Computational discovery of small drug-like compounds as potential inhibitors of SARS-CoV-2 main protease

Alexander M. Andrianov\textsuperscript{a*}, Yuri V. Kornoushenko\textsuperscript{a}, Anna D. Karpenko\textsuperscript{b}, Ivan P. Bosko\textsuperscript{b} & Alexander V. Tuzikov\textsuperscript{b}

\textsuperscript{a}Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, Minsk, Republic of Belarus
\textsuperscript{b}United Institute of Informatics Problems, National Academy of Sciences of Belarus, Minsk, Republic of Belarus

* Corresponding address: Email: alexande.andriano@yandex.ru
Abstract

A computational approach to in silico drug discovery was carried out to identify small drug-like compounds able to show structural and functional mimicry of the high affinity ligand X77, potent non-covalent inhibitor of SARS-COV-2 main protease (MPro). In doing so, the X77-mimetic candidates were predicted based on the crystal X77-MPro structure by a public web-oriented virtual screening platform Pharmit. Models of these candidates bound to SARS-COV-2 MPro were generated by molecular docking and optimized by the quantum chemical method PM7. At the final point, analysis of the interaction modes of the identified compounds with MPro and prediction of their binding affinity were carried out.

Calculation revealed 5 top-ranking compounds that exhibited a high affinity to the active site of SARS-CoV-2 MPro. Insights into the ligand–MPro models indicate that all identified compounds may effectively block the binding pocket of SARS-CoV-2 MPro, in line with the low values of binding free energy and dissociation constant. Mechanism of binding of these compounds to MPro is generally provided by hydrogen bonds and van der Waals interactions with the functionally important residues of the enzyme active site, such as His-41, Leu-141, His-163, Met-165, and Glu-166. In addition, individual ligands form salt bridges with the MPro residues His-163 or Glu-166 and participate in specific π-π interactions with the catalytic dyad residue His-41.

The data obtained show that the identified X77-mimetic candidates may serve as good scaffolds for the design of novel antiviral agents able to target the active site of SARS-CoV-2 MPro.

Keywords: Coronavirus SARS-CoV-2, COVID-19, Main Protease, SARS-CoV-2 Inhibitors, Virtual Screening, Molecular Docking, Quantum Chemical Calculations, Antiviral Drugs
Introduction

The recent outbreak of coronavirus infection in China caused by the SARS-CoV-2 virus associated with COVID-19 has become a matter of serious concern to the world community, as the number of infected people is constantly increasing with significant geographical spread. As of the beginning of May 2020, the World Health Organization reports over 3.5 million confirmed cases of infection and over 240 thousand deaths. Numerous attempts are being made to develop an effective antiviral vaccine and find new therapeutic agents against COVID-19. Studies of various aspects of SARS-CoV-2, including structure, mechanism of action, epidemiology and genome sequencing, have provided important information about the new virus (Boopathi, Poma & Kolandaivel, 2020; Lu et al., 2020; Chan et al., 2020). According to the data obtained (Lu et al., 2020; Chan et al., 2020), SARS-CoV-2 belongs to a large family of coronaviruses that infect humans and other animal species, causing many widespread and serious diseases, such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) (de Wit, van Doremalen, Falzarano & Munster, 2016). The SARS-CoV-2 coronavirus genome is positive-sense, single-stranded RNA and consists of ~ 30,000 nucleotides, and its replicase gene encodes two overlapping polyproteins, pp1a and pp1ab, required for virus replication and transcription (Chen et al., 2020). These polyproteins undergo extensive proteolytic processing by two cysteine proteases, namely papain-like protease PLpro and 3-chymotrypsin-like protease 3CLpro (also known as the main protease M<sub>Pro</sub>) which is essential for mediating viral replication and transcription (Anand et al., 2002; Yang et al., 2003). The main protease digests polyprotein at no less than 11 conserved sites, starting with the autolytic cleavage of this enzyme itself from pp1a and pp1ab (Hegyi & Ziebuhr, 2002). This indicates the extremely important role of M<sub>Pro</sub> in the life virus cycle and makes this enzyme one of the most attractive targets for the development of effective antiviral drugs (Pillaiyar et al., 2016).

In the newest studies, SARS-CoV-2 M<sub>Pro</sub> has been used as a target for screening clinically approved drugs as potential virus inhibitors in the hope of identifying drugs that are effective against COVID-19 (e.g., Rismanbaf, 2020; Liu & Wang, 2020; Yan et al., 2020; Zhijian et al., 2020; Enmozhi et al., 2020; Khan et al., 2020; Muralidharan, Sakthivel, Velmurugan & Gromiha, 2020; Zhou et al., 2020). Since the safety profiles of these drugs are well-documented, such an
approach combining the structural design of drugs with virtual screening and molecular modeling methods can significantly facilitate and accelerate the detection of antiviral compounds with clinical potential in order to re-profile them for the treatment of patients infected with a new type of coronavirus. However, taking into account SARS-CoV-2 mutations (Forster et al., 2020; Pachetti et al., 2020; Yao et al., 2020; Khailany, Safdar & Ozaslan, 2020), studies on the development of novel antiviral compounds capable of blocking the functionally important sites of the viral proteins are also extremely significant.

Determination of the high-resolution X-ray structures of SARS-CoV-2 main protease (MPro) in ligand-bound and unbound states (Berman et al., 2000; http://www.rcsb.org/pdb/) laid the foundation not only for understanding the function and molecular mechanism of the enzyme action, but also for developing novel effective SARS-CoV-2 inhibitors by direct methods of computer-aided drug design (e.g., Jin et al., 2020; Ton et al., 2020; Fischer et al., 2020; Vega-Valdez et al., 2020; Olubiyi et al., 2020; Pant et al., 2020; Islam et al., 2020). In particular, SARS-CoV-2 MPro structure in the complex with the high affinity ligand X77 that is potent non-covalent inhibitor of SARS-COV-2 MPro was recently deposited in the Protein Data Bank (PDB ID: 6W63, http://www.rcsb.org/pdb/).

In this study, an integrated computational approach to in silico drug discovery was carried out to discover small drug-like compounds able to show structural and functional mimicry of the inhibitor X77. This computer-based approach included i) generation of pharmacophore model representing 3D-arrangements of chemical functionalities that make X77 active towards the active site of SARS-CoV-2 MPro, (ii) shape/pharmacophore-based identification of the X77-mimetic candidates by a web-oriented virtual screening platform Pharmit (http://pharmit.csb.pitt.edu) allowing one to search for small molecules based on their structural and chemical similarity to another small molecule (Sunseri & Koes, 2016), iii) identification of compounds satisfying the Lipinski’s “rule of five” (Lipinski, Lombardo, Dominy & Feeney, 2001) that recognizes molecules with drug-like properties, iv) molecular docking of these drug-like compounds with the enzyme active site, v) prediction of the interaction modes dominating the binding; vi) calculation of the values of binding free energy and dissociation constant (Kd) for the docking ligand–MPro models, vii) optimization of these models using the semiempirical quantum chemical method PM7.
(Stewart, 2013), viii) calculation of the values of binding enthalpy for the PM7-based ligand–M\(^{\text{Pro}}\) structures, and ix) selection of molecules most promising for biochemical trials.

As a result, an ensemble of hit compounds that bind to the active site of SARS-CoV-2 M\(^{\text{Pro}}\) and specifically interact with the functionally important residues of the enzyme was identified. These compounds are suggested to form good scaffolds for the development of novel, potent and broad drugs against COVID-19.

Methods and materials

Virtual screening: The Pharmit server software (Sunseri & Koes, 2016; http://pharmit.csb.pitt.edu) was used to generate the X77 pharmacophore model based on the X77–M\(^{\text{Pro}}\) complex in crystal (PDB ID: 6w63; https://www.rcsb.org). This model (Table 1) was applied for virtual screening of small-molecule compounds able to block the X77-binding site of SARS-CoV-2 M\(^{\text{Pro}}\). Virtual screening was performed in the 9 Pharmit molecular libraries containing over 213.5 million chemical structures (Sunseri & Koes, 2016; http://pharmit.csb.pitt.edu), resulting in a set of compounds that satisfied the X77 pharmacophore model (Table 1) and Lipinski's “rule of five” (Lipinski, Lombardo, Dominy & Feeney, 2001). These molecules were further screened by molecular docking and quantum chemical calculations to evaluate the affinity of their binding to SARS-CoV-2 M\(^{\text{Pro}}\) and identify molecules most promising for biochemical assays.

Table 1. Pharmacophore model of X77 used for virtual screening of the Pharmit chemical databases

| Pharmacophore type     | Pharmacophore coordinates X, Y, Z (Å) | Pharmacophore radius (Å) |
|------------------------|---------------------------------------|--------------------------|
| Aromatic group         | -20.75 17.39 -28.53                   | R = 1.1                  |
| Aromatic group         | -20.55 20.33 -31.86                   | R = 1.1                  |
| H-bond acceptor        | -16.19 21.86 -26.86                   | R = 0.5                  |
| H-bond acceptor        | -20.84 19.52 -32.66                   | R = 0.5                  |
| H-bond acceptor        | -19.75 22.16 -29.14                   | R = 0.5                  |
| H-bond acceptor        | -18.66 18.65 -25.94                   | R = 0.5                  |
| Hydrophobic group      | -20.55 20.33 -31.86                   | R = 1.0                  |
**Molecular docking:** Molecular docking of the predicted compounds with SARS-CoV-2 M\(^{\text{Pro}}\) was carried out by the QuickVina 2 program (Alhossary, Handoko, Mu & Kwoh, 2015) in the approximation of rigid receptor and flexible ligands. The X77 inhibitor (Figure 1) was used in the calculations as a positive control. The 3D structure of this compound was isolated from the crystal X77–M\(^{\text{Pro}}\) complex (the PDB ID: 6W63; http://www.rcsb.org/pdb/). The SARS-CoV-2 M\(^{\text{Pro}}\) and ligand structures were prepared by adding hydrogen atoms with the Open Babel software (http://openbabel.org/wiki/Main_Page) followed by their optimization in the UFF force field (Rappe et al., 1992). The ligands were docked to the crystal SARS-CoV-2 M\(^{\text{Pro}}\) structure using QuickVina 2 (Alhossary, Handoko, Mu & Kwoh, 2015). The grid box included the X77-binding site of SARS-CoV-2 M\(^{\text{Pro}}\) and had the following parameters: \(\Delta X = 19\ \text{Å}, \Delta Y = 21\ \text{Å}, \Delta Z = 23\ \text{Å}\) centered at \(X = -20\ \text{Å}, Y = 19\ \text{Å}, Z = -26\ \text{Å}\); that is, the box volume was \(19 \times 21 \times 23 = 9177\ \text{Å}^3\). The value of “exhaustiveness” parameter defining number of individual sampling “runs” was set to 1000 (Alhossary, Handoko, Mu & Kwoh, 2015).

![Chemical structure of X77](image)

**(R)-N-(4-tert-butylphenyl)-N-(2-cyclohexylamino)-2-oxo-1-(pyridin-3-yl)ethyl]-1H-imidazole-5-carboxamide**

**Figure 1.** Chemical structure of X77, potent non-covalent inhibitor of SARS-CoV-2 M\(^{\text{Pro}}\) (PDB ID: 6W63, http://www.rcsb.org/pdb/). The systematic name of this compound is given
Quantum chemical calculations: The quantum chemical optimization of the docked ligand–M\textsuperscript{Pro} structures was carried out using the semiempirical quantum chemical method PM7 (Stewart, 2013) associated with the MOPAC2016 software package (http://OpenMOPAC.net). Before the calculations, the ligand–M\textsuperscript{Pro} complexes were supplemented with hydrogen atoms and optimized in the UFF force field (Rappe et al., 1992). For this purpose, the Open Babel program (http://openbabel.org/wiki/Main_Page) was used. The calculations were performed in the COSMO solvation model (COnductor-like Screening MOdel) approximation (Klamt & Schüürmann, 1993; Klamt, 2005; Klamt, Moya & Palomar, 2015) in an implicit solvent with water's dielectric constant of 78.4 (http://OpenMOPAC.net). To speed up the calculations, the Localized Molecular Orbitals method (Høyvik, Jansik & Jørgensen, 2012; Lehtola & Jónsson, 2013) available in MOPAC in the form of the linear scaling SCF MOZYME algorithm (Stewart, 2013) was applied. The value of RMS gradient was set to 10 kcal/mol/Å.

Analysis of interaction modes and binding affinity profile: The binding modes of the predicted compounds to SARS-CoV-2 M\textsuperscript{Pro}, namely hydrogen bonds, salt bridges, van der Waals contacts, and π-π interactions were identified by the BINANA program (Durrant & McCammon, 2011\textsuperscript{a}). The ligand poses in the docking ligand–M\textsuperscript{Pro} models were visualized with the program UCSF Chimera (Pettersen et al., 2004). To visualize van der Waals contacts, the program LigPlot (McDonald & Thornton, 1994) was employed. The values of K\textsubscript{d} for the ligand–M\textsuperscript{Pro} structures were calculated using a neural-network-based scoring function NNScore 2.0 (Durrant & McCammon, 2011\textsuperscript{b}). The values of binding free energy were estimated from those of K\textsubscript{d} by the formula ∆G = R×T×ln(K\textsubscript{d}) (where ∆G is the binding free energy, R is the universal gas constant, T is the absolute temperature equal to 310 K) (Sharma & First, 2009).

For the PM7-based complexes, the ligand-binding affinity was estimated in terms of the values of binding enthalpy ∆H calculated as the differences between the heats of formation of the ligand–M\textsuperscript{Pro} complexes and heats of formation of the ligand and M\textsuperscript{Pro} in the unbound states (Stewart, 2013; http://OpenMOPAC.net). Quantum chemical calculations of the ligand and M\textsuperscript{Pro} structures in the unbound states were performed using the computational protocol described above for the docking ligand–M\textsuperscript{Pro} models.
Results and discussion

Shape/Pharmacophore-based virtual screening of the Pharmit databases resulted in 24 molecules that exhibited favorable binding energies (< −6 kcal/mol) and the values of root-mean-square deviations between the query features and the hit compound features less than 2 Å (Sunseri & Koes, 2016). Molecular docking of these molecules with the active site of M\textsuperscript{Pro} followed by quantum chemical calculations identified 5 top-ranking compounds (Figure 2) that showed a high-affinity binding in terms of $K_d$, binding free energy, and binding enthalpy (Table 2). This allowed one to consider these compounds as the most probable X77-mimetic candidates. Inspection of the physicochemical parameters of the predicted compounds (Table 3) providing such important characteristics for a potential drug as absorption, distribution, metabolism and excretion indicates that these molecules fully satisfy the requirements of the Lipinski's “rule of five” (Lipinski, Lombardo, Dominy & Feeney, 2001).

Insights into the docking ligand–M\textsuperscript{Pro} models (Figure 3) show that all identified X77-mimetic candidates form a wide network of intermolecular interactions involving amino acid residues of the binding pocket of M\textsuperscript{Pro}. In particular, compound I exhibiting the lower values of $K_d$ and binding free energy compared to the other predicted molecules and X77 (Table 2) forms 3 hydrogen bonds with the M\textsuperscript{Pro} residues Ser-144, His-163 and Glu-166, a salt bridge with His-163, and 20 van der Waals contacts with the active site residues His-41, Met-49, Leu-141, Asn-142, Met-165, Glu-166, Asp-187, and Gln-189 (Table 4). In addition to these direct interatomic contacts, compound I is also involved in specific $\pi$-\$\pi$ interaction with His-41 which is a part of the catalytic dyad of M\textsuperscript{Pro} formed by this residue and Cys-145 (Chang, 2020; Qamar, Alqahtani, Alamri & Chen, 2020). Examination of the intermolecular interaction profile calculated for the other identified compounds indicates (Table 4) that these molecules exhibit the modes of binding to SARS-CoV-2 M\textsuperscript{Pro} similar to those predicted for compound I. This binding is generally provided by hydrogen bonds, van der Waals contacts, salt bridges (compounds I and III) and $\pi$-$\pi$ interactions between $\pi$-conjugated systems of the ligands and the side chain of His-41 (compounds I, III and IV) (Table 4, Figure 4). Among these binding modes, intermolecular van der Waals interactions are the major contributors to the ligand–M\textsuperscript{Pro} interface including significant residues of the enzyme active pocket (Table 4, Figure 4).
Figure 2. Chemical structures of the potential SARS-CoV-2 M<sup>Pro</sup> inhibitors. The systematic names of the compounds, as well as the corresponding databases with codes for these molecules are given. The ligand functional groups participating in the formation of intermolecular hydrogen bonds are marked by superscript numbers (see the text and Table 4)
The efficiency of the intermolecular interactions of the X77-mimetic candidates with SARS-CoV-2 M<sup>Pro</sup> is supported by the low values of K<sub>d</sub> (0.006 μM – 2.56 μM) and binding free energies (ΔG ≤ –7.9 kcal/mol), indicating their high affinity with the catalytic site of the enzyme (Table 2). Analysis of the values of K<sub>d</sub> and binding free energy calculated for the identified compounds shows that, given the calculation errors, they are comparable with those obtained for X77 using the identical computational protocol (Table 2).

**Table 2. Values of dissociation constant (K<sub>d</sub>), binding free energy (ΔG) and binding enthalpy (ΔH) calculated for the identified compounds and X77 bound to SARS-CoV-2 M<sup>Pro</sup>**

| Ligand | I    | II   | III  | IV   | V    | X77 |
|--------|------|------|------|------|------|-----|
| K<sub>d</sub>, (μM) | 0.006 | 0.039 | 0.157 | 2.0  | 2.65 | 0.057 |
| ΔG, (kcal/mol) | -11.65 | -10.50 | -9.64 | -8.07 | -7.90 | -10.21 |
| ΔH, (kcal/mol) | -80.1 | -96.6 | -90.7 | -71.4 | -53.78 | -62.8 |

**Table 3. Physicochemical parameters of the X77-mimetic candidates associated with the Lipinski's “rule of five”**

| Ligand | Chemical formula | Molecular mass (Da) | LogP | Number of H-bond donors | Number of H-bond acceptors |
|--------|------------------|--------------------|------|-------------------------|---------------------------|
| I      | C<sub>21</sub>H<sub>25</sub>N<sub>3</sub>O<sub>4</sub> | 384.00             | 2.225 | 3                       | 6                         |
| II     | C<sub>19</sub>H<sub>26</sub>N<sub>5</sub>O<sub>5</sub> | 419.00             | -1.206 | 3                       | 9                         |
| III    | C<sub>19</sub>H<sub>21</sub>N<sub>5</sub>O<sub>3</sub> | 368.00             | 0.146 | 5                       | 6                         |
| IV     | C<sub>21</sub>H<sub>24</sub>N<sub>6</sub>O<sub>6</sub> | 456.00             | 0.071 | 5                       | 11                        |
| V      | C<sub>17</sub>H<sub>20</sub>ClN<sub>5</sub>O<sub>3</sub> | 378.00             | -1.090 | 3                       | 6                         |
Figure 3. Structural complexes of compounds I, II, III, IV, and V with SARS-CoV-2 M\textsuperscript{Pro} generated by molecular docking. The compounds are represented by a ball-stick-ball model. The enzyme residues forming interatomic contacts with the ligands are indicated (Table 4). Residues of M\textsuperscript{Pro} involved in hydrogen bonding are noted using a stick model. Hydrogen bonds are shown by solid lines. A wire model is used to designate residues forming van der Waals contacts, salt bridges, and π- or T-stacking.
Table 4. Intermolecular interactions appearing in the structural complexes of the identified compounds with SARS-CoV-2 M<sub>Pro</sub>

| Ligand | Hydrogen bond<sup>1</sup> | Van der Waals contacts<sup>2</sup> | Salt bridges and π-π interactions<sup>3</sup> |
|--------|--------------------------|---------------------------------|----------------------------------|
| I      | O<sup>1</sup>...**HN[S144] | H41(3), M49(1), L141(3), N142(1), M165(6), E166(3), D187(2), Q189(1) | COO...H163 H41 (T-stacking) |
|        | O<sup>2</sup>...**HN[H163] |                                 |                                  |
|        | O<sup>3</sup>...HN[E166]  |                                 |                                  |
| II     | O<sup>1</sup>H...*O[M49]  | T25(1), H41(2), M165(3), L167(1), P168(1), Q189(1), Q192(1) | –                                 |
|        |                           |                                 |                                  |
| III    | O<sup>1</sup>H...*N[L141] | H41(4), C44(1), M49(1), M165(4), L141(2), N142(1), E166(2), Q189(5) | NCHC...E166 H41 (T- и π-stacking) |
|        | O<sup>2</sup>...*N[G143] |                                 |                                  |
|        | O<sup>2</sup>...NH[G143] |                                 |                                  |
|        | O<sup>1</sup>...**HN[H163] |                                 |                                  |
| IV     | O<sup>1</sup>H...*O[C44]  | T25(2), H41(1), C44(1), M49(1), M165(3), Q189(4) | H41 (T-stacking) |
|        | N<sup>1</sup>...*HN[E166] |                                 |                                  |
|        | O<sup>2</sup>...*HN[T190] |                                 |                                  |
|        | O<sup>2</sup>...**HN[Q192] |                                 |                                  |
| V      | O<sup>1</sup>...*HN[E166] | C44(2), M49(1), L141(1), N142(1), M165(4), E166(1), R188(1) | –                                 |

Footnotes: 1Atoms of the ligands are shown first, followed by the corresponding atoms of SARS-CoV-2 M<sub>Pro</sub> (M<sub>Pro</sub> residues are in brackets in one-letter code). Symbol * denotes the atoms of the residue main chain, and symbol ** marks the atoms of the residue side chain. 2Amino acids of SARS-CoV-2 M<sub>Pro</sub> forming van der Waals contacts with the ligands. The number of contacts is given in brackets. 3For salt bridges, the functional groups of ligands are shown first, followed by the residues of SARS-CoV-2 M<sub>Pro</sub>. For π- or T-stacking, residue of SARS-CoV-2 M<sub>Pro</sub> involved in these interactions is shown.

So, the data obtained show that all identified compounds (Figure 2) may effectively block the key residues of the M<sub>Pro</sub> catalytic site, which is confirmed by the low values of binding free energy and K<sub>d</sub> calculated for the docking ligand–M<sub>Pro</sub> models (Table 2). This is also evidenced by the data of quantum chemical calculations which show that, excluding compound V, the values of binding enthalpy of the analyzed molecules to MPro are lower than that predicted for the control inhibitor X77 by the same computational parameters (Table 2).
Figure 4. The SARS-CoV-2 M^Pro^ residues making direct interatomic contacts with compounds I, II, III, IV, and V. Residues involved in hydrogen bonding are marked by ellipses and highlighted in darker color.

However, when analyzing the calculation data, it is necessary to keep in mind that all computational approaches for modeling ligand–protein complexes and estimating the binding affinity involve various approximations. They vary from simplified forms of the first-principles equations that are easier or faster to solve, to approximations limiting the size of the system, to fundamental approximations to the underlying equations that are required to achieve any solution to them at all. Nevertheless, the findings of comparative analysis of the X77–M^Pro^ complexes constructed by the X-ray crystallography and molecular docking (Figure 5) indicate good
prediction accuracy of the computational algorithm used in the calculations, suggesting that the
data obtained for the identified compounds adequately describe the principal geometric and energy
c characteristics of their complexes with SARS-CoV-2 M\textsuperscript{Pro}.

**Figure 5.** Superposition of the X77–M\textsuperscript{Pro} complexes constructed using X-ray crystallography and
molecular docking. The root-mean-square deviation between the atomic coordinates of the X77
inhibitor in the calculated and experimental structures is 0.63 Å. Residues of M\textsuperscript{Pro} forming direct
interatomic contacts with X77 are marked by a wire model

Among the identified X77-mimetic candidates, it should be emphasized compound I (Figure 2)
that exposes the lowest value of $K_d$ and $\Delta G$ compared with the other identified molecules and X77
and demonstrates the low value of binding enthalpy (Table 2). This high-affinity binding is
provided by a large number of intermolecular interactions associated with the important residues
of the enzyme active site, such as His-41, Asn-142, His-163, and Glu-166 (Jin et al., 2020, Zhang
et al., 2020) (Table 4, Figure 4). Furthermore, compound I mimics most of the key interactions of
the inhibitor X77 with the binding pocket of SARS-CoV-2 M\textsuperscript{Pro}. According the data obtained, 7
out of 11 residues of the X77–M\textsuperscript{Pro} interface (Figure 5), namely His-41, Leu-141, His-163, Met-
165, Glu-166, Asp-187, and Gln-189, are involved in the direct interatomic contacts with the
analyzed molecule (compare the data of Table 4 and Figure 5). This small-molecule hit is therefore
a higher-priority candidate for further detailed experimental evaluation.

14
However, it is clear that, despite promising *in silico* profile, the analyzed compound is only a starting point for the development of new highly potent drug candidates. In this connection, before biochemical assays, this compound should go through a lead optimization, iterative process of altering the molecule structure to identify their chemical modifications with improved antiviral potency and ADMET parameters. For this purpose, the modern QSAR methods commonly used as a lead optimization approach in drug discovery may be applied (Golbraikh, Wang, Zhu & Tropsha, 2017; Diderich, Kuseva & Mekenyan, 2018; Kuseva et al., 2019).

**Conclusions**

Shape/Pharmacophore-based virtual screening combined with molecular docking and quantum chemical calculation revealed 5 top-ranking compounds that exhibited a high affinity to the catalytic site of SARS-CoV-2 M<sup>Pro</sup>, allowing one to consider these small drug-like molecules as the most promising X77-mimetic candidates. Insights into the ligand–M<sup>Pro</sup> models indicate that all identified compounds may specifically and effectively block the active site of SARS-CoV-2 M<sup>Pro</sup>, in line with the low values of binding free energy, dissociation constant and binding enthalpy. Mechanism of binding of these compounds to M<sup>Pro</sup> is generally provided by hydrogen bonds and van der Waals interactions with the functionally important residues of the enzyme active site, such as His-41, Leu-141, His-163, Met-165, and Glu-166. In addition, individual ligands form salt bridges with the M<sup>Pro</sup> residues His-163 (compound I) or Glu-166 (compound III) and participate in specific π-π interactions with the catalytic dyad residue His-41 (compounds I, III, IV).

Taken together, the data obtained show that the identified X77-mimetic candidates may serve as good scaffolds for the design of novel antiviral agents able to target the active pocket of SARS-CoV-2 M<sup>Pro</sup>.

**Acknowledgments**

This study was supported by a grant of the State Program of Scientific Research “Convergence 2020” (subprogram “Consolidation”, project 3.08).
References:

Alhossary, A., Handoko, S.D., Mu, Y., & Kwoh, C.K. (2015). Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics, 31*(13), 2214-2216. doi: 10.1093/bioinformatics/btv082

Anand, K., Palm, G.J., Mesters, J.R., Siddell, S.G., Ziebuhr, J., & Hilgenfeld, R. (2002). Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra α-helical domain. *The EMBO Journal, 21*(13), 3213-3224. doi: 10.1093/emboj/cdf327

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., & Bourne, P.E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*(1), 235–242. https://doi.org/10.1093/nar/28.1.235

Boopathi, S., Poma, A.B., & Kolandaivel, P. (2020): Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *Journal of Biomolecular Structure and Dynamics*. doi: 10.1080/07391102.2020.1758788

Chan, J.F.-W., Yuan, S., Kok, K.-H., To, K. K.-W, Chu, H., Yang, J.,…Yuen. K-Y. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet, 395*(10223), 514-523. doi.org/10.1016/S0140-6736(20)30154-9

Chang, G.G. (2020). Quaternary structure of the SARS coronavirus main protease. *Molecular Biology of the SARS-Coronavirus*: 115–128. doi:10.1007/978-3-642-03683-5_8

Chen, Y., Liu, Q., & Guo, D. (2020). Coronaviruses: genome structure, replication, and pathogenesis. *Journal of Medical Virology, 92*(4), 418-423. doi: 10.1002/jmv.25681

De Wit, E., van Doremalen, N., Falzarano, D., & Munster, V. J. (2016). SARS and MERS: recent insights into emerging coronaviruses. *Nature Reviews Microbiology, 14*(8), 523-534. doi: 10.1038/nrmicro.2016.81

Diderich, R., Kuseva, C.D., & Mekenyan, O.G. (2018). The OECD QSAR Toolbox starts its second decade. *Methods in Molecular Biology, 1800*, 55–77. doi: 10.1007/978-1-4939-7899-1_2
Durrant, J.D., & McCammon, J.A. (2011).^a BINANA: A novel algorithm for ligand-binding characterization. *Journal of Molecular Graphics & Modeling, 29*(6), 888–893. doi: 10.1016/j.jmgm.2011.01.004

Durrant, J.D., & McCammon, J.A. (2011).^b NNScore 2.0: A neural-network receptor–ligand scoring function. *Journal of Chemical Information & Modeling, 51*(11), 2897–2903. doi: 10.1021/ci2003889

Enmozhi, S.K., Raja, K., Sebastine, I., & Joseph J. (2020). Andrographolide as a potential inhibitor of SARS-CoV-2 main protease: An in silico approach. *Journal of Biomolecular Structure & Dynamics*. doi: 10.1080/07391102.2020.1760136

Fischer, A., Sellner, M., Neranjan, S., Lill, M. A., & Smiesko, M. Inhibitors for novel coronavirus protease identified by virtual screening of 687 million compounds. *ChemRxiv*. doi.org/10.26434/chemrxiv.11923239.v1

Forster, P., Forster, L., Renfrew, C., & Forster, M. (2020). Phylogenetic network analysis of SARS-CoV-2 genomes. *Proceedings of the National Academy of Sciences of the United States of America, (Latest Articles)*. doi.org/10.1073/pnas.2004999117

Golbraikh, A, Wang, X.S., Zhu, H., & Tropsha, A. (2017). Predictive QSAR modeling: Methods and applications in drug discovery and chemical risk assessment. In: Leszczynski J, Kaczmarek-Kedziera A, Puzyn T, G. Papadopoulos M, Reis H, K. Shukla M (eds) Handbook of Computational Chemistry. Springer, Cham.

Hegyi, A., & Ziebuhr, J. (2002). Conservation of substrate specificities among coronavirus main proteases. *Journal of General Virology, 83*(3), 595-599. doi: 10.1099/0022-1317-83-3-595

Høyvik, I.-M., Jansik, B., & Jørgensen, P. (2012). Trust region minimization of orbital localization functions. *Journal of Chemical Theory & Computation, 8*(9), 3137–3146. doi.org/10.1021/ct300473g

Islam, R., Parves, R., Paul, A.S., Uddin, N., Rahman, M.S., Mamun, A.A., Hossain, M.N., Ali, M.A., & Halim, M.A. (2020). A molecular modeling approach to identify effective antiviral
phytochemicals against the main protease of SARS-CoV-2. *Journal of Biomolecular Structure & Dynamics.* doi: 10.1080/07391102.2020.1761883

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., …Yang, H. (2020). Structure of MPro from 1 COVID-19 virus and discovery of its inhibitors. *Nature.* https://doi.org/10.1038/s41586-020-2223-y

Khailany, R.A., Safdar, M., & Ozaslan, M. (2020). Genomic characterization of a novel SARS-CoV-2. *Gene Reports, 19,* 100682. doi:10.1016/j.genrep.2020.100682

Khan, R.J., Jha, R. K., Amera, G. M., Jain, M., Singh, E., Pathak, A., Singh, R.P., Muthukumaran, J., & Singh, A.K. (2020). Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like protease and 2′-O-ribose methyltransferase. *Journal of Biomolecular Structure & Dynamics.* doi: 10.1080/07391102.2020.1753577

Muralidharan, N., Sakthivel, R., Velmurugan D., & Gromiha, M.M. (2020). Computational studies of drug repurposing and synergism of lopinavir, oseltamivir and ritonavir binding with SARS-CoV-2 Protease against COVID-19. *Journal of Biomolecular Structure & Dynamics.* doi: 10.1080/07391102.2020.1752802

Klamt, A. COSMO-RS: From quantum chemistry to fluid phase thermodynamics and drug design, 1st ed.; Elsevier: Boston, MA, USA, 2005; 246 pages.

Klamt, A., & Schüürmann, G. (1993). COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions, 2,* 799–805.

Klamt, A., Moya, C., & Palomar, J. (2015). A comprehensive comparison of the IEFPCM and SS(V)PE continuum solvation methods with the COSMO approach. *Journal of Chemical Theory & Computation, 11*(9), 4220–4225. doi.org/10.1021/acs.jctc.5b00601

Kuseva, C., Schultz, T.W., Yordanova, D., Tankova, K., Kutsarova, S., Pavlov, T., Chapkanov, A., Georgiev, M., Gissi, A., Sobanski, T., & Mekenyan, O.G. (2019). The implementation of RAAF in the OECD QSAR Toolbox. *Regular Toxicology & Pharmacology, 105,* 51–61. doi: 10.1016/j.yrtph.2019.03.018
Lehtola, S., & Jónsson, H. (2013). Unitary optimization of localized molecular orbitals. *Journal of Chemical Theory & Computation*, 9(12), 5365–5372. doi.org/10.1021/ct400793q

Lipinski, C.A., Lombardo, F., Dominy, B.W., & Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 46(1-3), 3–26. PMID: 11259830

Liu, X., & Wang, X.-J. (2020). Potential inhibitors for 2019-nCoV coronavirus M protease from clinically approved medicines. *BioRxiv*. doi.org/10.1101/2020.01.29.924100

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H,...,Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*, 395(10224), 565–574. doi.org/10.1016/S0140-6736(20)30251-8

McDonald, I.K., & Thornton, J.M. (1994). Satisfying hydrogen bonding potential in proteins. *Journal of Molecular Biology*, 238(5), 777–793. doi:10.1006/jmbi.1994.1334

Olubiyi, O.O., Olagunju, M., Keutmann, M., Loschwitz, J., & Strodel, B. (2020). High throughput virtual screening to discover inhibitors of the main protease of the coronavirus SARS-CoV-2. *Preprint from Preprints.org*. doi: 10.20944/preprints202004.0161.v1

Pant, S., Singh, M., Ravichandiran, V., Murty U.S.N., & Srivastava. H.K. (2020). Peptide-like and small-molecule inhibitors against Covid-19. *Journal of Biomolecular Structure and Dynamics*. doi: 10.1080/07391102.2020.1757510

Pachetti, M., Marini, B., Benedetti, F., Giudici, F. Mauro, E., Storici, P,..., & Ippodrino, R. (2020). Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *Journal of Translational Medicine*, 18 (1):179. doi: 10.1186/s12967-020-02344-6

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., & Ferrin, T.E. (2004). UCSF Chimera – a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. doi: 10.1002/jcc.20084

Pillaiyiar, T., Manickam, M., Namasivayam, V., Hayashi, Y., & Jung, S. H. (2016). An overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL protease inhibitors:
peptidomimetics and small molecule chemotherapy. *Journal of Medicinal Chemistry, 59*(14), 6595-6628, doi:10.1021/acs.jmedchem.5b01461

Qamar, M.T., Alqahtani, S.M., Alamri, M.A., & Chen, L.-L. (2020). Structural basis of SARS-CoV-2 3CL\(^{pro}\) and anti-COVID-19 drug discovery from medicinal plants. *Journal of Pharmaceutical Analysis*. doi.org/10.1016/j.jpha.2020.03.009

Rappe, A.K., Casewit, C.J., Colwell, K.S., Goddard, III W.A., & Skiff, W.M. (1992). UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of American Chemical Society, 114*(25), 10024–10035. doi.org/10.1021/ja00051a040

Rismanbaf, A. (2020). Potential treatments for COVID-19; a narrative literature review. *Archives of Academic Emergency Medicine, 8*(1), e29

Sharma, G., & First, E.A. (2009). Thermodynamic analysis reveals a temperature-dependent change in the catalytic mechanism of bacillus stearothermophilus tyrosyl-tRNA synthetase. *Journal of Biological Chemistry, 284*(7), 4179–4190. doi: 10.1074/jbc.M808500200

Stewart, J.J.P. (2013). Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *Journal of Molecular Modeling, 19*(1), 1–32. doi: 10.1007/s00894-012-1667-x

Sunseri, J., & Koes, D.R. (2016). Pharmit: interactive exploration of chemical space. *Nucleic Acids Research, 44* (Web Server issue), W442–W448. doi.org/10.1093/nar/gkw287

Ton, A.-T., Gentile, F., Hsing, M., Ban, F., & Cherkasov, A. (2020). Rapid identification of potential inhibitors of SARS-CoV-2 main protease by deep docking of 1.3 billion compounds. *Molecular Informatics*. doi.org/10.1002/minf.202000028

Vega-Valdez, I.R., Santiago-Quintana J.M., Rosalez, M.N., Farfan-Garcia, E.D., & Soriano-Ursua, M.A. (2020). Theoretical evaluation of bortezomib and other boron-containing compounds as inhibitors of SARS-CoV-2 main protease. ChemRxiv. doi.org/10.26434/chemrxiv.12047346.v1

Yan, L., Zhang, J., Wang, N., Li, H., Shi, Y., Guo,…& Zou, Q. (2020). Therapeutic drugs targeting 2019-nCoV main protease by high-throughput screening. *BioRxiv*. doi.org/10.1101/2020.01.28.922922
Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., …, & Rao, Z. (2003). The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proceedings of the National Academy of Sciences of the United States of America, 100*(23), 13190–13195. https://doi.org/10.1073/pnas.1835675100

Yao, H., Lu, X., Chen, O., Xu, K., Chen, Y., Cheng, L., & Li, L. (2020). Patient-derived mutations impact pathogenicity of SARS-CoV-2. *MedRxiv*. doi.org/10.1101/2020.04.14.20060160

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld R. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science, 368*(6489), 409-412. doi: 10.1126/science.abb3405

Zhijian, X., Peng, C., Shi, Y., Zhu, Z., Mu, K., Wang, X., & Zhu, W. (2020). 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*. doi.org/10.1101/2020.01.27.921627

Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., & Cheng, F. (2020). Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery, 6*(1). doi.org/10.1038/s41421-020-0153-3