

## Article

# Computer Aided Structure-Based Drug Design of Novel SARS-CoV-2 Main Protease Inhibitors: Molecular Docking and Molecular Dynamics Study

Dmitry S. Kolybalov <sup>1,2,\*</sup>, Evgenii D. Kadtsyn <sup>1,3</sup> and Sergey G. Arkhipov <sup>1,2</sup><sup>1</sup> SRF "SKIF", 630559 Koltsovo, Russia; e.kadtsyn@g.nsu.ru (E.D.K.); arksergey@gmail.com (S.G.A.)<sup>2</sup> Scientific Educational Center "Institute of Chemical Technology", Novosibirsk State University, 630090 Novosibirsk, Russia<sup>3</sup> Voevodsky Institute of Chemical Kinetics and Combustion, 630090 Novosibirsk, Russia

\* Correspondence: d.s.kolybalov@srf-skif.ru

**Abstract:** Severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2) virus syndrome caused the recent outbreak of COVID-19 disease, the most significant challenge to public health for decades. Despite the successful development of vaccines and promising therapies, the development of novel drugs is still in the interests of scientific society. SARS-CoV-2 main protease M<sup>Pro</sup> is one of the key proteins for the lifecycle of the virus and is considered an intriguing target. We used a structure-based drug design approach as a part of the search of new inhibitors for SARS-CoV-2 M<sup>Pro</sup> and hence new potential drugs for treating COVID-19. Four structures of potential inhibitors of (4S)-2-(2-(1H-imidazol-5-yl)ethyl)-4-amino-2-(1,3-dihydroxypropyl)-3-hydroxy-5-(1H-imidazol-5-yl)pentanal (**L**<sup>1</sup>), (2R,4S)-2-((1H-imidazol-4-yl)methyl)-4-chloro-8-hydroxy-7-(hydroxymethyl)octanoic acid (**L**<sup>2</sup>), 1,9-dihydroxy-6-(hydroxymethyl)-6-(((1S)-1,7,7-trimethylbicyclo [2.2.1]heptan-2-yl)amino)nonan-4-one (**L**<sup>3</sup>), and 2,4,6-tris((4H-1,2,4-triazol-3-yl)amino)benzotrile (**L**<sup>4</sup>) were modeled. Three-dimensional structures of ligand–protein complexes were modeled and their potential binding efficiency proved. Docking and molecular dynamic simulations were performed for these compounds. Detailed trajectory analysis of the ligands' binding conformation was carried out. Binding free energies were estimated by the MM/PBSA approach. Results suggest a high potential efficiency of the studied inhibitors.

**Keywords:** gromacs; structure-based drug design; SBDD; SARS-CoV-2; M<sup>Pro</sup>

**Citation:** Kolybalov, D.S.; Kadtsyn, E.D.; Arkhipov, S.G. Computer Aided Structure-Based Drug Design of Novel SARS-CoV-2 Main Protease Inhibitors: Molecular Docking and Molecular Dynamics Study. *Computation* **2024**, *12*, 18. <https://doi.org/10.3390/computation12010018>

Academic Editors: Gennady Bocharov, Fabrizio Mafessoni and Rainer Breitling

Received: 11 October 2023

Revised: 8 January 2024

Accepted: 16 January 2024

Published: 20 January 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

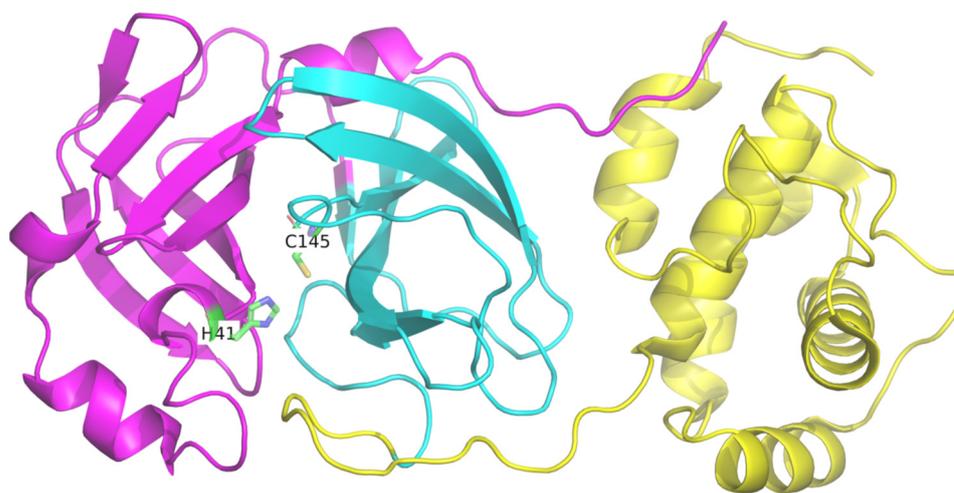
SARS-CoV-2, originally identified at the end of 2019 in Wuhan, China, causes a severe respiratory disease, COVID-19 [1–3]. It spread all over the world and was stated as a pandemic by the World Health Organization in March 2020 [4]. The disease resulted in millions of victims, giant losses in world economics, and became one of the most challenging crises in public health in decades. In May 2023, WHO removed the global health emergency status of COVID-19, however, it noted that a transition from emergency response to long-term disease management is required [5]. SARS-CoV-2 has a high mutability rate [6], and new variants have been documented [7]. However, the danger has not disappeared, and COVID-19 treatment remains a challenge. The development of new safe, efficient, and affordable drugs against COVID-19 is still an important task for scientists around the world.

SARS-CoV-2 belongs to the clade *b* of the genus Betacoronavirus family Coronaviridae [1,2,6]. It means a positive-sense single-stranded RNA virus with a genome size of about 30 kb [2,8,9]. SARS-CoV-2 virion consists of nucleocapsid (N) and envelope (E) proteins and membrane (M) and spike (S) glycoproteins [8,10]. S-glycoproteins form spikes located at the outer surface of the virion. That it is similar to a crown-like shape gives

the virus its name, “corona”. The S-protein binds directly to angiotensin-converting enzyme 2 (ACE2) of the human cell receptor, triggering a mechanism of virus entry into a cell [11–13]. After entering, the virus releases its RNA into the cell cytoplasm, exploiting cell machinery for replication [6]. The viral genome contains two overlapping open reading frames, ORF1a and ORF1ab, which are translated to polyproteins pp1a and pp1ab [14]. The polyproteins’ cleavage during the maturation process produces non-structural proteins (NSPs) necessary for the viral lifecycle [15,16]. The process is mediated by 3-chymotrypsin-like protease (3CL<sup>pro</sup>), also known as the main protease (M<sup>pro</sup>), and papain-like protease (PL<sup>pro</sup>). Several NSPs are released in the process, among them—M<sup>pro</sup> (NSP5), PL<sup>pro</sup> (NSP3), RNA-dependent-RNA-polymerase (RdRp, NSP12), helicase (NSP13) etc. [14,17,18].

Some of these proteins, especially S-protein, RdRp, and M<sup>pro</sup>, are often considered to be targets for drug development [17,19–26]. The main protease plays a key role in virus replication and its inhibition helps to suppress the propagation of disease. Viral protease inhibition is a well-known strategy for the treatment of viral infections such as hepatitis C or HIV [27,28]. M<sup>pro</sup> is known to be highly conservative among coronaviruses in sequence and also its 3D structure has no similarity to any human protease and thus is an excellent target for anti-COVID-19 drug therapy [29–31].

A 3D structure of SARS-CoV-2 M<sup>pro</sup> has been established for unliganded protein, as well as for numerous complexes (more than 700 reports in RCSB PDB by summer 2023, see, e.g., 6LU7 as one of the first entries [17]). Its homodimer is asymmetrical and heart shaped. A protomer has 306 residues and consists of three domains [32], see Figure 1. Domains I (residues 8–101) and II (residues 102–184) are mostly beta strands while domain III (residues 185–306) consists of five alpha helices and is connected with domain II via a long loop. M<sup>pro</sup>’s catalytic dyad, namely His41 and Cys145, is located on an active site in a gap between domains I and II, see Figure 1. Cysteine residue Cys145 acts as a nucleophile, while histidine His41 is a proton acceptor [17,33]. Nucleophilic attack of the S atom of Cys145 on the main chain carbonyl C atom of pp1a or pp1ab leads to the formation of acyl-enzyme intermediate and the subsequent proteolytic cleavage; in total, 11 conserved cleavage sites are known [17,34].



**Figure 1.** A model of the 3D structure of SARS-CoV-2 M<sup>pro</sup> (PDB code 7NG6). Cartoon style is used for the protein, domain I is magenta, domain II is cyan, and domain III is yellow. A catalytic dyad is represented in licorice style with elementwise CPK coloring. Ligands and water molecules were removed from the figure for clarity.

A lot of effort was put into suggesting suitable M<sup>pro</sup> inhibitors during the COVID-19 pandemic, and numerous candidates were tested (e.g., [18,30,33,35–39]). One of the most famous inhibitors is N3 (PubChem CID 6323191) [17]. Many studies focus on repurposing known antiviral drugs [36,40–45], or on studying natural compounds [46–50]. Such an approach has several advantages, namely, reducing the chance of candidate failure or

shortening the development time and cost. However, success with treating a new viral disease is not guaranteed in this way, and side effects still might be rather serious. Examples of remdesivir and hydroxychloroquine are well known [43–45,51–53]. During a large-scale outbreak, repurposing could indeed be considered as a right strategy, because saving development time meant saving human lives. However, a rational drug design seems to be a more appropriate strategy.

Computational structure-based drug design (SBDD) is a powerful approach, which has been already shown to be promising in developing drugs against SARS-CoV-2 as well as against other viruses [54–59]. For example, one of the inhibitors against SARS-CoV-1, namely N3, was developed by drug modeling. Computational drug design may include QSAR, screening techniques, or building candidate molecules directly in the target active sites. What is more, docking and molecular dynamics are used for validation of results.

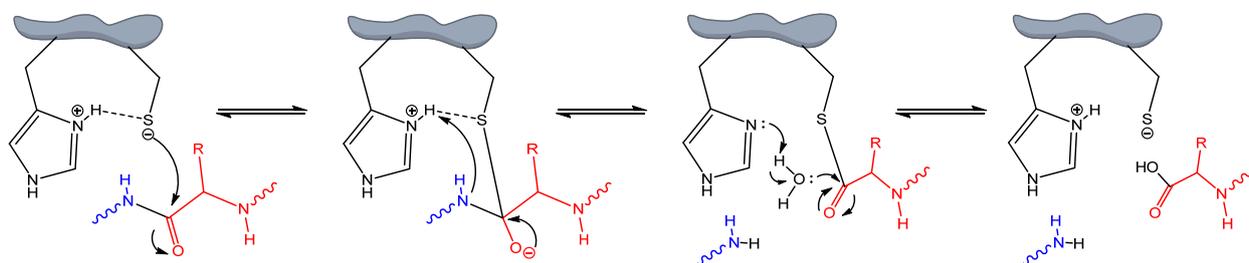
In this work, we used a 3D structural model of M<sup>Pro</sup> (PDB code 7NG6 [60]), which contains information about its active site and cleavage mechanism, in order to design novel candidate molecules. Different candidates represent different electrophilic warheads to bind catalytic Cys145 residue of M<sup>Pro</sup>, the side fragments of the molecules were chosen to fill known M<sup>Pro</sup> recognition pockets. The size and shape of candidates' side fragments were refined using the 3D structure of M<sup>Pro</sup>—for better structure-to-structure correspondence to recognition pockets. Docking was used to find the complex structure. Molecular dynamics and MM/PBSA techniques were used to check configurations of complexes, their stability, and binding free energies. Results show the stability of all the complexes, and binding free energies are high enough to provide a large value of the inhibition constant. The conclusion has been drawn: the novel structures show potential to further in vitro and in vivo research. SBDD also allows us to predict some improvement to the candidates' structure that can be made.

## 2. Materials and Methods

### 2.1. Molecular Design

Molecular modeling of four ligands was carried out by the standard structure-based drug design (SBDD) technique [61]. First, a molecular target was chosen and the active site was found. In this study, the main protease of SARS-CoV-2 (M<sup>Pro</sup>, 3CL) was chosen as a target, see above.

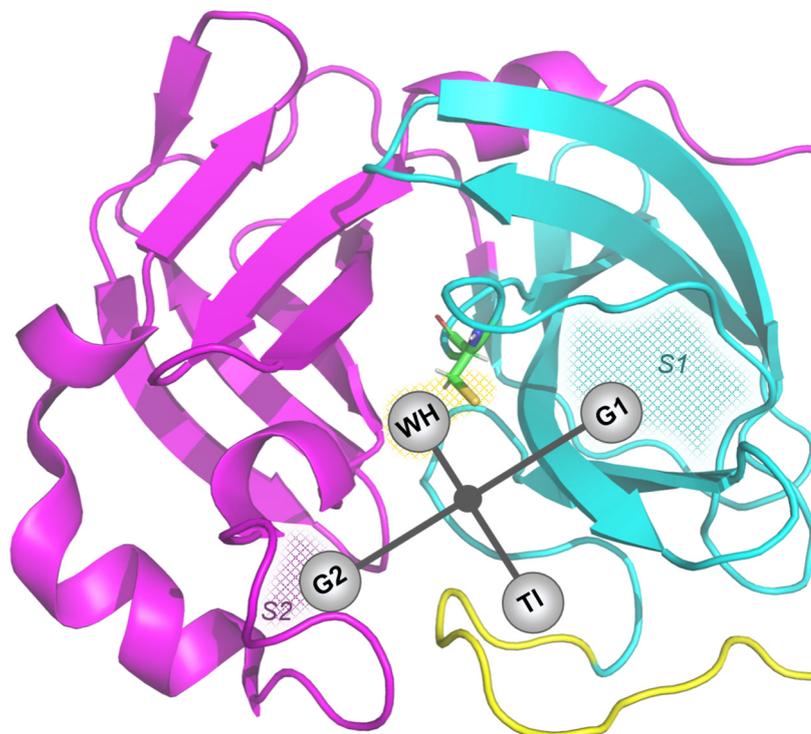
The catalytic mechanism of M<sup>Pro</sup> is known [62] and based on the Cys145 sulfur atom nucleophilic attack at the carbonyl carbon atom of the target protein main chain (Figure 2). Covalent bonding of the cysteine residue blocks the active site and leads to the inactivation of protease. Deprotonated cysteine is highly nucleophilic, one should hence consider the electrophilic group as a warhead (i.e., functional groups, which create a covalent bond with the amino acid, involved in biocatalysis) of any inhibitor.



**Figure 2.** Mechanism of cysteine protease action [62].

M<sup>Pro</sup>'s recognition mechanism is provided mainly by two pockets, **S1** and **S2**, located on both sides of the peptide binding groove between domains I and II (Figures 1 and 3) [63–66]. **S1** belongs to domain II and is located near the catalytic dyad; glutamine residue is always found at the first cleavage position **P1**. Phe140 and His163 provide H-bonds with the glutamine side amide group while Ser144 and Glu166 are responsible for main chain binding. The **P2** position

is more variable, but in most cases, it is a leucine. The **S2** recognition pocket is located mostly in domain I, which includes His41 residue in its part. The pocket is rather hydrophobic, but H-bonds with His41 or with Gln189 are also possible.



**Figure 3.** A general scheme of the ligand at the binding site. Domains I and II of  $MP^{pro}$  are shown as in Figure 1, catalytic Cys145 is shown by sticks and colored elementwise. **WH** is the warhead located near Cys145, **G1** and **G2** are recognition groups occupying **S1** and **S2** pockets, respectively, and **T1** is an optional tail. A dark circle is an attachment point connecting the ligand's parts.

The inhibitor should occupy both **S1** and **S2** pockets to bind with the protease. A general scheme of the ligand molecule is depicted in Figure 3. **WH** here means a warhead group, which should be located near Cys145, **G1** and **G2** are substituents designed to fill **S1** and **S2** pockets, respectively, and **T1** is an optional tail that can fill the groove or be just an artifact of a molecule's synthesis. The size of **G1** should be similar to the one of glutamine, i.e., the chain length varies from 3 to 5 carbons with single atomic ending functional groups or 1–2 carbons if the ending group is somewhat larger. Ending groups provide H-bonding with Phe140 and/or with His163 at the bottom of the pocket, other possible groups can bind with Ser144 and Glu166 on its entering. **G2** should be either a hydrophobic group or an H-bond aromatic donor/acceptor forming an H-bond with His41 or Gln189. An attachment point (the black circle in Figure 3) could be a tertiary carbon atom as well as a small chain group (1–3 carbons) or a 5/6-membered cycle. Having such a scaffold, inhibitor will be similar in size to the first amino acid dyad of the cleaved peptide and will fill the binding position. Using  $MP^{pro}$ 's 3D structure, one can refine chain lengths manually by any molecular builder like the Avogadro locating inhibitor in the required position and adding/removing carbon atoms until the desired distances are obtained.

Warheads are small nucleophilic groups that should be stable in physiological conditions. Four groups are commonly used in drug design, namely aldehyde, ketone, halogenide, and nitrile. Considering recognition functional groups, one should follow Lipinski's rule of 5: no more than 5 H-bond donors, no more than 10 H-bond acceptors, octanol–water coefficient of no more than 5, and a molecular weight not over 500 Da [67]. In the recent work of Kenny [68], H-bond donors commonly used in drug design were considered. Following this work, we considered using a hydroxyl group (OH) and amino groups (NH<sub>2</sub>,

NH) as strong H-donors as well as imidazolyl ( $C_3H_3N_2$ ) and triazolyl groups ( $C_2H_2N_3$ ), which can act as both a H-donor and H-acceptor. The last groups are suitable for the **G2** position due to their aromaticity. We considered also the camphor-based substitute as a hydrophobic group for the **G2** position due to the known antiviral effect of camphor derivatives and their activity regarding viruses of the Coronaviridae family, which was separately noted [69].

Combining all those reasonings, we can see that there are tens of possibly suitable scaffolds and hundreds of functional group combinations that give us a decision space of thousands of potential candidates. The common method in such a case is the creation of a candidate library covering the whole space followed by a screening procedure. It is clear, however, that most of the variants would be inadequate from the chemical point of view. The other important note is that the candidate molecule should have a rather simple synthetic scheme, otherwise, its drug significance will be poor. Hence a systematic search over the whole library has a low meaning and manual structure selection is required anyway. We decided to generate structures manually, trying to build them to be highly different from each other, thus covering possibly far areas in the decision space. Structure building followed the next principles: we built four structures,  $L^1$ – $L^4$ , one for each warhead (aldehyde, chloride, ketone, and nitrile); we divided the structures into two sets, with OH ( $L^1$ – $L^3$ ) and with heterocycle ( $L^1$ ,  $L^2$ ,  $L^4$ ) as the main H-donor, we used tertiary carbon ( $L^1$ ), small chains ( $L^2$ ,  $L^3$ ) and cycle ( $L^4$ ) as an attachment point; at last, we tried many ( $L^1$ ,  $L^2$ ) groups inside **G1** to bind with both the bottom and entry of **S1** and one group ( $L^3$ ,  $L^4$ ) group to bind with the bottom only. While  $L^3$  contains no heterocycles, the camphoryl group was used as **G2**. Tail **T1** was chosen to obtain synthetically accessible structures and increase the structure solubility.

The following structures were obtained (see Table 1).

For  $L^1$  the following combination was obtained: **G1** is a three-carbon chain having  $NH_2$  and OH side groups to bind with Ser144 or Glu166 at the entering of **S1** and imidazole ending group to make the H-bond with Phe140 or His163 at the **S1** bottom (see Table 1); **G2** is an imidazole ring at the end of a three-carbon chain; the warhead is aldehyde; tertiary carbon is the attachment point; tail **T1** improves the structure solubility by OH groups.

In  $L^2$  a chlorine warhead is used, **G1** is a three-carbon chain with double  $CH_2OH$  endings, **G2** is an imidazole ring, and carboxyl **T1** should increase structure solubility. The attachment point here is the three-carbon chain.

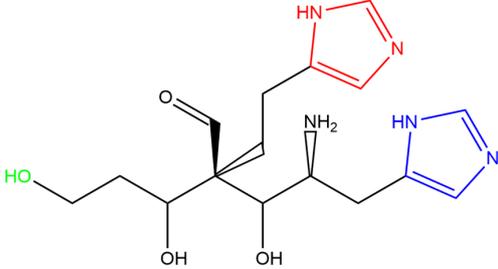
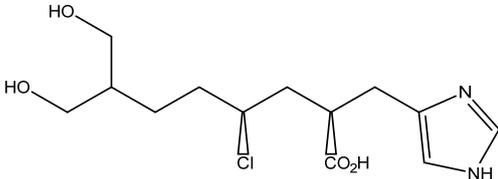
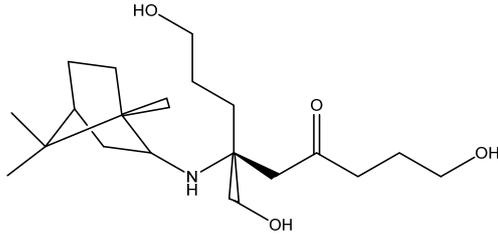
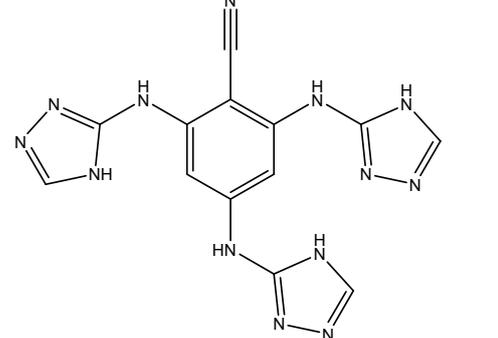
In  $L^3$  a ketone warhead is used, the attachment point is the two-carbon chain, and no heterocycles are used. **G1** is a three-carbon chain with an OH end, **G2** is a camphoryl group, and the **T1** tail contains OH groups for solubility increase.

$L^4$  has a nitrile warhead and cyclic attachment point, in contains no OH groups, thus both **G1** and **G2** are heterocycles. We choose triazole cycles to increase H-bonding ability and attach a third ring (see Table 1) as a tail to provide a possibility of binding by any of the heterocycles.

It should be emphasized that such a “chaotic” structure construction method gives not the best ligands, but rather allows one to find suitable structures for further improvement by analyzing protein–ligand interactions. If some of the constructed molecules show high binding energy values, then further structure refinement can be carried out. Such a refinement includes a small variation of the molecular scaffold (adding/removing 1–2 carbon atoms) or changing some of the functional groups. MD analysis of ligand–protein interactions gives us information about necessary modifications. One can consider such a kind of structure design as a form of Monte-Carlo simulation annealing rather than brute force search. We will show below that  $L^2$  and  $L^3$  are not the best candidates,  $L^4$  required some modifications, but  $L^1$  has an almost suitable geometry. However experimental verification of the results obtained is required to make further structure modification meaningful.

To reach the energy minimum, geometric optimizations of potential inhibitor molecular fragments were carried out using the Gaussian 16 package.

**Table 1.** Abbreviation, structures, and binding groups of (4S)-2-(2-(1H-imidazol-5-yl)ethyl)-4-amino-2-(1,3-dihydroxypropyl)-3-hydroxy-5-(1H-imidazol-5-yl)pentanal (**L**<sup>1</sup>), (2R,4S)-2-((1H-imidazol-4-yl)methyl)-4-chloro-8-hydroxy-7-(hydroxymethyl)octanoic acid (**L**<sup>2</sup>), 1,9-dihydroxy-6-(hydroxymethyl)-6-(((1S)-1,7,7-trimethylbicyclo[2.2.1]heptan-2-yl)amino)nonan-4-one (**L**<sup>3</sup>), and 2,4,6-tris((4H-1,2,4-triazol-3-yl)amino)benzotrile (**L**<sup>4</sup>).

Abbreviation	Structure	Binding Groups	Amino Acids Residue
<b>L</b> <sup>1</sup>		5-(1H-imidazol-5-yl) (blue) 2-(1H-imidazol-5-yl) (red)	Ser144, His163 Gln189
		OH-group (green)	His41
<b>L</b> <sup>2</sup>		imidazolyl ring carboxyl group	His41, His164 Met165
		OH-groups	Ser144, His163, Glu166
<b>L</b> <sup>3</sup>		main OH-containing group with CO short OH-containing group without CO camphoryl group	Glu166, Met165 His163 located in S2 groove
		OH-containing group without CO	Asp142
<b>L</b> <sup>4</sup>		triazolyl rings in orto-position	His41
		triazolyl ring in para-position	Ser144

## 2.2. Molecular Docking

Molecular docking was performed using the GOLD 2022.3 by CCDC software [70]. The crystal structure of a complex of the main protease dimer with an MG-132 covalent inhibitor (PDB code 7NG6) was used for molecular docking. Ligands and water molecules in the protein structure were removed, and the protonation state was identified using Hermes' protonation tool by the CCDC 2022.3 software.

The geometry of potential inhibitors was optimized by using the Gaussian 16 package [71], which was carried out with the B3LYP functional and def2-TZVP basis set. The IR vibrational frequencies were calculated, no imaginary values were found, proving that a true minimum was reached. Calculation results were visualized and analyzed with ChemCraft v.1.8 graphical software [72].

Four ligands were docked in the active sites of one of the M<sup>Pro</sup> protein chains in a 10 Å spherical cavity (center of cavity −17.909, −2.355, −27.038) with the ASP scoring function; the ChemScore scoring function was used to rescore. For each ligand, 200 docking runs were carried out; the search efficiency parameter of GOLD was set to 200%.

Docking results were visualized and analyzed with Hermes. The selection of leader poses for each compound was achieved by summing the ASP and ChemScore values. Poses with the highest summing score were chosen to be leaders, wherein a negative ASP or ChemScore value for the docking pose was considered to be unacceptable and was removed from further analysis. BIOVIA Discovery Studio [73] was used to identify the interactions between selected ligands and the M<sup>Pro</sup> protein.

### 2.3. Molecular Dynamic Simulation

Molecular dynamic (MD) simulation of protein–ligand complexes and free protease was carried out using Gromacs 2023 [74]. The topology of SARS-CoV-2 M<sup>Pro</sup> was prepared by the Gromacs 2023 tools with an OPLS-AA/L force field; ligands were parameterized using the LigParGen web-based service [75]. Protein–ligand complexes were placed in a periodic cubic box with a minimum distance of 1 nm between the complex and the box wall. Then a complex was dissolved in SPC/E water [76]. Negative overcharge was neutralized by replacing 8 water molecules by Na<sup>+</sup> ions. The system energy minimization was achieved by the steepest descent algorithm with a convergence criterion of 10.0 kJ/mol. After that, the system was equilibrated during 100 ps in an NVT ensemble, and then again during 100 ps in an NPT ensemble. Position restraints for protein and ligand heavy atoms were used at both equilibration stages. Two temperature coupling groups were used: protein–ligand complex and water with ions. Temperature was maintained by a modified Berendsen thermostat at 300 K [77], and pressure was maintained by C-rescale barostat at 1 bar [78]. H-bonds were constrained using the LINCS algorithm [79]. The production run was performed in a NPT ensemble, a constant temperature of 300 K (Berendsen thermostat) and a constant of 1 bar pressure (Parrinello-Rahman barostat [80]), the trajectory length was 100 ns with an integration step of 2 fs and recording step of 10 ps. PME scheme was used for long-range electrostatics [81]; the cut-off radius for short-range interactions was 1.0 nm.

Molecular dynamic trajectories were analyzed using the tools of the GROMACS suite, including root mean square deviation (RMSD), root mean square fluctuation (RMSF), solvent-accessible surface area (SASA), number of hydrogen bonds, and radius of gyration (Rg).

### 2.4. MM/PBSA

Binding free energies for M<sup>Pro</sup>-ligand complexes were estimated using the MM/PBSA approach with the aid of the gmxpbsa tool [82]. The binding free energy is defined as

$$\Delta G = \Delta E_{\text{MM}} + \Delta G_{\text{polar}} + \Delta G_{\text{apolar}} - T\Delta S, \quad (1)$$

$E_{\text{MM}}$  term is calculated for an instant configuration using force-field parameters from MD simulation,  $G_{\text{polar}}$  is calculated from the Poisson–Boltzmann equation with an environment dielectric constant of 80 and inner constant of 1, and  $G_{\text{apolar}}$  is calculated from the surface area value.  $\Delta$  means a difference between the value for the complex and the sum of values for the separated protein and ligand. The  $T\Delta S$  term is usually neglected. For each trajectory, the last half (50–100 ns) was extracted for MM/PBSA analysis, frames were taken with a timestep of 250 ps, and an averaging over 200 independent configurations was thus applied. For each frame all three terms ( $\Delta E_{\text{MM}}$ ,  $\Delta G_{\text{polar}}$ , and  $\Delta G_{\text{apolar}}$ ) as well as each amino-acid contribution were estimated.

### 3. Results

#### 3.1. Molecular Docking

Potential inhibitors were docked into the active site of one of the two polypeptide chains in the GOLD program package. ASP and ChemScore functions were used to estimate the binding interactions. Configurations with the highest sum value of the two scoring functions were selected for each structure. The structures obtained were used for molecular dynamic simulation with the Gromacs program package. The sum of the two scoring functions was used to compare the interaction strengths for the highest scoring configurations (Table 2).

**Table 2.** Non-dimensional values of scoring functions for L<sup>1</sup>–L<sup>4</sup>.

	L <sup>1</sup>	L <sup>2</sup>	L <sup>3</sup>	L <sup>4</sup>
ASP	30.2693	27.4447	23.2161	36.1228
ChemScore	1.9658	6.3231	7.6813	15.7884
Sum	32.2351	33.7678	30.8974	51.9112

BIOVIA Discovery Studio was used for the interaction analysis of a ligand in the M<sup>Pro</sup> active site. L<sup>1</sup> forms two hydrogen bonds with His164 and Asn142, weak  $\pi$ -sulfur and  $\pi$ -alkyl interactions and one unfavorable bump between Gln189 and the imidazole fragment of the ligand. L<sup>2</sup> forms four conventional hydrogen bonds with His163, His164, and Asp187 of the A chain and Ser1 of the B chain, the imidazole fragment has  $\pi$ - $\pi$  stacking interaction with His41, Met49, and Cys44 showing weak  $\pi$ -sulfur and  $\pi$ -alkyl interactions, respectively. L<sup>3</sup> forms four conventional hydrogen bonds with Phe140, Asn142, His164, and Glu166, as well as two carbon-hydrogen bonds and one  $\pi$ -hydrogen bond. L<sup>4</sup> builds three strong hydrogen bonds with Ser144, His163, and Glu166, three  $\pi$ -sulfur interactions with Cys44, Met165, and Cys145, as well as weak carbon-hydrogen and  $\pi$ -alkyl interactions. It is interesting to note, that L<sup>1</sup>–L<sup>3</sup> contain large fragments, causing some difficulties during docking (Figure 4).

#### 3.2. Molecular Dynamics

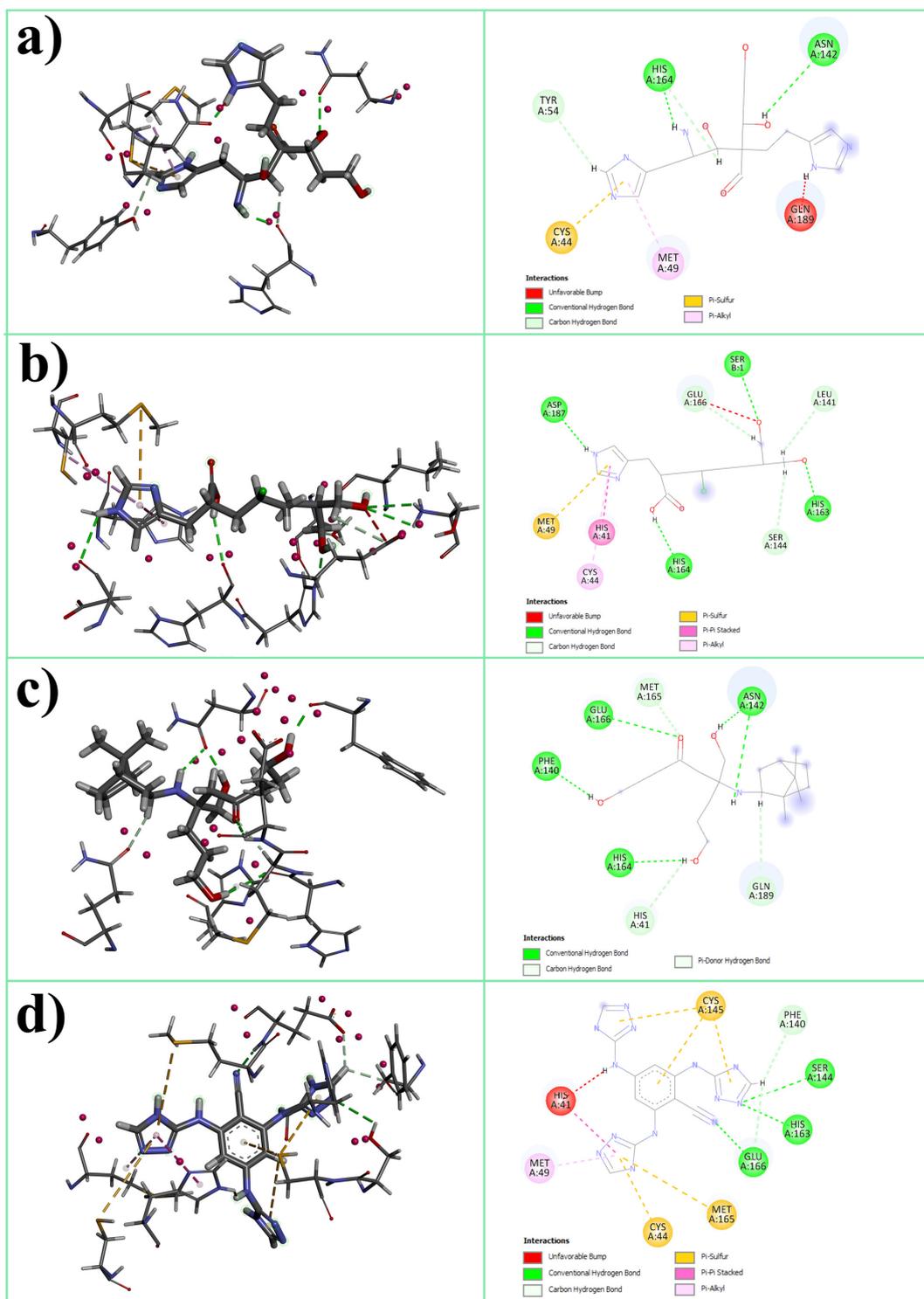
Molecular docking is a widely used method, which gives the configuration of the protein–ligand complex and lets us estimate binding energy. The method, however, has some restrictions, such as the inability to take into account possible effects of solvent, and neglecting the possibility of conformational dynamics of the ligand and the protein active site during the docking process. Molecular dynamics (MD), in contrast, takes these effects into account and thus complements molecular docking.

In this work, MD of unliganded protein and of protein–ligand complexes was performed during 100 ns for each system. Root mean square deviation (RMSD), root mean square fluctuation (RMSF), solvent-accessible surface area (SASA), radius of gyration (Rg), and H-bonds analysis were used, in order to check the stability of the model system.

##### 3.2.1. Root Mean Square Deviation

Unliganded protein and protein–ligand complexes, which have a higher docking score, were simulated in a water box for 100 ns, in order to describe better the interactions between the protein and the ligand. Trajectories were analyzed by using RMSD and RMSF to check the stability of the complexes. RMSD values were calculated for protein C $\alpha$  atoms and for all ligand atoms. Mean RMSD values for protein and ligands in all systems are listed in Table 3. It was observed that the RMSD value of protein in four protein–ligand complexes and of the unliganded protein fluctuates around 0.2 nm for all of the systems, indicating their stability (Figure 5a). For all ligands, except L<sup>1</sup>, the mean RMSD value for the protein was slightly lower than the value for unliganded protein, suggesting some stabilization. For L<sup>4</sup>, the decrease in RMSD was rather high, about 20%. The RMSD value of the ligands fluctuated around 0.2 nm (Figure 5b), indicating the stability of the ligands at the active site.

Figure 5b illustrates that  $L^4$  is the most stable one among other ligands. For  $L^2$  and  $L^3$  some large fluctuations were observed. One can suggest that the ligand underwent significant conformational changes. Visual analysis supports this viewpoint, see Section 4 Discussion.

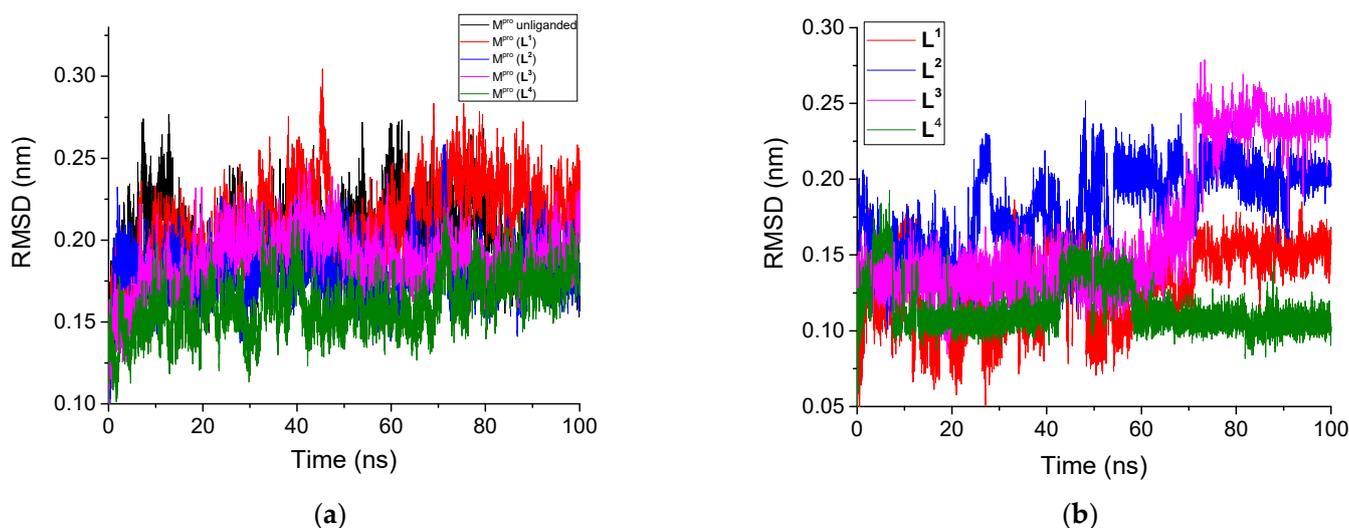


**Figure 4.** The 3D structures (left) and 2D interaction maps (right) for  $MP^{pro}$  in complexes with the inhibitor in the active site after docking: (a)  $L^1$ , (b)  $L^2$ , (c)  $L^3$ , (d)  $L^4$ . Dashed lines designate interactions in both 2D and 3D images, reds are repulsive, greens are H-bonds, yellows are sulfur atoms with  $\pi$ -electrons, magentas are pure electronic.  $MP^{pro}$ 's residues are labeled circles in the right pictures and represented by sticks in the left ones; red spheres on the left are lone pairs.

**Table 3.** The average values of RMSD, RMSE, SASA, and number of H-bonds for protein–ligand complexes.

System	RMSD, nm		RMSE, nm	$R_g$ , nm	SASA, nm <sup>2</sup>		H-Bonds Number	
	Protein	Ligand			Protein	Buried *	Intra-Protein	Ligand-Protein
Unliganded M <sup>Pro</sup>	0.201	-	0.109	2.58	261.49	-	452.8	-
M <sup>Pro</sup> + L <sup>1</sup>	0.211	0.130	0.107	2.59	263.64	6.66	425.5	4.3
M <sup>Pro</sup> + L <sup>2</sup>	0.180	0.181	0.100	2.59	262.58	6.97	430.7	2.3
M <sup>Pro</sup> + L <sup>3</sup>	0.188	0.167	0.096	2.59	262.03	6.93	424.6	2.4
M <sup>Pro</sup> + L <sup>4</sup>	0.161	0.115	0.088	2.59	264.03	7.53	433.1	3.2

$$* \text{SASA}_{\text{buried}} = \text{SASA}_{\text{protein}} + \text{SASA}_{\text{ligand}} - \text{SASA}_{\text{complex}}$$

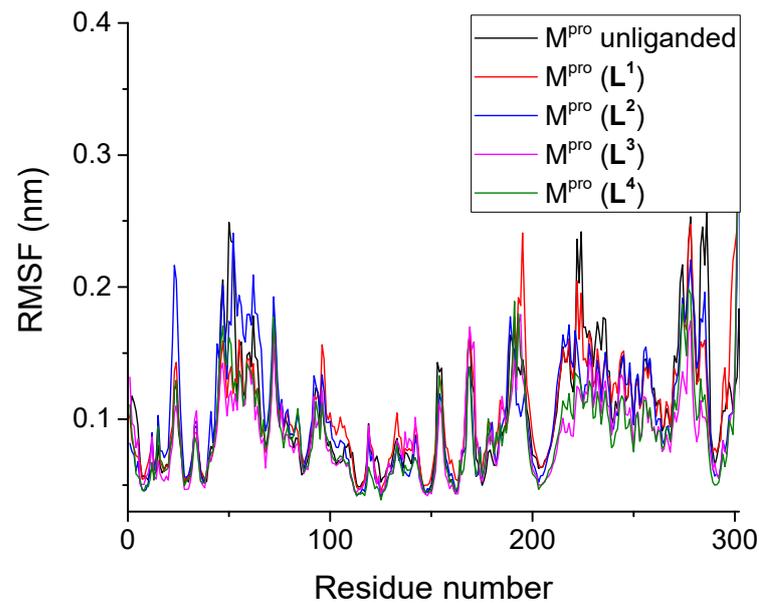
**Figure 5.** Root mean square deviation values for (a) C $\alpha$  atoms of unliganded M<sup>Pro</sup> and four protein–ligand complexes, (b) all atoms of four ligands for protein–ligand complexes.

### 3.2.2. Root Mean Square Fluctuation

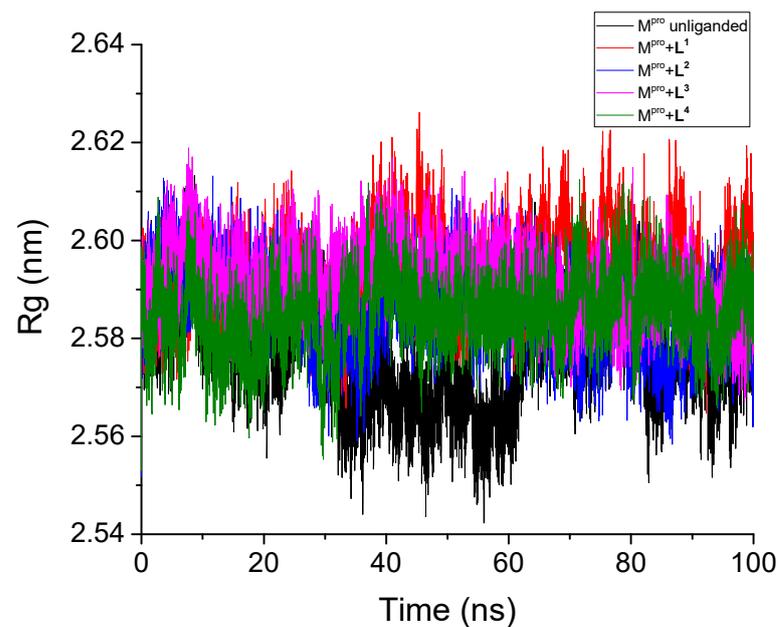
The C $\alpha$  atoms' RMSF was used to calculate the fluctuation of each amino acid of the protein (Figure 6). A higher RMSF value means a higher amino acid fluctuation, while a lower RMSF value means a higher stability of the residue. The average value of RMSF was calculated to be 0.109 for unliganded M<sup>Pro</sup>, 0.107 for M<sup>Pro</sup> + L<sup>1</sup>, 0.100 for M<sup>Pro</sup> + L<sup>2</sup>, 0.096 for M<sup>Pro</sup> + L<sup>3</sup>, and 0.088 for M<sup>Pro</sup> + L<sup>4</sup> (Table 3). In general, RMSF plots for each complex were similar, as well as for unliganded protein. We can speculate that ligands do not perturb the protein stability during the 100 ns MD simulation. Moreover, a comparison of average RMSF values showed that the presence of a ligand led to some stabilization of the protein, especially in the case of L<sup>4</sup>.

### 3.2.3. Radius of Gyration

The radius of gyration is a measure of protein size and compactness. Mean  $R_g$  values for all studied systems are listed in Table 3.  $R_g$  time dependencies are shown in Figure 7. It is seen that there is no variance between different complexes and between the complexes and unliganded protein, so the geometry of M<sup>Pro</sup> is the same for all systems. No significant changes occurred during the simulation, only a small decrease in  $R_g$  could be noted for pure M<sup>Pro</sup> at a time interval of 30–60 ns.



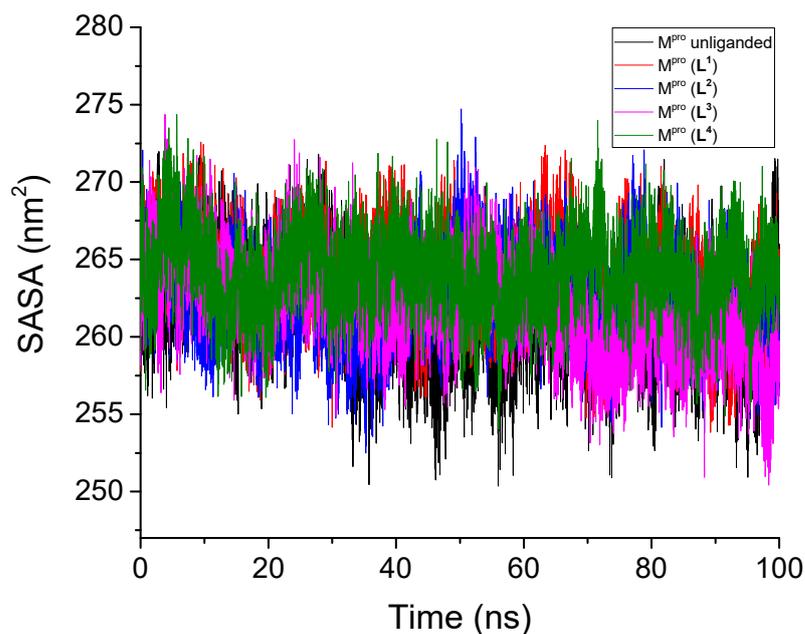
**Figure 6.** Root mean square fluctuation values for C $\alpha$  atoms of unliganded M<sup>pro</sup> and four protein–ligand complexes.



**Figure 7.** Radius of gyration values for unliganded M<sup>pro</sup> and of the chains in four protein–ligand complexes.

### 3.2.4. Solvent-Accessible Surface Area

Solvent-accessible surface area (SASA) is the surface area where the protein or a ligand interacts with solvent molecules. It correlates with the interactions between the complex and the solvent during the simulation analysis. The SASA values for unliganded M<sup>pro</sup>, four protein–ligand complexes, as well as for four ligands, were calculated and plotted vs. simulation time (Figure 8); the corresponding mean values are listed in Table 3. A change in SASA value can be referred to as a change in protein conformation. What is more, SASA values are stable, and no drastic changes occur. A comparison of the average values for different systems revealed that they are not different from each other.



**Figure 8.** Protein solvent-accessible surface area values for unliganded  $M^{pro}$  and four protein–ligand complexes.

Buried SASA is the surface area where a ligand interacts with the protein. It can be considered as a measure of protein–ligand interaction. Buried SASA was calculated using the expression:

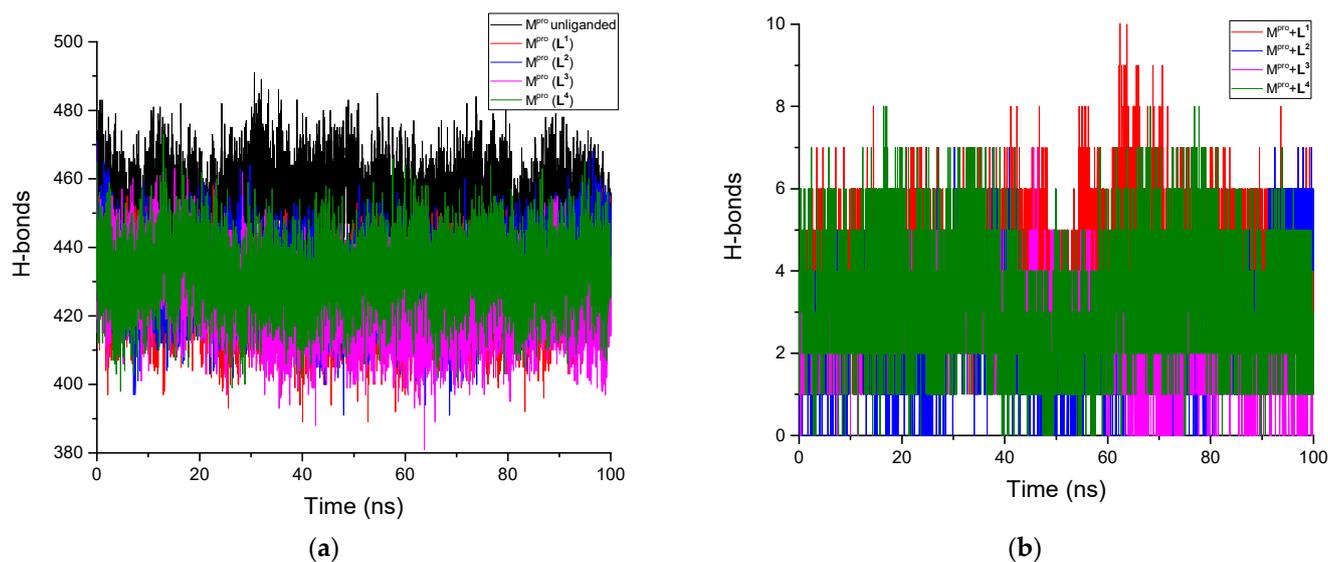
$$SASA_{buried} = SASA_{protein} + SASA_{ligand} - SASA_{complex} \quad (2)$$

The mean values are listed in Table 3. It is seen that  $L^2$  and  $L^3$  values are almost the same, while the value for  $L^1$  is slightly lower, and that for  $L^4$  is much higher. The latter fact can be related to the geometry of the ligand molecules and the presence of four cycles. One can suggest that  $L^4$  is pressed down to  $M^{pro}$  more tightly than other ligands. A low SASA value for  $L^1$  should point out the exposure of  $L^1$  from protein to solvent and a less tight protein–ligand contact.

### 3.2.5. H-Bonds Analysis

Hydrogen bonds (H-bonds) are the key interactions that are responsible for preserving the protein structure and for protein–ligand binding. A notable change in the number of inter-protein H-bonds usually correlates with the change in ternary protein structure. Time dependencies of the number of inter-protein H-bonds are shown in Figure 9a; average numbers are listed in Table 3. The average number can be expected to be lower for a ligated system, because some protein H-bond donors and acceptors switch to binding with the ligand. However, no dramatic changes occur (Figure 9a), and it confirms that the ternary structure does not change on ligation.

Figure 9b shows the time dependencies of the number of protein–ligand H-bonds; average numbers are listed in Table 3. The  $L^1$  forms the largest number of H-bonds, which is related to a large number of donor and acceptor groups, and probably with their favorable spatial distribution.  $L^4$  shows also a large number of H-bonds, whereas the binding of  $L^2$  and  $L^3$  is much poorer.



**Figure 9.** Number of H-bonds: (a) interprotein, (b) protein–ligand.

### 3.3. Binding Free Energy

M<sup>Pro</sup>–ligand binding free energies were calculated as described in MM/PBSA; the results are listed in Table 4. Molecular mechanic contributions (van der Waals and electrostatic terms) are always negative, whereas the interaction energy is always higher at short distances than at longer ones. The polar term trend was positive, solvation had led to lowering of the energy. It was more significant for the separated protein and a ligand than for their complex due to a larger solvation area. The apolar term is negative as long as cavity creation costs energy. The effect was larger for the separated protein and a ligand than for their complex due to a larger cavity area. Note that the absolute value of the apolar term is always small, and does not influence the results significantly.

**Table 4.** Mean binding free energies (kJ/mol) of ligands and their contributions \*.

Energy Term, kJ/mol	$E_{vdW}$	$E_{elec}$	$E_{pol}$	$E_{apol}$	$E_{tot}$
System					
L <sup>1</sup>	−141.203 ± 2.142	−94.304 ± 2.490	167.428 ± 2.311	−16.778 ± 0.135	−84.858 ± 2.662
L <sup>2</sup>	−130.295 ± 2.078	−61.892 ± 3.220	143.489 ± 2.465	−17.239 ± 0.152	−65.937 ± 3.063
L <sup>3</sup>	−125.815 ± 1.630	−37.209 ± 2.021	108.833 ± 1.962	−17.010 ± 0.158	−71.201 ± 2.110
L <sup>4</sup>	−164.479 ± 1.641	−45.442 ± 1.985	134.414 ± 1.843	−18.085 ± 0.137	−93.592 ± 1.994

\* Mean values were calculated over 200 configurations recorded from 50 ns to 100 ns with a step of 250 ns. A confidence interval was calculated by the standard approach for a confidence level  $\alpha = 0.95$ , Student’s *t* value of 1.972 was used.

For the L<sup>1</sup> ligand, the values of both polar and electrostatic terms were the highest, which can be related to the large number of highly charged oxygen and nitrogen atoms and to the formation of hydrogen bonds. The values for all contributions of the L<sup>2</sup> ligand, except for the apolar one, were smaller, possibly due to its elongated form with a small number of charged atoms. This combination had led to the lowest value of total binding energy. L<sup>3</sup> had the smallest value of both electrostatic and polar terms, apparently due to the presence of a large hydrophobic camphoryl group; its van der Waals interaction energy was also smaller than that for other ligands. It is interesting, however, that a reduction of both electrostatic attractive and polar repulsive terms leads in combination to a rather high total interaction energy, higher than that in the L<sup>2</sup> case. Finally, an L<sup>4</sup> ligand had the largest value of van der Waals contribution, due to the presence of four aromatic cycles and

a large value of  $SASA_{buried}$ . Due to the same reason, it had the largest value of the apolar contribution. At the same time, it had low values of electrostatic and polar contributions, lower than the other ligands, except for  $L^3$ . Due to a large value of the van der Waals term, it had the highest binding energy among all the ligands and can, therefore, be considered the most promising inhibitor.

We note that docking results (Table 2) lead to the same conclusion: the highest score value was obtained for  $L^4$ . For other ligands, however, the docking score values were close to each other and suggested another binding priority order. The fact is most probably related to the poor estimation of solvent effects by the docking technique; the solvation is less important for  $L^4$  than for other ligands. On the other hand, docking can give a rather good estimate of the van der Waals interaction energy, which is the most important contribution in the case of  $L^4$ ; therefore, the docking and MM/PBSA results match.

Per residue contribution analysis revealed two amino acid blocks, Leu141-Cys145 and His163-Met165, in the  $L^1$  binding. Also, Ser46 and Met49 showed strong attraction to the ligand (interaction energy lower than  $<-2.5$  kJ/mol). At the same time, Glu166 showed a high repulsive interaction (about 17 kJ/mol).

In the  $L^2$  case, binding energies were high for Phe140, Asn142, Cys145, Met165, and Ser46. For His41 the energy was about  $-14$  kJ/mol, and the values of the interaction energies for Leu27 and Thr25 were also significant. Again, Glu166 manifested a high repulsive interaction with the ligand, His163 had an antibonding interaction, and, interestingly, a repulsive interaction was observed with Ser1 from the second chain of the dimer.

For the  $L^3$  ligand, no repulsive interactions were found, apparently due to its hydrophobic nature and low electrostatic repulsion. Attraction energies were also lower than in other cases. The highest interaction value was found with Met165 (about  $-10$  kJ/mol); notable interaction with the ligand ( $<-2.5$  kJ/mol) was also identified for Gly143, Cys145, Gln189, Thr25, Leu27, Ser46, Glu47, and Met49.

For the  $L^4$  ligand, the only repulsive interaction was found with Ser144. A very strong attraction ( $<-10$  kJ/mol) was found with His41, Met49, Met165, Asp187, and Gln189. Among other highly interactive amino acids (energy lower than  $-2.5$  kJ/mol) were Cys44, Asp48, Lue50, Pro52, Glu166, and Arg188. Note that no catalytic Cys145 residue was found among them.

#### 4. Discussion

The availability of the crystallographic structure for SARS-CoV-2  $M^{pro}$  makes an implementation of computer-aided drug design possible. In this work, the catalytic site of protease was targeted both in ligand building and docking. Docking analysis suggested the following order of ligand affinity:  $L^4 \gg L^2 > L^1 > L^3$ . The analysis of both 3D docked configurations and the 2D docking interaction map revealed that all ligands occupy both **S1** and **S2** recognition pockets, closing thus the catalytic site.

Molecular dynamic simulation proved that all complexes are stable. Protein RMSDs showed similar values for pure  $M^{pro}$  and complexes, there were no significant deviations from mean values along the trajectory. The absence of protein conformation changes was also seen on Rg and SASA. RMSF showed similar behavior for all systems, which meant the similarity of secondary and ternary structures for  $M^{pro}$  in all complexes as well as for the unliganded protein. The hydrogen bond analysis of the protein also revealed no significant differences between the systems. Altogether, these results showed that all ligands can form stable complexes with  $M^{pro}$ .

Ligand RMSD demonstrated the stability of  $L^1$  and  $L^4$  in the binding position and the existence of at least two binding positions for  $L^2$  and  $L^3$ ; two shells were observed for both RMSD time dependencies. MM/PBSA showed also highly negative binding energy values for  $L^1$  and  $L^4$ , which explains their relative stability. The nature of such stability differs for  $L^1$  and  $L^4$ . In the case of  $L^4$ , based on a large  $SASA_{buried}$ , as well as a large van der Waals contribution to the total binding energy, one can conclude that  $L^4$  occupies some geometrically suitable positions in protein pockets, forming a large number of weak

contacts.  $L^1$ , on the contrary, has a large electrostatic contribution to the binding energy and forms the highest number of H-bonds. Its stability is provided by a small number of strong bonds. It has a suitable distribution of H-bond donors and acceptors, rather than a suitable size and form of molecular scaffold. For  $L^2$  and  $L^3$  we see that  $SASA_{buried}$  is as small as for  $L^1$ , but the H-bonds number is even smaller than for  $L^4$ . That makes their binding with protein less favorable, and their configuration less stable. A closer analysis of a MD trajectory can reveal details of such binding.

For the  $L^1$  ligand MM/PBSA showed a large negative value of binding free energy with the highest electrostatic term among all the ligands. This can be related to the highest number of protein–ligand hydrogen bonds. The binding was stable during MD simulation, which follows from RMSD time dependence. From the docking structure, it follows that  $L^1$  occupies predominantly the **S1** pocket. In the docking structure, one of the OH groups formed an H-bond with Asn142, but during the simulation it was switching between Met143 and Glu166 residues, while the terminal OH group formed H-bonds with Met143, Ser144, and His163. The OH-containing branch was stretched in **S1** during almost all the simulation time, being strongly fixed by two hydrogen bonds from both edges of the pocket. OH and  $NH_2$  groups on the second branch fixed a catalytic His41 residue, forming H-bonds with it (mostly  $NH_2$ ) and with Glu166, capping thus a catalytic site. The imidazole ring of that branch was located in the docked configuration in the **S2** pocket. During the simulation, however, it went out soon, and was located between **S2** and **S4** pockets most of the time. The last imidazole branch was located in **S4** between Gln189 and Glu166, forming an H-bond with the latter. A weak  $\pi$ -alkyl interaction with Met49 (mentioned in the 2D docking interaction map) was also observed sometimes during the simulation. Per residue MM/PBSA calculation supported the conclusion that these residues stabilized the ligand. A large positive value was also observed for Glu166. This may be related to the disturbing H-bonds between Glu166 and Asn142, covering the **S1** pocket in an unliganded protein.  $L^1$  thus covered a binding groove and was anchored in the **S1** pocket by an acyclic branch. Its aldehyde warhead was located near the target Cys145, forming occasionally the H-bond with Asp142 and Met143. We suggest that  $M^{Pto}$  can be inhibited by  $L^1$ .

For  $L^2$ , MM/PBSA gave the value of the binding energy, which was the smallest among the other ligands, but still highly negative. It formed a comparatively low number of H-bonds with the protein and demonstrated a strong fluctuation in RMSD (between 0.15 nm and 0.20 nm). At least two positions could be selected on the RMSD plot, which allowed us to assume that there are two binding conformations.

In the docked conformation, a 2D interaction map for the OH-containing branch showed hydrogen bonds with Glu166, Ser1B, His163, Ser144, and Leu141. Among them, only H-bonds with Glu166 and Ser144 existed more than 10% of the time during an MD simulation. The branch forming such H-bonds is located in the **S1** pocket, which corresponds to the first type of bound conformations. The branch in the **S1** pocket can either stretch along the pocket or bend inside it, forming an intramolecular H-bond linking OH-groups. The imidazole ring was located in the docked conformation in the **S2** pocket, forming an H-bond with Asp187 and weak  $\pi$ -interactions with Hys41, Cys44, and Met49. During the first third of the simulation, the ring position was preserved, and an H-bond was formed mostly with Asp187.

In the second bound conformation the imidazole ring went out of **S2** and was located between the **S2** and **S4** positions, forming H-bonds with Cys44 via an NH group and with Ser46 via an N atom; a carboxyl group was binding with Thr25. The OH-containing branch at the same time fell out of the **S1** and was located near the catalytic site in the binding groove. It was stabilized by H-bonds with Ser144 and Met143. Note that, in the elongated start configuration, the electrophilic C atom is far from the catalytic Cys145 residue, but during bending and conformation change they became very close in some configurations. Per residue MM/PBSA confirmed these results.

For  $L^3$ , the ligand RMSD was stable until 70 ns, after which it grew sharply and came to a plateau. The existence of two positions can be thus inferred. At the beginning of the

trajectory, the docked configuration was preserved. In this configuration, the CO-containing branch was located in the **S1** pocket, the OH group formed an H-bond with Phe140, and the CO-group was H-bonded to Glu166. During the simulation, CO remained connected with Glu166 almost all the time during which this conformation was preserved, while the -OH was switching its H-bond to His163. In the bent sub-configuration, the -OH could be placed near the catalytic site, forming H-bonds with Gly143 and Cys145. A short OH-containing chain was located near the catalytic site in the docked configuration, forming an H-bond with Asn142; however, during the simulation, the H-bond existed less than 9% of the time. The largest OH-containing branch was directed to the **S4** pocket and fixed by an H-bond with His164; in the simulation it was H-bonded mostly with Gln189 and Glu166. The camphoryl group was placed under the **S2** pocket. Such a position was preserved during the whole simulation. It is clear that the size of the group was too large to be placed into the **S2**. Significant conformation changes occurred after approximately 70 ns of the simulation and were related to the opening of the **S1** pocket by moving Glu166 outwards. After that, the CO-containing chain fell out of it, and took place near **S4**, interacting mostly with solvent. The fact explains the lowest solvent contribution to the binding free energy, even if the interaction of the bound state with the solvent was large enough. Per residue MM/PBSA supported most of these findings. Note, that there were no H-bonds formed by the ligand with Met165, but it made the largest contribution to the interaction energy, which was related apparently to a change in intra-protein interactions.

For **L<sup>4</sup>**, the binding interaction energy was the highest, and an RMSD was stable, suggesting a single binding conformation. In the docked conformation, one of the orthotriazole rings was located in **S1**, and the second one in **S2**. The first one was stabilized by H-bonds with Ser144 and His163, and the second one was stabilized by  $\pi$ -interactions with Met165, Cys44, Met49, and His41. Per residue H-bond statistics over the trajectory showed that the abovementioned H-bonds existed for a large percentage of time. Strong H-bonds were also formed with Glu166 in **S1** and with Asp187 in **S2**. A triazole ring in para-position was placed near the catalytic site between **S1** and **S2**, interacting strongly with Asn142. The 2D interaction map of the docked state revealed also a  $\pi$ -sulfur bond with Cys145. The “inversed” position of a ligand with a cyano-group in **S4** and a para-triazole ring at the catalytic site was thus preferred. Per residue MM/PBSA supported these results and revealed that  $\pi$ -interactions existed not only in the docked configuration, but were also preserved during the whole simulation.

To conclude, the geometry and functional group distribution in **L<sup>1</sup>** is optimal for M<sup>Pro</sup> inhibition. **L<sup>4</sup>** has a proper geometry and can inhibit M<sup>Pro</sup> just by closing the recognition pockets, but it is necessary to change the position of the electrophile in order to form a covalent bond with the catalytic Cys145 residue. **L<sup>2</sup>** is too long, and this does not allow it to fit properly into the pockets. **L<sup>3</sup>** has too many chains and a too large hydrophobic camphor group. Taking this consideration into account, it is possible to design much more efficient M<sup>Pro</sup> inhibitors.

## 5. Conclusions

Computer-aided drug design is a powerful approach that allows one to save research time and resources. In the present work, structure-based drug design, molecular docking, molecular dynamics, and MM/PBSA methods were used in combination, in order to design novel inhibitors against the SARS-CoV-2 main protease and check their efficiency. In summary, four potential inhibitors were constructed on a base of the known M<sup>Pro</sup> 3D structure, recognition, and catalytic mechanism. Docking presented a 3D structure of ligand-protein complexes and proved their potential binding efficiency. Molecular dynamics and MM/PBSA were used to validate the docking results. RMSD, RMSF, Rg, and SASA values showed that all complexes were stable. The analysis of hydrogen bonds revealed the nature of protein-ligand interactions. All the ligands have highly negative binding free energies, lower than  $-60$  kJ/mol.

All the ligands are bound in a recognition groove. Side groups of ligands occupy **S1** and **S2** pockets, interacting in a way similar to natural peptides. The binding energies of the ligands show that the inhibition constant should be high to provide deactivation of  $M^{Pro}$ , thus stopping virus reproduction. In the cases of **L<sup>2</sup>**, **L<sup>3</sup>**, and **L<sup>4</sup>**, further optimization is required; however, in the case of **L<sup>1</sup>** the warhead was located near the catalytic Cys145 residue. Thus favors a covalent bonding of the ligand with the protease. It may be considered as a good candidate for further in vitro and in vivo studies of their activities. We believe that the outcome of this work can be exploited in the development of antiviral candidates against SARS-CoV-2.

Although someone might assume that COVID-19 has been defeated, we need to remain vigilant. New emergencies may appear, and we have to be ready. As WHO says, “. . . we cannot become complacent and fall victim to the cycle of panic and neglect” [5].

**Author Contributions:** Conceptualization, D.S.K. and E.D.K.; methodology, E.D.K.; formal analysis and investigation, E.D.K. and D.S.K.; writing—original draft preparation, D.S.K. and E.D.K.; writing—review and editing, D.S.K.; funding acquisition, S.G.A.; project administration, S.G.A.; supervision, E.D.K. and S.G.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by the Ministry of Science and Higher Education of the Russian Federation within the governmental order for SRF SKIF Boreskov Institute of Catalysis (project FWUR-2024-0040) and the Ministry of Science and Higher Education of the Russian Federation in the implementation of the research program “Use of synchrotron radiation for virological research” within the framework of the Federal Scientific and Technical Program for the Development of Synchrotron and Neutron Research and Research Infrastructure for 2019–2027 (Agreement No. 075-15-2021-1355 (12 October 2021)).

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** Authors thanks SRF “SKIF” and Novosibirsk State University for providing the computational resources and Professor Elena Boldyreva for support.

**Conflicts of Interest:** The authors have no relevant financial or non-financial interests to disclose. The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

Abbreviation	Definition
ACE2	Angiotensin-Converting Enzyme 2
CCDC	Cambridge Crystallographic Data Center
DFT	Density Functional Theory
HIV	Human Immunodeficiency Viruses
IR	Infrared Radiation
MD	Molecular Dynamic
MM/PBSA	Molecular Mechanics/Poisson–Boltzmann Surface Area
NSPs	Non-Structural Proteins
PME	Particle Mesh Ewald
RdRp	RNA-Dependent RNA Polymerase
Rg	Radius of Gyration
RMSD	Root-Mean-Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic Acids
SASA	Solvent-Accessible Surface Area
SBDD	Structure-Based Drug Design
QSAR	Quantitative Structure–Activity Relationship
WHO	World Health Organization

## References

1. Zhou, P.; Yang, X.L.; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; et al. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* **2020**, *579*, 270–273. [CrossRef] [PubMed]
2. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579*, 265–269. [CrossRef] [PubMed]
3. Hui, D.S.; Azhar, E.I.; Madani, T.A.; Ntoumi, F.; Kock, R.; Dar, O.; Ippolito, G.; Mchugh, T.D.; Memish, Z.A.; Drosten, C.; et al. The Continuing 2019-NCoV Epidemic Threat of Novel Coronaviruses to Global Health—The Latest 2019 Novel Coronavirus Outbreak in Wuhan, China. *Int. J. Infect. Dis.* **2020**, *91*, 264–266. [CrossRef] [PubMed]
4. WHO Director-General’s Opening Remarks at the Media Briefing on COVID-19 11 March 2020. Available online: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (accessed on 18 August 2023).
5. World Health Organization. *From Emergency Response to Long-Term COVID-19 Disease Management: Sustaining Gains Made during the COVID-19 Pandemic*; WHO: Geneva, Switzerland, 2023.
6. Ortiz-Prado, E.; Simbaña-Rivera, K.; Gómez-Barreno, L.; Rubio-Neira, M.; Guaman, L.P.; Kyriakidis, N.C.; Muslin, C.; Jaramillo, A.M.G.; Barba-Ostria, C.; Cevallos-Robalino, D.; et al. Clinical, Molecular, and Epidemiological Characterization of the SARS-CoV-2 Virus and the Coronavirus Disease 2019 (COVID-19), a Comprehensive Literature Review. *Diagn. Microbiol. Infect. Dis.* **2020**, *98*, 115094. [CrossRef]
7. Pan, Y.; Wang, L.; Feng, Z.; Xu, H.; Li, F.; Shen, Y.; Zhang, D.; Liu, W.J.; Gao, G.F.; Wang, Q. Characterisation of SARS-CoV-2 Variants in Beijing during 2022: An Epidemiological and Phylogenetic Analysis. *Lancet* **2023**, *401*, 664–672. [CrossRef]
8. Wu, A.; Peng, Y.; Huang, B.; Ding, X.; Wang, X.; Niu, P.; Meng, J.; Zhu, Z.; Zhang, Z.; Wang, J.; et al. Genome Composition and Divergence of the Novel Coronavirus (2019-NCoV) Originating in China. *Cell Host Microbe* **2020**, *27*, 325–328. [CrossRef]
9. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727–733. [CrossRef]
10. Wu, C.; Liu, Y.; Yang, Y.; Zhang, P.; Zhong, W.; Wang, Y.; Wang, Q.; Xu, Y.; Li, M.; Li, X.; et al. Analysis of Therapeutic Targets for SARS-CoV-2 and Discovery of Potential Drugs by Computational Methods. *Acta Pharm. Sin. B* **2020**, *10*, 766–788. [CrossRef]
11. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural Basis for the Recognition of SARS-CoV-2 by Full-Length Human ACE2. *Science* **2020**, *367*, 1444–1448. [CrossRef]
12. Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.Y.; et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181*, 894–904.e9. [CrossRef]
13. Cao, Y.; Li, L.; Feng, Z.; Wan, S.; Huang, P.; Sun, X.; Wen, F.; Huang, X.; Ning, G.; Wang, W. Comparative Genetic Analysis of the Novel Coronavirus (2019-NCoV/SARS-CoV-2) Receptor ACE2 in Different Populations. *Cell Discov.* **2020**, *6*, 11. [CrossRef] [PubMed]
14. Boopathi, S.; Poma, A.B.; Kolandaivel, P. Novel 2019 Coronavirus Structure, Mechanism of Action, Antiviral Drug Promises and Rule out against Its Treatment. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3409–3418. [CrossRef] [PubMed]
15. Marra, M.A.; Jones, S.J.M.; Astell, C.R.; Holt, R.A.; Brooks-Wilson, A.; Butterfield, Y.S.N.; Khattra, J.; Asano, J.K.; Barber, S.A.; Chan, S.Y.; et al. The Genome Sequence of the SARS-Associated Coronavirus. *Science* **2003**, *300*, 1399–1404. [CrossRef]
16. Thiel, V.; Ivanov, K.A.; Putics, Á.; Hertzog, T.; Schelle, B.; Bayer, S.; Weißbrich, B.; Snijder, E.J.; Rabenau, H.; Doerr, H.W.; et al. Mechanisms and Enzymes Involved in SARS Coronavirus Genome Expression. *J. Gen. Virol.* **2003**, *84*, 2305–2315. [CrossRef] [PubMed]
17. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**, *582*, 289–293. [CrossRef]
18. Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K.M.; O’Meara, M.J.; Rezelj, V.V.; Guo, J.Z.; Swaney, D.L.; et al. A SARS-CoV-2 Protein Interaction Map Reveals Targets for Drug Repurposing. *Nature* **2020**, *583*, 459–468. [CrossRef]
19. Babadaei, M.M.N.; Hasan, A.; Vahdani, Y.; Bloukh, S.H.; Sharifi, M.; Kachooei, E.; Haghghat, S.; Falahati, M. Development of Remdesivir Repositioning as a Nucleotide Analog against COVID-19 RNA Dependent RNA Polymerase. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3771–3779. [CrossRef]
20. Basit, A.; Ali, T.; Rehman, S.U. Truncated Human Angiotensin Converting Enzyme 2; a Potential Inhibitor of SARS-CoV-2 Spike Glycoprotein and Potent COVID-19 Therapeutic Agent. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3605–3614. [CrossRef]
21. Elfiky, A.A. SARS-CoV-2 RNA Dependent RNA Polymerase (RdRp) Targeting: An in Silico Perspective. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3204–3212. [CrossRef]
22. Hasan, A.; Paray, B.A.; Hussain, A.; Qadir, F.A.; Attar, F.; Aziz, F.M.; Sharifi, M.; Derakhshankhah, H.; Rasti, B.; Mehrabi, M.; et al. A Review on the Cleavage Priming of the Spike Protein on Coronavirus by Angiotensin-Converting Enzyme-2 and Furin. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3025–3033. [CrossRef]
23. Sinha, S.K.; Shakya, A.; Prasad, S.K.; Singh, S.; Gurav, N.S.; Prasad, R.S.; Gurav, S.S. An In-Silico Evaluation of Different Saikosaponins for Their Potency against SARS-CoV-2 Using NSP15 and Fusion Spike Glycoprotein as Targets. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3244–3255. [CrossRef]
24. Badavath, V.N.; Kumar, A.; Samanta, P.K.; Maji, S.; Das, A.; Blum, G.; Jha, A.; Sen, A. Determination of Potential Inhibitors Based on Isatin Derivatives against SARS-CoV-2 Main Protease (m<sup>PTO</sup>): A Molecular Docking, Molecular Dynamics and Structure-Activity Relationship Studies. *J. Biomol. Struct. Dyn.* **2022**, *40*, 3110–3128. [CrossRef]

25. Fan, H.H.; Wang, L.Q.; Liu, W.L.; An, X.P.; Liu, Z.D.; He, X.Q.; Song, L.H.; Tong, Y.G. Repurposing of Clinically Approved Drugs for Treatment of Coronavirus Disease 2019 in a 2019–Novel Coronavirus-Related Coronavirus Model. *Chin. Med. J.* **2020**, *133*, 1051–1056. [[CrossRef](#)]
26. Liu, X.; Wang, X.J. Potential Inhibitors against 2019-NCoV Coronavirus M Protease from Clinically Approved Medicines. *J. Genet. Genom.* **2020**, *47*, 119–121. [[CrossRef](#)]
27. Ghosh, A.K.; Osswald, H.L.; Prato, G. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *J. Med. Chem.* **2016**, *59*, 5172–5208. [[CrossRef](#)]
28. Yang, S.; Chen, S.J.; Hsu, M.F.; Wu, J.D.; Tseng, C.T.K.; Liu, Y.F.; Chen, H.C.; Kuo, C.W.; Wu, C.S.; Chang, L.W.; et al. Synthesis, Crystal Structure, Structure-Activity Relationships, and Antiviral Activity of a Potent SARS Coronavirus 3CL Protease Inhibitor. *J. Med. Chem.* **2006**, *49*, 4971–4980. [[CrossRef](#)]
29. Ullrich, S.; Nitsche, C. The SARS-CoV-2 Main Protease as Drug Target. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 127377. [[CrossRef](#)]
30. Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved  $\alpha$ -Ketoamide Inhibitors. *Science* **2020**, *368*, 409–412. [[CrossRef](#)]
31. St. John, S.E.; Tomar, S.; Stauffer, S.R.; Mesecar, A.D. Targeting Zoonotic Viruses: Structure-Based Inhibition of the 3C-like Protease from Bat Coronavirus HKU4—The Likely Reservoir Host to the Human Coronavirus That Causes Middle East Respiratory Syndrome (MERS). *Bioorg. Med. Chem.* **2015**, *23*, 6036–6048. [[CrossRef](#)]
32. Mengist, H.M.; Dilnessa, T.; Jin, T. Structural Basis of Potential Inhibitors Targeting SARS-CoV-2 Main Protease. *Front. Chem.* **2021**, *9*, 622898. [[CrossRef](#)] [[PubMed](#)]
33. Dai, W.; Zhang, B.; Jiang, X.M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F.; et al. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* **2020**, *368*, 1331–1335. [[CrossRef](#)]
34. Lee, J.; Kenward, C.; Worrall, L.J.; Vuckovic, M.; Gentile, F.; Ton, A.-T.; Ng, M.; Cherkasov, A.; Strynadka, N.C.J.; Paetzel, M. X-Ray Crystallographic Characterization of the SARS-CoV-2 Main Protease Polypeptide Cleavage Sites Essential for Viral Processing and Maturation. *Nat. Commun.* **2022**, *13*, 5196. [[CrossRef](#)] [[PubMed](#)]
35. Tong, J.B.; Luo, D.; Xu, H.Y.; Bian, S.; Zhang, X.; Xiao, X.C.; Wang, J.; Tong, J.B. A Computational Approach for Designing Novel SARS-CoV-2 Mpro Inhibitors: Combined QSAR, Molecular Docking, and Molecular Dynamics Simulation Techniques. *N. J. Chem.* **2021**, *45*, 11512–11529. [[CrossRef](#)]
36. Kumar, Y.; Singh, H.; Patel, C.N. In Silico Prediction of Potential Inhibitors for the Main Protease of SARS-CoV-2 Using Molecular Docking and Dynamics Simulation Based Drug-Repurposing. *J. Infect. Public Health* **2020**, *13*, 1210–1223. [[CrossRef](#)] [[PubMed](#)]
37. Tejera, E.; Munteanu, C.R.; López-Cortés, A.; Cabrera-Andrade, A.; Pérez-Castillo, Y. Drugs Repurposing Using QSAR, Docking and Molecular Dynamics for Possible Inhibitors of the SARS-CoV-2 M<sup>pro</sup> Protease. *Molecules* **2020**, *25*, 5172. [[CrossRef](#)] [[PubMed](#)]
38. Ma, C.; Sacco, M.D.; Hurst, B.; Townsend, J.A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M.T.; Chen, Y.; et al. Boceprevir, GC-376, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease. *Cell Res.* **2020**, *30*, 678–692. [[CrossRef](#)] [[PubMed](#)]
39. Sacco, M.D.; Ma, C.; Lagarias, P.; Gao, A.; Townsend, J.A.; Meng, X.; Dube, P.; Zhang, X.; Hu, Y.; Kitamura, N.; et al. Structure and Inhibition of the SARS-CoV-2 Main Protease Reveal Strategy for Developing Dual Inhibitors against M<sup>pro</sup> and Cathepsin L. *Sci. Adv.* **2020**, *6*, eabe0751. [[CrossRef](#)] [[PubMed](#)]
40. Tsuji, M.; Tsuji, C.M. Potential Anti-SARS-CoV-2 Drug Candidates Identified through Virtual Screening of the ChEMBL Database for Compounds That Target the Main Coronavirus Protease. *FEBS Open Bio* **2020**, *10*, 995–1004. [[CrossRef](#)]
41. Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F. Network-Based Drug Repurposing for Novel Coronavirus 2019-NCoV/SARS-CoV-2. *Cell Discov.* **2020**, *6*, 14. [[CrossRef](#)]
42. Nukoolkarn, V.; Lee, V.S.; Malaisree, M.; Aruksakulwong, O.; Hannongbua, S. Molecular Dynamic Simulations Analysis of Ritonavir and Lopinavir as SARS-CoV 3CL<sup>pro</sup> Inhibitors. *J. Theor. Biol.* **2008**, *254*, 861–867. [[CrossRef](#)]
43. Wang, Y.; Zhang, D.; Du, G.; Du, R.; Zhao, J.; Jin, Y.; Fu, S.; Gao, L.; Cheng, Z.; Lu, Q.; et al. Remdesivir in Adults with Severe COVID-19: A Randomised, Double-Blind, Placebo-Controlled, Multicentre Trial. *Lancet* **2020**, *395*, 1569–1578. [[CrossRef](#)] [[PubMed](#)]
44. Eastman, R.T.; Roth, J.S.; Brimacombe, K.R.; Simeonov, A.; Shen, M.; Patnaik, S.; Hall, M.D. Remdesivir: A Review of Its Discovery and Development Leading to Emergency Use Authorization for Treatment of COVID-19. *ACS Cent. Sci.* **2020**, *6*, 672–683. [[CrossRef](#)] [[PubMed](#)]
45. Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G. Remdesivir and Chloroquine Effectively Inhibit the Recently Emerged Novel Coronavirus (2019-NCoV) In Vitro. *Cell Res.* **2020**, *30*, 269–271. [[CrossRef](#)] [[PubMed](#)]
46. Sharma, P.; Joshi, T.; Mathpal, S.; Joshi, T.; Pundir, H.; Chandra, S.; Tamta, S. Identification of Natural Inhibitors against Mpro of SARS-CoV-2 by Molecular Docking, Molecular Dynamics Simulation, and MM/PBSA Methods. *J. Biomol. Struct. Dyn.* **2022**, *40*, 2757–2768. [[CrossRef](#)]
47. Aanouz, I.; Belhassan, A.; El-Khatibi, K.; Lakhlifi, T.; El-Idrissi, M.; Bouachrine, M. Moroccan Medicinal Plants as Inhibitors against SARS-CoV-2 Main Protease: Computational Investigations. *J. Biomol. Struct. Dyn.* **2021**, *39*, 2971–2979. [[CrossRef](#)] [[PubMed](#)]
48. Krishna Cheerala, V.S.; Ghanta, P.; Neelakantan, S.C. Design, Synthesis and in Silico Screening of Benzoxazole–Thiazolidinone Hybrids as Potential Inhibitors of SARS-CoV-2 Proteases. *RSC Adv.* **2021**, *11*, 39328–39342. [[CrossRef](#)]

49. Patel, U.; Desai, K.; Dabhi, R.C.; Maru, J.J.; Shrivastav, P.S. Bioprospecting Phytochemicals of *Rosmarinus officinalis* L. for Targeting SARS-CoV-2 Main Protease (M<sup>Pro</sup>): A Computational Study. *J. Mol. Model.* **2023**, *29*, 161. [CrossRef]
50. Ghosh, R.; Chakraborty, A.; Biswas, A.; Chowdhuri, S. Evaluation of Green Tea Polyphenols as Novel Corona Virus (SARS CoV-2) Main Protease (M<sup>Pro</sup>) Inhibitors—An in Silico Docking and Molecular Dynamics Simulation Study. *J. Biomol. Struct. Dyn.* **2021**, *39*, 4362–4374. [CrossRef]
51. Aggarwal, G.; Henry, B.M.; Aggarwal, S.; Bangalore, S. Cardiovascular Safety of Potential Drugs for the Treatment of Coronavirus Disease 2019. *Am. J. Cardiol.* **2020**, *128*, 147–150. [CrossRef]
52. Khan, Z.; Karataş, Y.; Rahman, H. Anti COVID-19 Drugs: Need for More Clinical Evidence and Global Action. *Adv. Ther.* **2020**, *37*, 2575–2579. [CrossRef]
53. Javorac, D.; Grahovac, L.; Manić, L.; Stojilković, N.; Anđelković, M.; Bulat, Z.; Đukić-Čosić, D.; Curcic, M.; Djordjevic, A.B. An Overview of the Safety Assessment of Medicines Currently Used in the COVID-19 Disease Treatment. *Food Chem. Toxicol.* **2020**, *144*, 111639. [CrossRef] [PubMed]
54. Muralidharan, N.; Sakthivel, R.; Velmurugan, D.; Gromiha, M.M. Computational Studies of Drug Repurposing and Synergism of Lopinavir, Oseltamivir and Ritonavir Binding with SARS-CoV-2 Protease against COVID-19. *J. Biomol. Struct. Dyn.* **2021**, *39*, 2673–2678. [CrossRef] [PubMed]
55. Avti, P.; Chauhan, A.; Shekhar, N.; Prajapat, M.; Sarma, P.; Kaur, H.; Bhattacharyya, A.; Kumar, S.; Prakash, A.; Sharma, S.; et al. Computational Basis of SARS-CoV 2 Main Protease Inhibition: An Insight from Molecular Dynamics Simulation Based Findings. *J. Biomol. Struct. Dyn.* **2022**, *40*, 8894–8904. [CrossRef] [PubMed]
56. Khan, S.A.; Zia, K.; Ashraf, S.; Uddin, R.; Ul-Haq, Z. Identification of Chymotrypsin-like Protease Inhibitors of SARS-CoV-2 via Integrated Computational Approach. *J. Biomol. Struct. Dyn.* **2021**, *39*, 2607–2616. [CrossRef] [PubMed]
57. Kumar, R.; Serrette, J.M.; Thompson, E.B. Osmolyte-Induced Folding Enhances Tryptic Enzyme Activity. *Arch. Biochem. Biophys.* **2005**, *436*, 78–82. [CrossRef] [PubMed]
58. Kumari, R.; Kumar, V.; Dhankhar, P.; Dalal, V. Promising Antivirals for PL<sup>Pro</sup> of SARS-CoV-2 Using Virtual Screening, Molecular Docking, Dynamics, and MMPBSA. *J. Biomol. Struct. Dyn.* **2023**, *41*, 4650–4666. [CrossRef] [PubMed]
59. Kumar, K.A.; Sharma, M.; Dalal, V.; Singh, V.; Tomar, S.; Kumar, P. Multifunctional Inhibitors of SARS-CoV-2 by MM/PBSA, Essential Dynamics, and Molecular Dynamic Investigations. *J. Mol. Graph. Model.* **2021**, *107*, 107969. [CrossRef] [PubMed]
60. Costanzi, E.; Kuzikov, M.; Esposito, F.; Albani, S.; Demitri, N.; Giabbai, B.; Camasta, M.; Tramontano, E.; Rossetti, G.; Zaliani, A.; et al. Structural and Biochemical Analysis of the Dual Inhibition of MG-132 against SARS-CoV-2 Main Protease (M<sup>Pro</sup>/3CL<sup>Pro</sup>) and Human Cathepsin-L. *Int. J. Mol. Sci.* **2021**, *22*, 11779. [CrossRef]
61. Batool, M.; Ahmad, B.; Choi, S. A Structure-Based Drug Discovery Paradigm. *Int. J. Mol. Sci.* **2019**, *20*, 2783. [CrossRef]
62. Banerjee, R.; Perera, L.; Tillekeratne, L.M.V. Potential SARS-CoV-2 Main Protease Inhibitors. *Drug Discov. Today* **2021**, *26*, 804–816. [CrossRef]
63. Fischer, C.; Feys, J.R. SARS-CoV-2 M<sup>Pro</sup> Inhibitors: Achieved Diversity, Developing Resistance and Future Strategies. *Fut. Pharmacol.* **2023**, *3*, 80–107. [CrossRef]
64. Citarella, A.; Scala, A.; Piperno, A.; Micale, N. SARS-CoV-2 M<sup>Pro</sup>: A Potential Target for Peptidomimetics and Small-Molecule Inhibitors. *Biomolecules* **2021**, *11*, 607. [CrossRef] [PubMed]
65. Zabo, S.; Lobb, K.A. In Silico Substrate-Binding Profiling for SARS-CoV-2 Main Protease (M<sup>Pro</sup>) Using Hexapeptide Substrates. *Viruses* **2023**, *15*, 1480. [CrossRef] [PubMed]
66. Stoddard, S.V.; Stoddard, S.D.; Oelkers, B.K.; Fitts, K.; Whalum, K.; Whalum, K.; Hemphill, A.D.; Manikonda, J.; Martinez, L.M.; Riley, E.G.; et al. Optimization Rules for SARS-CoV-2 M<sup>Pro</sup> Antivirals: Ensemble Docking and Exploration of the Coronavirus Protease Active Site. *Viruses* **2020**, *12*, 942. [CrossRef] [PubMed]
67. Lipinski, C.A. Lead- and Drug-like Compounds: The Rule-of-Five Revolution. *Drug Discov. Today Technol.* **2004**, *1*, 337–341. [CrossRef] [PubMed]
68. Kenny, P.W. Hydrogen-Bond Donors in Drug Design. *J. Med. Chem.* **2022**, *65*, 14261–14275. [CrossRef]
69. Yarovaya, O.I.; Shcherbakov, D.N.; Borisevich, S.S.; Sokolova, A.S.; Gureev, M.A.; Khamitov, E.M.; Rudometova, N.B.; Zybkina, A.V.; Mordvinova, E.D.; Zaykovskaya, A.V.; et al. Borneol Ester Derivatives as Entry Inhibitors of a Wide Spectrum of SARS-CoV-2 Viruses. *Viruses* **2022**, *14*, 1295. [CrossRef]
70. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D. Improved Protein-Ligand Docking Using GOLD. *Proteins Struct. Funct. Bioinform.* **2003**, *52*, 609–623. [CrossRef]
71. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. Gaussian 16 2016. Available online: <https://gaussian.com/> (accessed on 8 January 2024).
72. Andrienko, G.A. Chemcraft—Graphical Software for Visualization of Quantum Chemistry Computations. Available online: <https://www.chemcraftprog.com/> (accessed on 1 July 2023).
73. BIOVIA Discovery Studio 2021 Visualizer. Available online: <https://www.3ds.com/products/biovia> (accessed on 8 January 2024).
74. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [CrossRef]
75. Dodda, L.S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W.L. LigParGen Web Server: An Automatic OPLS-AA Parameter Generator for Organic Ligands. *Nucleic Acids Res.* **2017**, *45*, W331–W336. [CrossRef]

76. Berendsen, H.J.C.; Grigera, J.R.; Straatsma, T.P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271. [[CrossRef](#)]
77. Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101. [[CrossRef](#)] [[PubMed](#)]
78. Bernetti, M.; Bussi, G. Pressure Control Using Stochastic Cell Rescaling. *J. Chem. Phys.* **2020**, *153*, 114107. [[CrossRef](#)] [[PubMed](#)]
79. Hess, B.; Bekker, H.; Berendsen, H.J.C.; Fraaije, J.G.E.M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472. [[CrossRef](#)]
80. Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190. [[CrossRef](#)]
81. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593. [[CrossRef](#)]
82. Kumari, R.; Kumar, R.; Lynn, A. G\_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.