.1016/0006-2952(84)90083-2)
20. Dominguez JM, Fuertes A, Orozco L, et al (2012) Evidence for irreversible inhibition of glycogen synthase kinase-3beta by tideglusib. *J Biol Chem* 287:893–904. <https://doi.org/10.1074/jbc.M111.306472>
21. Eldar-Finkelman H, Martinez A (2011) GSK-3 Inhibitors: Preclinical and Clinical Focus on CNS. *Front Mol Neurosci* 4:32. <https://doi.org/10.3389/fnmol.2011.00032>
22. Jordan BF, Runquist M, Raghunand N, et al (2005) The thioredoxin-1 inhibitor 1-methylpropyl 2-imidazolyl disulfide (PX-12) decreases vascular permeability in tumor xenografts monitored by dynamic contrast enhanced magnetic resonance imaging. *Clin Cancer Res* 11:529–536
23. Bouma MJ, Snowdon D, Fairlamb AH, Ackers JP (1998) Activity of disulfiram (bis(diethylthiocarbamoyl)disulphide) and ditiocarb (diethyldithiocarbamate) against metronidazole-sensitive and -resistant *Trichomonas vaginalis* and *Tritrichomonas foetus*. *J Antimicrob Chemother* 42:817–820. <https://doi.org/10.1093/jac/42.6.817>
24. Wang H, Liu Z, Li X, et al (2018) Shikonin causes apoptosis by disrupting intracellular calcium homeostasis and mitochondrial function in human hepatoma cells. *Exp Ther Med* 15:1484–1492. <https://doi.org/10.3892/etm.2017.5591>
25. Yu W, Goddard C, Clearfield E, et al (2011) Design, synthesis, and biological evaluation of triazolopyrimidine derivatives as novel inhibitors of hepatitis B virus surface antigen (HBsAg) secretion. *J Med Chem* 54:5660–5670. <https://doi.org/10.1021/jm200696v>
26. Vansteenkiste JF, Van De Kerkhove C, Wauters E, Van Mol P (2019) Capmatinib for the treatment of non-small cell lung cancer. *Expert Review of Anticancer Therapy* 19:659–671. <https://doi.org/10.1080/14737140.2019.1643239>
27. Kamata K, Suetsugu T, Yamamoto Y, et al (2006) Bisindole alkaloids from myxomycetes *Arcyria denudata* and *Arcyria obvelata*. *J Nat Prod* 69:1252–1254. <https://doi.org/10.1021/np060269h>

28. Yamada Y, Fujii T, Ishijima R, et al (2011) DR396, an apoptotic DNase gamma inhibitor, attenuates high mobility group box 1 release from apoptotic cells. *Bioorg Med Chem* 19:168–171. <https://doi.org/10.1016/j.bmc.2010.11.037>
29. Hussein Z, Mulford DJ, Bopp BA, Granneman GR (1993) Differences in the stereoselective pharmacokinetics of pazinaclone (DN-2327), a new anxiolytic, and its active metabolite after intravenous and oral single doses to dogs. *Drug Metab Dispos* 21:805–810
30. Ames BD, Haynes SW, Gao X, et al (2011) Complexity generation in fungal peptidyl alkaloid biosynthesis: oxidation of fumiquinazoline A to the heptacyclic hemiaminal fumiquinazoline C by the flavoenzyme Af12070 from *Aspergillus fumigatus*. *Biochemistry* 50:8756–8769. <https://doi.org/10.1021/bi201302w>
31. Thangadurai A, Minu M, Wakode S, et al (2012) Indazole: a medicinally important heterocyclic moiety. *Medicinal Chemistry Research* 21:1509–1523. <https://doi.org/10.1007/s00044-011-9631-3>
32. Bibian M, Rahaim RJ, Choi JY, et al (2013) Development of highly selective casein kinase 1delta/1epsilon (CK1delta/epsilon) inhibitors with potent antiproliferative properties. *Bioorg Med Chem Lett* 23:4374–4380. <https://doi.org/10.1016/j.bmcl.2013.05.075>
33. Hu E, Chen N, Bourbeau MP, et al (2014) Discovery of clinical candidate 1-(4-(3-(4-(1H-benzo[d]imidazole-2-carbon yl)phenoxy)pyrazin-2-yl)piperidin-1-yl)ethanone (AMG 579), a potent, selective, and efficacious inhibitor of phosphodiesterase 10A (PDE10A). *J Med Chem* 57:6632–6641. <https://doi.org/10.1021/jm500713j>
34. Nasab EM, Athari SM, Ghafarzade S, et al (2020) Immunomodulatory effects of two silymarin isomers in a Balb/c mouse model of allergic asthma. *Allergol Immunopathol (Madr)*. <https://doi.org/10.1016/j.aller.2020.01.003>
35. MarvinSketch. ChemAxon. <http://www.chemaxon.com>
36. Wishart YDF Roman Eisner, Craig Knox, Leonid Chepelev, Janna Hastings, Gareth Owen, Eoin Fahy, Christoph Steinbeck, Shankar Subramanian, Evan Bolton, Russell Greiner & David S (2016) ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J Cheminform* 61
37. Kim S, Chen J, Cheng T, et al (2019) PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res* 47:D1102–D1109. <https://doi.org/10.1093/nar/gky1033>
38. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46:3–26. [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0)
39. Veber DF, Johnson SR, Cheng H-Y, et al (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 45:2615–2623. <https://doi.org/10.1021/jm020017n>
40. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem* 1:55–68. <https://doi.org/10.1021/cc9800071>

41. O'Boyle NM, Banck M, James CA, et al (2011) Open Babel: An open chemical toolbox. J Cheminform 3:33. <https://doi.org/10.1186/1758-2946-3-33>
42. Berman HM, Westbrook J, Feng Z, et al (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242. <https://doi.org/10.1093/nar/28.1.235>
43. Burley SK, Berman HM, Kleywegt GJ, et al (2017) Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. Methods Mol Biol 1607:627–641. [https://doi.org/10.1007/978-1-4939-7000-1\\_26](https://doi.org/10.1007/978-1-4939-7000-1_26)
44. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461. <https://doi.org/10.1002/jcc.21334>
45. Schrödinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8
46. (2017) Discovery Studio Visualizer, Discovery Studio Modeling Environment, Release 2017. Dassault Systèmes BIOVIA

## Figures

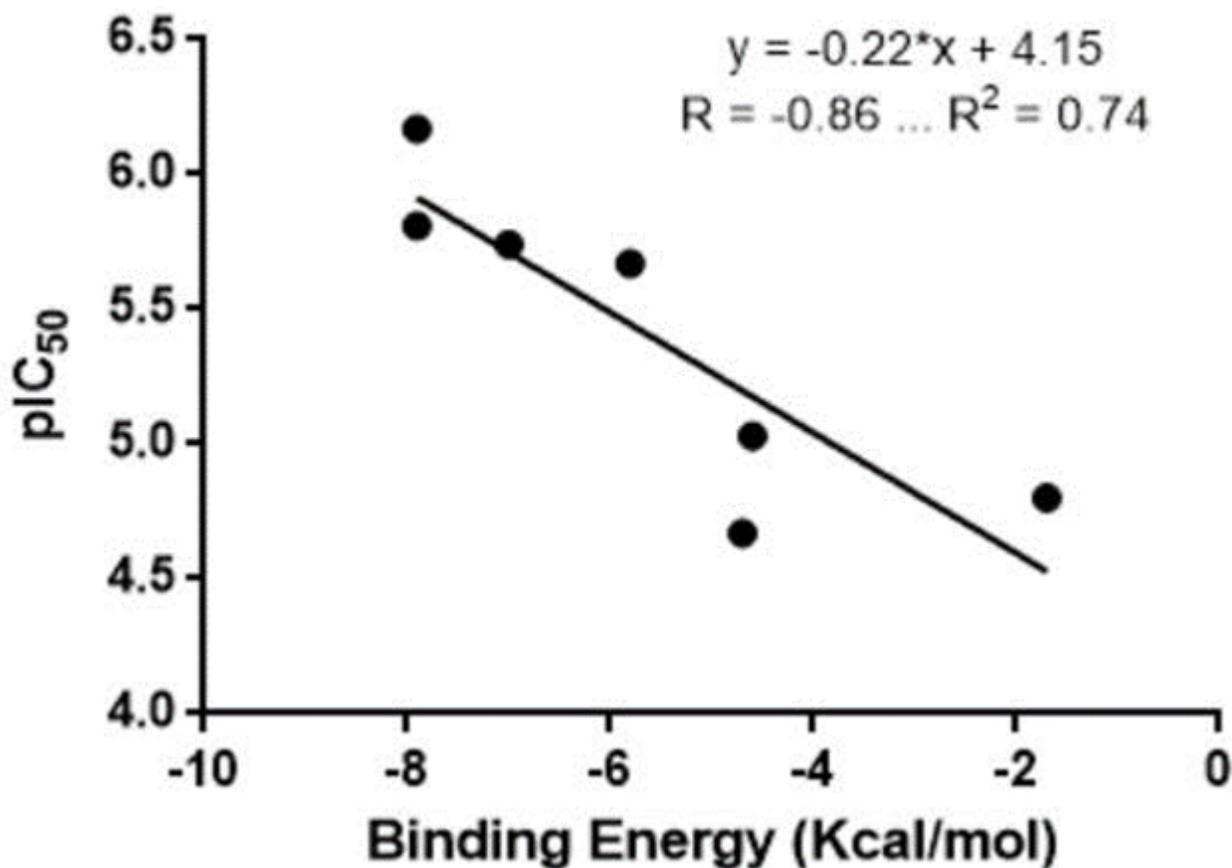


Figure 1

Simple linear regression plot of the docking scores and the pIC50 inhibition values of known Mpro protease inhibitor.

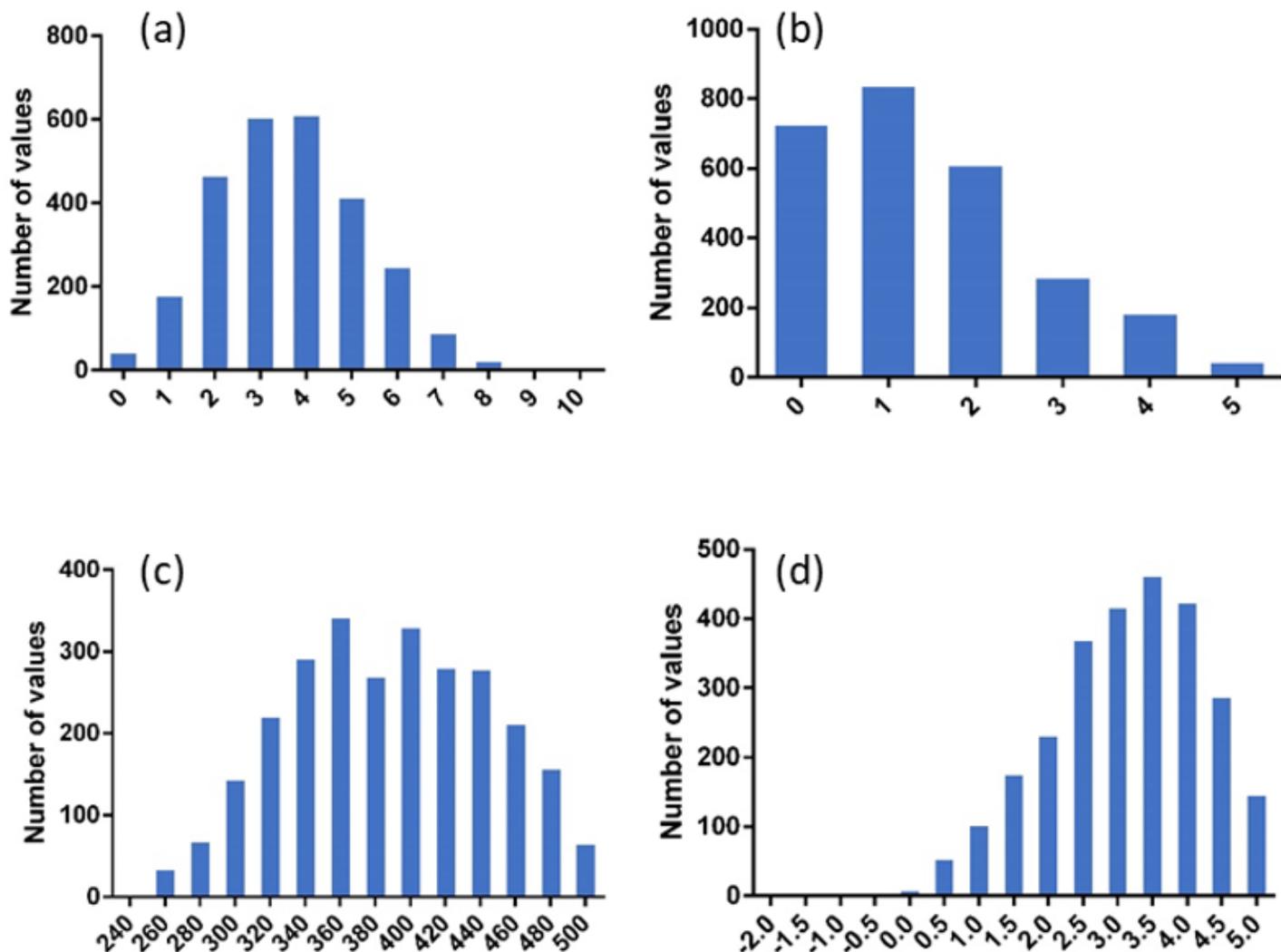
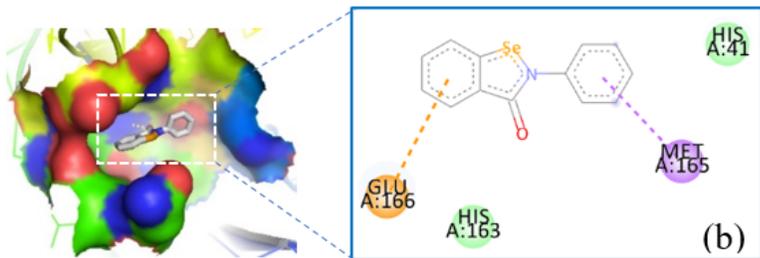
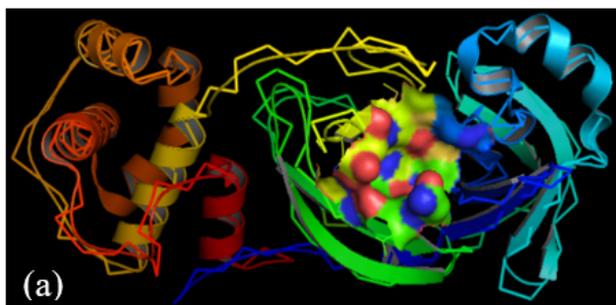


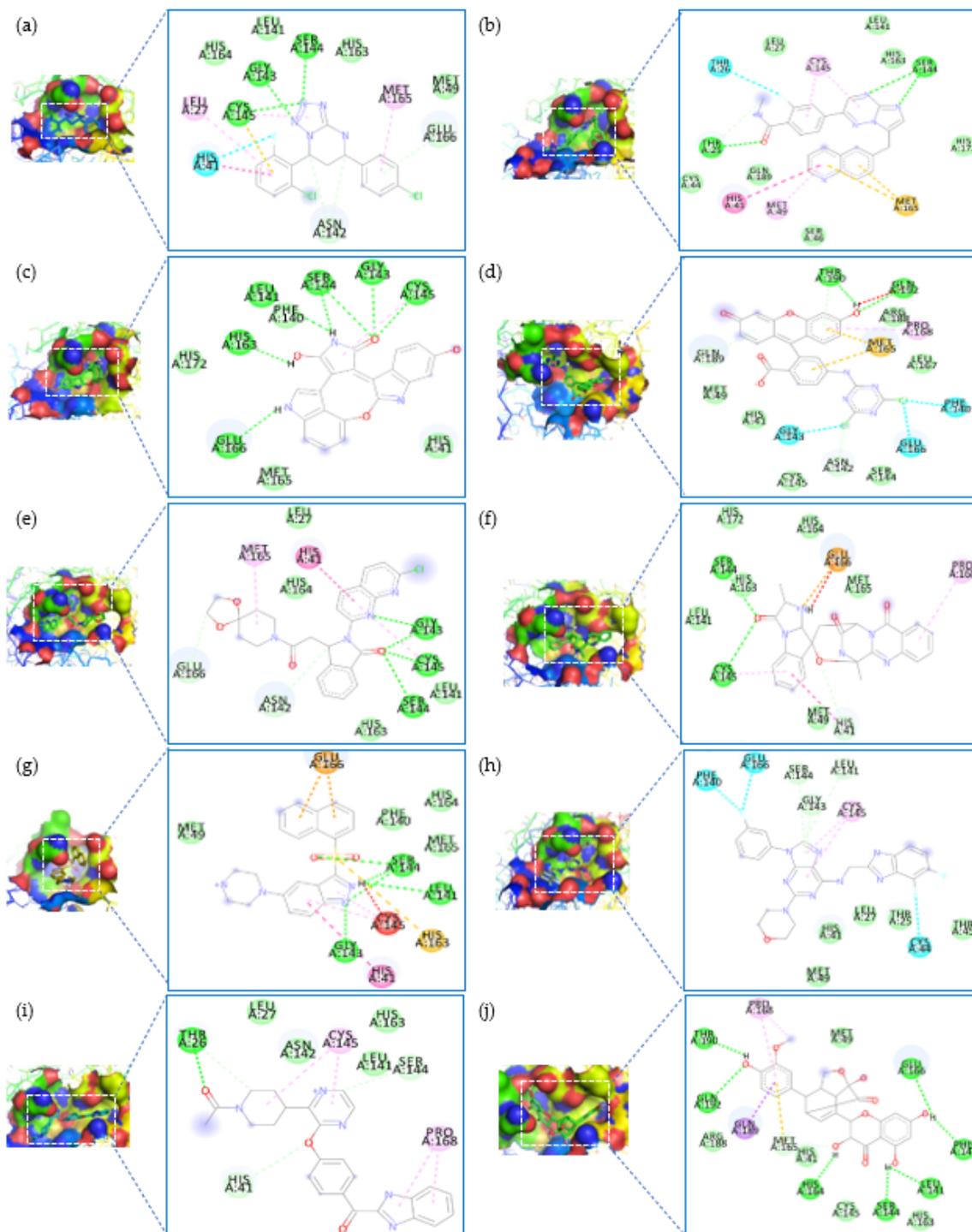
Figure 2

Distributions of the four Lipinski's rules within the selected set of compounds having a docking score  $\leq -9$  Kcal/mol. (a) Hydrogen bonds acceptors, (b) Hydrogen bonds donors, (c) LogP(o/w), (d) Molecular weight.



**Figure 3**

(a) Structure of Mpro protease showing the surface of its binding site. (b) Ebselen binding mode into Mpro protease active site and 2D receptor – ligand interactions map.



**Figure 4**

2D ligand-receptor interaction maps. (a) HBF-0259, (b) Capmatinib, (c) Arcyroxocin B, (d) DR396, (e) Pazinaclone, (f) Fumiquinazoline C, (g) Compound 57, (h) SR-3029, (i) AMG-579, (j) Sliydianin. Compounds and their assigned groups are listed in Table 1S.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.docx](#)