Ensemble Docking Coupled to Linear Interaction Energy Calculations for Identification of Coronavirus Main Protease (3CLpro) Non-Covalent Small-Molecule Inhibitors

Marko Jukič [1,2], **Dušanka Janežič [2,*]** and **Urban Bren [1,2,*]**

[1] Laboratory of Physical Chemistry and Chemical Thermodynamics, Faculty of Chemistry and Chemical Engineering, University of Maribor, Smetanova ulica 17, SI-2000 Maribor, Slovenia; marko.jukic@um.si

[2] Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Glagoljaška 8, SI-6000 Koper, Slovenia

* Correspondence: dusanka.janezic@upr.si (D.J.); urban.bren@um.si (U.B.);
  Tel.: +386-5-611-76-59 (D.J.); +386-2-22-94-421 (U.B.)

check for updates

Abstract: SARS-CoV-2, or severe acute respiratory syndrome coronavirus 2, represents a new strain of *Coronaviridae*. In the closing 2019 to early 2020 months, the virus caused a global pandemic of COVID-19 disease. We performed a virtual screening study in order to identify potential inhibitors of the SARS-CoV-2 main viral protease (3CLpro or Mpro). For this purpose, we developed a novel approach using ensemble docking high-throughput virtual screening directly coupled with subsequent Linear Interaction Energy (LIE) calculations to maximize the conformational space sampling and to assess the binding affinity of identified inhibitors. A large database of small commercial compounds was prepared, and top-scoring hits were identified with two compounds singled out, namely 1-[(R)-2-(1,3-benzimidazol-2-yl)-1-pyrrolidinyl]-2-(4-methyl-1,4-diazepan-1-yl)-1-ethanone and [({(S)-1-[(1H-indol-2-yl)methyl]-3-pyrrolidinyl}methyl)amino](5-methyl-2H-pyrazol-3-yl)formaldehyde. Moreover, we obtained a favorable binding free energy of the identified compounds, and using contact analysis we confirmed their stable binding modes in the 3CLpro active site. These compounds will facilitate further 3CLpro inhibitor design.

Keywords: COVID-19; SARS-CoV-2; Mpro; 3CLpro; 3C-like protease; virtual screening; inhibitors; in silico drug design; free-energy calculations

1. Introduction

Severe acute respiratory syndrome coronavirus 2 or SARS-CoV-2, the pathogen behind the 2019–2020 coronavirus pandemic (COVID-19), is a member of the *Coronaviridae* family, a positive-sense single-stranded (+ssRNA) RNA virus [1,2]. As there are only a handful of therapeutic options for this global threat, novel drug design is critical; thus, we performed a structure-based virtual screening in order to identify potential inhibitors of the main viral protease (3CLpro or Mpro) of SARS-CoV-2 [3].

In the 3CLpro homodimer, the P1 pocket of the substrate-binding site is formed and seems essential for the catalytically active form [4–6]. The enzyme is vital for the processing of coronavirus polyproteins (pp1a, ppa1ab) that are then cleaved by 3C-like and papain-like proteases to form mature non-structural proteins (NSPs), which are themselves involved in subsequent viral replication mechanisms [7]. 3CLpro represents a cysteine protease (EC 3.4.22.69, "3C" refers to the *Picornaviridae* Enterovirus protease 3C) [8] and shares 96% sequence identity with the SARS-CoV main protease (Supporting Information; aligned PDB ID: 6LU7 and 2QIQ with 288 identical residues out of 301) [3,4,8,9]. The substrate recognition

pockets in 3CL[pro] are named as P1–4, and the enzyme is currently the most studied representative in the context of drug design, mainly due to the availability of structural data. X-ray crystal structure of the 3CL[pro] in complex with the inhibitor N3 has been recently released with PDB IDs 6LU7 and 7BQY at 2.16 and 1.7 Å resolutions, respectively [3]. N3 is a covalent inhibitor of 3CL[pro], featuring a vinyl carboxyl ester as an electrophilic warhead that acts as a Michael-acceptor, reacting with the catalytic Cys145 nucleophile [3]. Substrate specificity is described as P1-Gln, P2-Leu (hydrophobic), P3-Val (or positively charged residues) or P4-Ala (small hydrophobic), but scientific literature also describes preference for His at the P1 binding pocket of the protease active site [9–13]. Proteolysis itself occurs via a catalytic dyad defined by Cys145 and His41 [14].

Considering the currently available structural data, standard in silico docking efforts towards novel potential inhibitors of SARS-CoV-2 main protease are underway [15]. However, only two peptide-like covalent inhibitors have been reported in scientific literature [3]. Due to drawbacks associated with covalent inhibitors, we opted for the identification of novel non-covalent protease inhibitors in a robust screening experiment [16]. We believe the non-covalent inhibitors offer synthetic availability, the flexibility of optimization and can also be used for the future design of covalent inhibitors, if necessary [17]. To this end, we developed a novel methodology directly coupling ensemble docking high-throughput virtual screening (HTVS) with subsequent Linear Interaction Energy (LIE) calculations. Ensemble docking affords viable starting ligand poses and ensemble protein conformations, thereby maximizing the conformational space sampling and yielding reliable ligand binding affinities in the following LIE step. To the best of our knowledge, only SARS-CoV 3CL[pro] small-molecule inhibitors are reported in the scientific literature and can be used as starting points, but no SARS-CoV-2 3CL[pro] small-molecule non-covalent inhibitors are available as of yet (Figure 1) [18].
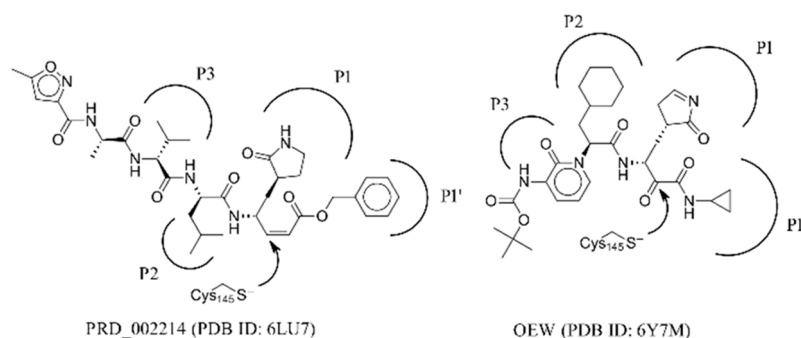

Figure 1. Existing inhibitors of SARS-CoV-2 supported by structural data. Depicted are binding pockets (Px) and the site of covalent reaction.

2. Results and Discussion

2.1. Database Preparation

In a contemporary VS (virtual screening) or HTVS (high-throughput virtual screening) scenario, database design is essential for efficient CPU-time usage in downstream calculations. In order to commence a robust HTVS scenario, we gathered commercially available databases (e.g., ENAMINE, Vitas-M, Chembridge, Maybridge, Ambinter, Otava, PrincetonBIO, Key-Organics, Life Chemicals, Uorsy, Specs) and pre-filtered all compounds in order to exclude small fragments or extra-large molecules, aggregators, and compounds with poor physico-chemical properties. This step was performed using OpenEye FILTER software (OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com). The following parameters were used: min_molwt 250, max_molwt 800, min_solubility moderately, eliminate known and predicted aggregators and allowed elements H, C, N, O, F, S, Cl, Br, I and P. This database was subsequently filtered for PAINS [19–21] and REOS structures in order to eliminate reactive and labile functional groups [22,23]. For this step we used KNIME software with RDKit software nodes to compare all structures in the library to the selection

of SMARTS-formatted flags and to remove hits from the database. We ended up with a collection of approximately 4 million compounds that was expanded in the subsequent step where final enumeration of undefined chiral centers, tautomeric structures, removal of structural faults, ionization at the pH of 7.4 and minimization (using OPLS 3 force-field) towards the final 3D conformation was performed. For this work, Ligprep tool by Schrödinger (Release 2018–3, Schrödinger, LLC, New York, NY, USA 2020) was employed [24,25]. The final database thus consisted of 8,190,951 molecules and was ultimately used for conformer 3D-database preparation using OpenEye OMEGA2 tool (OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com). A maximum number of conformations was set at 25, and rms threshold of 0.8 nm afforded approximately 205 million compound conformations ready for VS (Figure 2).
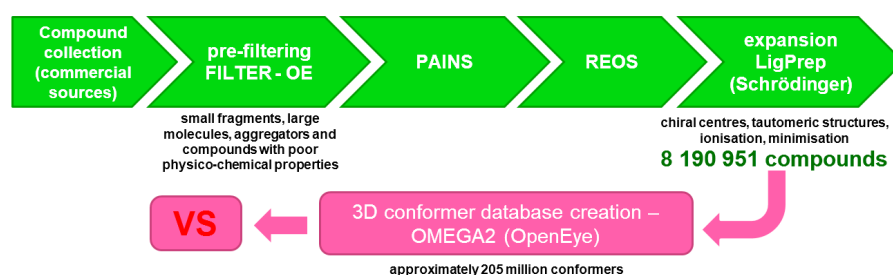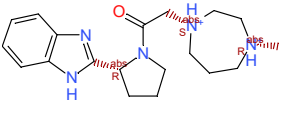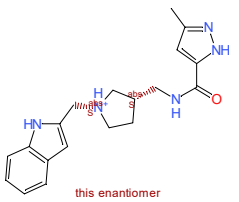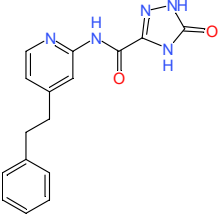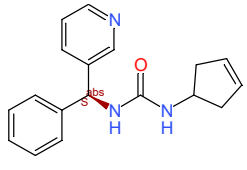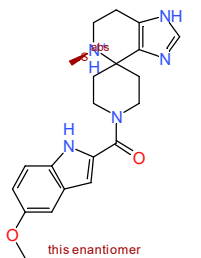

Figure 2. Database preparation for subsequent virtual screening (VS) on the SARS-CoV-2 main protease 3CL^pro or M^pro. The final database contained 8,190,951 molecules before conformer generation.

2.2. Target Preparation

Next, we examined the available experimental SARS-CoV-2 3CL^pro crystal structures and identified the main protease in complex with N3 peptide-like covalent inhibitor published by Yang, H. et al. (PDB ID: 6LU7) [3]. The ligand possessed a Michael-acceptor as an electrophilic warhead to react with the active site cysteine. The covalent bond was cleaved, the N3 residue removed, the cysteine amino-acid residue regenerated (Open Source PyMOL, release 2.1) and the target prepared using Schrödinger Small-Molecule Discovery Suite (Schrödinger LLC, New York). Residue conformation at the active site was checked by superposition with PDB ID: 6M2N with an all atom RMSD of 0.538 Å [26]. As our main goal represented the identification of potential 3CL^pro non-covalent inhibitors, we did not want to limit the VS search to a binding pocket occupied by the only reported N3 inhibitor. Thus, we examined similar protein complexes published in the PDB database using ProBiS server (https://probis.nih.gov/) to identify all possible small-molecule binding modes in the vicinity of the catalytic Cys145 (Figure 3). Therefore, a PDB ID: 6LU7 3CL^pro was used as an input for ProBiS calculation and one binding site identified (binding site 1 in ProBiS; proximity of Cys145). ProBiS server thus produced local superimposition on the defined binding site, and the ligands from the locally superimposed proteins with structural data (PDB IDs: 2op9, 2gz8, 4mds, 3v3m, 4twy, 3vb4, 2hob, 3tnt, 3vb6, 2gx4 and 2gtb) were used for extended receptor space definition. [27,28].

The postulated binding site, located in the proximity of Cys145, was furnished with several superimposed ligands from locally aligned similar protein structures (PDB IDs: 2op9, 2gz8, 4mds, 3v3m, 4twy, 3vb4, 2hob, 3tnt, 3vb6, 2gx4 and 2gtb), which along with the N3-3CL^pro crystal complex (PDB entry: 6LU7) served the subsequent docking definition. With the binding site defined, the receptor structure was generated using OEDocking 3.2.0.2 software package (OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com; details are in Supporting Information).

<table>
<tr><td>(**a**)</td><td>(**b**)</td></tr>
</table>

Figure 3. (**a**) Identification of the active site in the vicinity of the Cys145 using the ProBiS server and superposition of the small-molecule ligands from the locally aligned complexes at the 3CL^{pro} active site; (**b**) the resulting extended docking grid calculated by Make Receptor of the OpenEye software.

2.3. Ensemble Docking

For the HTVS step, we performed a robust ensemble docking experiment coupled to a subsequent Linear Interaction Energy (LIE) calculation workflow into the prepared receptor binding site to afford 1‰ of top-scoring compounds, as depicted in Figure 4 (Fred; OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com).

Figure 4. HTVS workflow incorporating ensemble docking coupled to LIE calculations initiated from multiple MD ensemble complexes for the two top-scoring compounds.

In the next step, 3CL^{pro} ensemble was prepared via a 100 ns molecular dynamics (MD) simulation (PDB ID: 6LU7) followed by a pairwise average-linkage Hierarchical clustering step (ClusCO software, clustering considered the backbone atoms of the whole target protein) to afford 7 representative protein conformations [29,30]. Using OEDocking software (3.2.0.3, OpenEye Scientific Software, Inc., Santa Fe, NM, USA; www.eyesopen.com), mean FRED scores from the ensemble docking experiment highlighted top-scoring compounds with the top five compounds **1** to **5** collected in Tables 1 and S5 [31–34]. FRED scores for the top 100 scoring hits resided below −12.1 with the top-scorer **1** at −14.5, compound **2** at −13.5 and all other hits above −13.0. No molecular weight or cLogP FRED score bias could be observed for the top-scoring compounds, so no additional docking score normalization step had to be performed (Supporting Information, Figures S19–S23). Ensemble docking was somewhat analogous to a classical docking experiment with all residues in the active site kept flexible during the docking runs. Our objective was to verify the predicted binding poses and to postulate that the compounds can adopt a favorable binding mode even with protein flexibility considered as opposed to performing a classical docking experiment with a single rigid protein conformation (Supporting Information, Figure S18).

Table 1. Identified top-scoring compounds in the ensemble docking VS on the SARS-CoV-2 main protease (for the complete list see Table S5 in the Supporting Information).

No.	Structure	Mr (g/mol)	Binding Mode	Fred Docking Score [1]
1	this enantiomer	343.5	P2–P3	−14.5
2	this enantiomer	338.4	P2–P3	−13.5
3		309.3	P2–P3	−13.0
4	this enantiomer	293.4	P1′–P2	−13.0
5	this enantiomer	380.5	P1′–P2	−13.0

[1] Cleaved native ligand N3 from PDB ID: 6LU7 served as a re-docking validation.

2.4. Free-Energy Calculations and Contact Analysis

The Linear Interaction Energy (LIE) method was applied for calculating the binding free energy of the identified two top-scoring compounds [35,36]. Compounds **1** and **2** with the FRED scores of −14.5 and −13.5, respectively, displayed significantly lower values (below −13.5) and were followed by 7 compounds with scores in the narrow range of −13.0 to −12.8. Compounds **1** and **2** were, therefore, selected for subsequent LIE calculation based on the predicted docking score and a favorable binding pose generation. The docked conformations of the two top-scoring compounds exhibited an analogous positioning in the S2–S3 pocket of the 3CL^{pro} active site interacting mainly with His41, Met49, His164, Glu166, Asp187 and Gln189 (Figure 5) [37]. The predicted binding mode is in accordance with the crystal ligand N3 (PDB ID: 6LU7) and analogous to the reported binding mode of a potential SARS-CoV-2 3CL^{pro} inhibitor dipyridamole reported by Li et al. [38]. Othar et al. identified similar interaction profiles upon MD experiments using FDA-approved compounds [39].
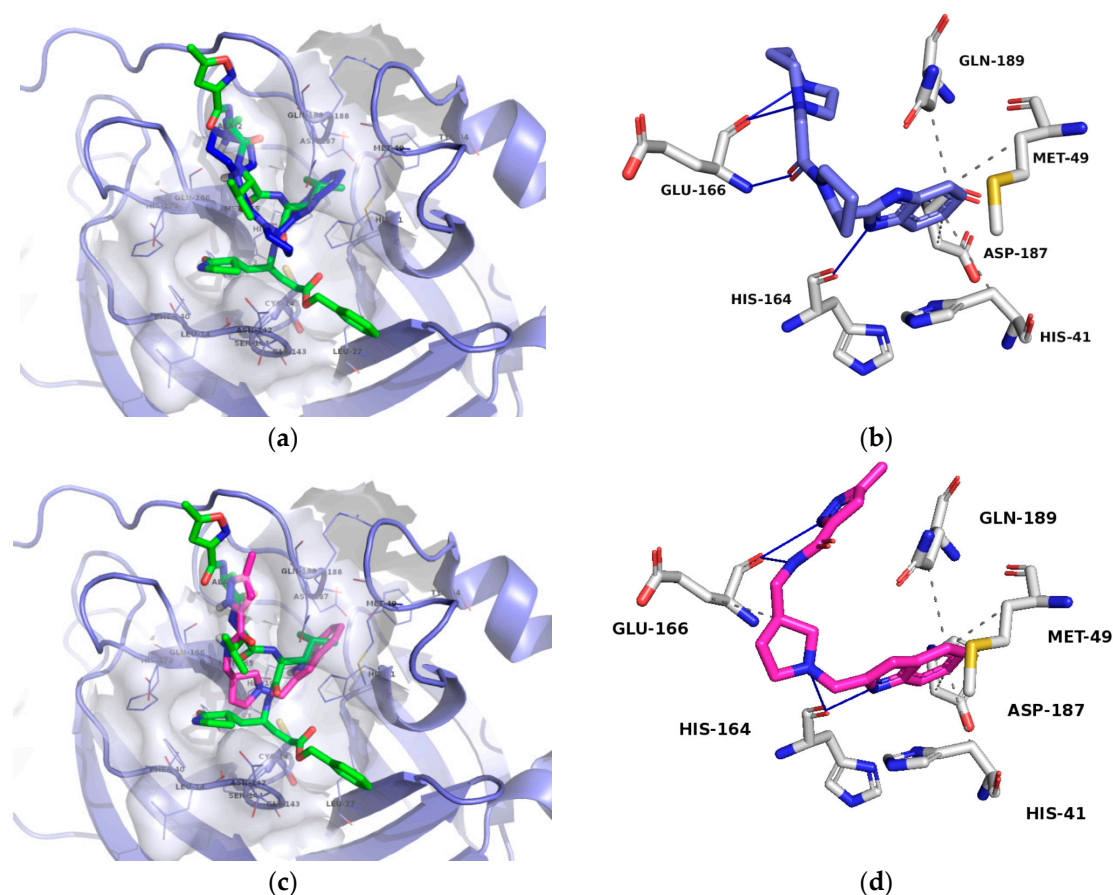
(a)

(b)

(c)

(d)

Figure 5. (**a**): Compound **1** binding mode presented in blue colored stick model; (**b**): Compound **1** key interactions; (**c**): Compound **2** binding mode presented in cyan colored stick model. Reference N3 ligand from the PDB ID: 6LU7 is depicted in green-colored stick model and the protein surface around the ligand calculated. (**d**): Compound **2** key interactions. Hydrogen bonds are denoted as blue lines and hydrophobic interactions as grey lines.

Clustered complexes from the ensemble docking were used as starting ligand-receptor bound complexes, i.e., 7 starting complexes for each compound, subjected to the 100 ns MD production run along with corresponding 100 ns MD simulations in ligand-free state for maximal conformational space coverage. Weighted LIE calculation demonstrated the affinity towards 3CL^{pro} for compound **1** with $_\Delta G_{LIE-BIND}$ value of -8.2 ± 1.9 and for compound **2** with $_\Delta G_{LIE-BIND}$ value of -3.5 ± 1.7 kcal/mol using pre-optimized α and β parameters (Table 2). Moreover, ligand-protein contact analyses along all 7 production runs of compound **1** confirmed the hydrogen bond formation with Glu166, Asp187, Gln189, His41 (for more than 90% of simulation time) and His164 (for more than 50% of simulation time) along with an average of 7 hydrophobic contacts along all 100 ns production runs. Similar observations in all MD simulation runs were made for compound **2** as well. Hydrogen bonds with Gln192, Glu166, Gln189, His164 (for more than 90% of simulation time), Glu166, Val186, Arg188 and Thr190 (for more than 50% of simulation time) along with an average of 9 hydrophobic contacts were formed (details on individual MD replicas can be found in Supporting Information, Figures S2–S15). MD simulations thus place compounds **1** and **2** near P1, tightly into the P2–P3 pockets of the 3CL^{pro} active site in a close proximity (under 4 Å) to the catalytic Cys145 and His41 residues.

Table 2. Calculation of binding free energies for compounds **1** and **2** using LIE methodology.

Compound	Free VdW (kcal/mol)	Free Coulomb (kcal/mol)	Complex VdW Weighted Sum (kcal/mol)	Complex Coulomb Weighted Sum (kcal/mol)	ΔG_{LIE}^{BIND} (kcal/mol)
1	−16.2 ± 0.2	−32.3 ± 0.1	−22.0 ± 1.4	−37.3 ± 2.4	−8.2 ± 1.9
2	−14.7 ± 0.2	−19.0 ± 0.1	−22.5 ± 2.4	−18.7 ± 2.6	−3.5 ± 1.7

3. Materials and Methods

3.1. MD and Ensemble Docking

Crystal complex (PDB ID: 6LU7) was prepared with Yasara software [29]. Missing hydrogens were added, overlapping atoms adjusted, hydrogen bonds optimized and residue ionization assigned at pH = 7.4 [40,41]. The system was solvated using TIP3P (cubic; 10 Å padding, periodic boundary conditions applied) water model, and a physiological concentration of 0.9% of NaCl ions was added with an appropriate excess of either Na$^+$ or Cl$^-$ to neutralize the cell (long-range Coulomb forces calculated using particle-mesh Ewald algorithm). After steepest descent and simulated annealing minimizations to remove sterical clashes, the MD simulation was run for 100 ns using the AMBER14 force field for the protein, GAFF for the ligands as well as TIP3P for water [42,43]. Ligand charges were assigned using AM1-BCC [44]. The equations of motion were integrated with a multiple timestep of 1.25 fs for bonded interactions and 2.5 fs for non-bonded interactions at a temperature of 298K and a pressure of 1 atm (NPT ensemble) using algorithms described in detail previously with snapshots saved every 100 ps [45]. Hydrogen atom bonds were not constrained during the simulation. Energy parameters of the system were stable throughout the production run as were root-mean-square deviation (RMSD) values of protein backbone. Protein conformation models (MD snapshots, 100 structures) were aligned using Theseus and iteratively clustered using ClusCo software (hierarchical clustering, pairwise average-linkage manner, rmsd score) in order to obtain clusters covering a representative portion of the trajectory (with all cluster occupancies above 1% and no single-structure clusters formed) [30,46]. We identified 7 clusters and centroid structures were selected for target preparation as described in the previous chapter.

Ensemble docking was performed using FRED software and the final score calculated as the median FRED score (7 docking experiments on 7 identified protein conformations). Docked complexes of the top-scoring compounds **1** and **2** (Table S5) in all 7 target conformations from ensemble docking were used in subsequent LIE calculations. We generated 14 independent 100 ns MD trajectories totaling to 1.4 µs of simulation time for each ligand in protein and water environments in order to thoroughly explore the conformational space of the system. Moreover, the starting coordinates of ligands and ligand–protein complexes were as varied as possible due to preceding ensemble docking step, thereby effectively enhancing the conformation space sampling.

Simulations of free ligands in water were performed as well to estimate potential energies for calculating binding affinities using LIE methodology. Standard MD simulations were run on a ligand in TIP3P water. MD simulations of 100 ns were carried out analogously as described above for receptor–ligand complexes (no restraints were applied). Based on generated snapshots (1000 snapshots per run), we ran MD energetics analyses to get the VDW and electrostatic interactions of the ligands, and their average values were taken for the subsequent LIE calculation.

3.2. Free Energy Calculation Using LIE Methodology

A number of different computational approaches are available for predicting or estimating binding free energies. Linear Interaction Energy (LIE) methodology was proposed by Aqvist et al. [47]. This is a semiempirical method based on the linear response theory. This methodology focuses on the starting and end states of the binding process, which are the free and the bound state of the ligand. LIE is usually

less computationally intensive than FEP but on the contrary with the popular MM-PB(GB)SA, uses an explicit solvent model; therefore, de-solvation can be handled in an explicit manner. The concept of the LIE approach is to separately calculate the VdW and electrostatic interaction energies of the ligand in water and of the ligand in complex with solvated protein. Then, average interaction energies between the ligand and its surroundings are analyzed using Equation (1).

$$\Delta G_{bind} = \alpha \, \Sigma_i^N \, \Delta V_i^{VdW} + \beta \, \Sigma_i^N \, \Delta V_i^{coulomb} \tag{1}$$

In Equation (1) the Δ term indicates the change in potential energy between the ligand bound and ligand free (in water) states. The α and β represent LIE empirical parameters, determined by comparing calculated and experimentally measured binding affinities. Their values are optimized by Aqvist et al. [47].

For the LIE approach, the MD simulations of 100 ns of receptor-bound ligands (7 complexes for 2 ligands, 14 in total) as well as of free ligands in water were carried out to obtain the VdW and electrostatic interaction energies between the ligand and its surroundings. Each MD simulation afforded 1000 snapshots (using 100 ps intervals) of each compound in different environments which were used for calculating average electrostatic and VdW energies. Energies were simply averaged for the ligand-free simulation (2 sets separately for compounds **1** and **2**), while ligand bound potential energies were initially weighted according to Equation (2) and subsequently used in the general LIE Equation (1) (Table 2) [35,36,48]. The employed method was reported as suitable for calculation of ligand binding free energy, compared to MM/PBSA and reported to produce relevant results even with shorter simulation times when compared to alternative methodologies [49].

$$W_i = \frac{e^{-\Delta G_{calc,i} \, /k_b T}}{\Sigma_i e^{-\Delta G_{calc,i} \, /k_b T}} \tag{2}$$

4. Conclusions

This work presents an extensive and robust virtual screening scenario on the SARS-CoV-2 main protease 3CL^{pro} or M^{pro}, a potential therapeutic target for COVID-19. We report two top-scoring compounds, 1-[(R)-2-(1,3-benzimidazol-2-yl)-1-pyrrolidinyl]-2-(4-methyl-1,4-diazepan-1-yl)-1-ethanone and [({(S)-1-[(1H-indol-2-yl)methyl]-3-pyrrolidinyl}methyl)amino](5-methyl-2H-pyrazol-3-yl)formaldehyde, as viable binders supported by LIE calculations starting from multiple ensemble 3CL^{pro} conformations. This is an in silico work that warrants experimental support in future research and, if successful, the presented set of compounds could be validated for further development of novel non-covalent inhibitors of SARS-CoV-2 main protease (or be used as probes or experimental decoys). Last but not least, the reported hits possess a relatively low molecular weight, their scaffolds are suitable for synthetic optimization and should be synthetically accessible and even commercially available.

Moreover, the presented methodology where ensemble docking is used to identify viable starting ligand poses and ensemble protein conformations directly coupled to subsequent Linear Interaction Energy (LIE) calculations represents a novel advantageous approach to maximize the conformational space sampling thereby facilitating reliable ligand binding free energies.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

HTVS	High-Throughput Virtual Screening
LIE	Linear Interaction Energy
MD	Molecular Dynamics
M^{pro}	Main Protease
PAINS	Pan-assay interference compounds
REOS	Rapid elimination of swill filter
VS	Virtual Screening

References

1. Gorbalenya, A.E.; Baker, S.C.; Baric, R.S.; de Groot, R.J.; Drosten, C.; Gulyaeva, A.A. The species severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536–544.

2. Wang, C.; Horby, P.W.; Hayden, F.G.; Gao, G.F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470–473. [CrossRef]

3. 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 COVID-19 virus and discovery of its inhibitors. *Nature* **2020**, *582*, 289–293. [CrossRef] [PubMed]

4. 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 α-ketoamide inhibitors. *Science* **2020**, *368*, 409–412. [CrossRef]

5. Grottesi, A.; Bešker, N.; Emerson, A.; Manelfi, C.; Beccari, A.R.; Frigerio, F.; Lindahl, E.; Cerchia, C.; Talarico, C. Computational Studies of SARS-CoV-2 3CLpro: Insights from MD Simulations. *Int. J. Mol. Sci.* **2020**, *21*, 5346. [CrossRef]

6. Shi, J.; Sivaraman, J.; Song, J. Mechanism for Controlling the Dimer-Monomer Switch and Coupling Dimerization to Catalysis of the Severe Acute Respiratory Syndrome Coronavirus 3C-Like Protease. *J. Virol.* **2008**, *82*, 4620–4629. [CrossRef]

7. Thiel, V.; Ivanov, K.A.; Putics, A.; Hertzig, 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]

8. Malcolm, B.A. The picornaviral 3C proteinases: Cysteine nucleophiles in serine proteinase folds. *Prot. Sci.* **1995**, *4*, 1439–1445. [CrossRef]

9. Xia, S.; Liu, M.; Wang, C.; Xu, W.; Lan, Q.; Feng, S.; Feng, S.; Qi, F.; Bao, L.; Du, L.; et al. Inhibition of SARS-CoV-2 infection (previously 2019-nCoV) by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell. Res.* **2020**, *30*, 343–355. [CrossRef]

10. Shanker, A.; Bhanu, D.; Alluri, A. Analysis of Whole Genome Sequences and Homology Modelling of a 3C Like Peptidase and a Non-Structural Protein of the Novel Coronavirus COVID-19 Shows Protein Ligand Interaction with an Aza-Peptide and a Noncovalent Lead Inhibitor with Possible Antiviral Properties. *ChemRxiv* **2020**, 1–39. [CrossRef]

11. Goetz, D.H.; Choe, Y.; Hansell, E.; Chen, Y.T.; McDowell, M.; Jonsson, C.B.; Roush, W.R.; McKerrow, J.; Craik, C.S. Substrate specificity profiling and identification of a new class of inhibitor for the major protease of the SARS coronavirus. *Biochemistry* **2007**, *46*, 8744–8752. [CrossRef]

12. Chuck, C.P.; Chong, L.T.; Chen, C.; Chow, H.F.; Wan, D.C.C.; Wong, K.B. Profiling of substrate specificity of SARS-CoV 3CLpro. *PLoS ONE* **2010**, *5*, e13197. [CrossRef]

13. Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, L.; Pei, J.; Liu, Y.; Chen, J.; et al. Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. *J. Biol. Chem.* **2004**, *279*, 1637–1642. [CrossRef]

14. Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J.R.; Hilgenfeld, R. Coronavirus Main Proteinase (3CLpro) Structure: Basis for Design of Anti-SARS Drugs. *Science* **2003**, *300*, 1763–1767. [CrossRef]

15. Fischer, A.; Sellner, M.; Neranjan, S.; Smieško, M.; Lill, M.A. Potential Inhibitors for Novel Coronavirus Protease Identified by Virtual Screening of 606 Million Compounds. *Int. J. Mol. Sci.* **2020**, *21*, 3626. [CrossRef]

16. Bauer, R.A. Covalent inhibitors in drug discovery: From accidental discoveries to avoided liabilities and designed therapies. *Drug Disc. Today* **2015**, *20*, 1061–1073. [CrossRef]

17. Baillie, T.A. Targeted covalent inhibitors for drug design. *Angew. Chem. Int. Ed.* **2016**, *55*, 13408–13421. [CrossRef]

18. Ghosh, A.K.; Gong, G.; Grum-Tokars, V.; Mulhearn, D.C.; Baker, S.C.; Coughlin, M.; Prabhakar, B.S.; Sleeman, K.; Johnson, M.E.; Mesecar, A.D. Design, synthesis and antiviral efficacy of a series of potent chloropyridyl ester-derived SARS-CoV 3CLpro inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5684–5688. [CrossRef]

19. Shoichet, B.K. Interpreting steep dose-response curves in early inhibitor discovery. *J. Med. Chem.* **2006**, *49*, 7274–7277. [CrossRef]

20. Baell, J.B.; Holloway, G.A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, *53*, 2719–2740. [CrossRef]

21. Saubern, S.; Guha, R.; Baell, J.B. KNIME workflow to assess PAINS filters in SMARTS format. Comparison of RDKit and Indigo cheminformatics libraries. *Mol. Inform.* **2011**, *30*, 847–850. [CrossRef]

22. Walters, W.P.; Stahl, M.T.; Murcko, M.A. Virtual screening—An overview. *Drug Disc. Today* **1998**, *3*, 160–178. [CrossRef]

23. Zhu, T.; Cao, S.; Su, P.C.; Patel, R.; Shah, D.; Chokshi, H.B.; Hevener, K.E. Hit identification and optimization in virtual screening: Practical recommendations based on a critical literature analysis: Miniperspective. *J. Med. Chem.* **2013**, *56*, 6560–6572. [CrossRef]

24. Greenwood, J.R.; Calkins, D.; Sullivan, A.P.; Shelley, J.C. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aid. Mol. Des.* **2010**, *24*, 591–604. [CrossRef]

25. Shelley, J.C.; Cholleti, A.; Frye, L.L.; Greenwood, J.R.; Timlin, M.R.; Uchimaya, M. Epik: A software program for pK a prediction and protonation state generation for drug-like molecules. *J. Comput. Aid. Mol. Des.* **2007**, *21*, 681–691. [CrossRef]

26. Su, H.X.; Yao, S.; Zhao, W.F.; Li, M.J.; Liu, J.; Shang, W.J.; Yu, K.Q. Anti-SARS-CoV-2 activities in vitro of Shuanghuanglian preparations and bioactive ingredients. *Acta Pharmacol. Sin.* **2020**, *41*, 1167–1177. [CrossRef]

27. Konc, J.; Miller, B.T.; Stular, T.; Lesnik, S.; Woodcock, H.L.; Brooks, B.R.; Janežič, D. ProBiS-CHARMMing: Web interface for prediction and optimization of ligands in protein binding sites. *J. Chem. Inf. Model.* **2015**, *55*, 2308–2314. [CrossRef]

28. Konc, J.; Janežič, D. ProBiS tools (algorithm, database, and web servers) for predicting and modeling of biologically interesting proteins. *Prog. Biophys. Mol. Biol.* **2017**, *128*, 24–32. [CrossRef]

29. Krieger, E.; Vriend, G. New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 996–1007. [CrossRef] [PubMed]

30. Jamroz, M.; Kolinski, A. ClusCo: Clustering and comparison of protein models. *BMC Bioinform.* **2013**, *14*, 1–6. [CrossRef]

31. McGann, M. FRED and HYBRID docking performance on standardized datasets. *J. Comput. Aid. Mol. Des.* **2012**, *26*, 897–906. [CrossRef] [PubMed]

32. Gan, J.L.; Kumar, D.; Chen, C.; Taylor, B.C.; Jagger, B.R.; Amaro, R.E.; Lee, C.T. Benchmarking ensemble docking methods as a scientific outreach project. *bioRxiv* **2020**, 1–16. [CrossRef]

33. Lionta, E.; Spyrou, G.; KVassilatis, D.; Cournia, Z. Structure-based virtual screening for drug discovery: Principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014**, *14*, 1923–1938. [CrossRef] [PubMed]

34. Du, J.; Bleylevens, I.W.; Bitorina, A.V.; Wichapong, K.; Nicolaes, G.A. Optimization of Compound Ranking for Structure-Based Virtual Ligand Screening Using an Established FRED–Surflex Consensus Approach. *Chem. Biol. Drug. Des.* **2014**, *83*, 37–51. [CrossRef] [PubMed]

35. Stjernschantz, E.; Oostenbrink, C. Improved ligand-protein binding affinity predictions using multiple binding modes. *Biophys. J.* **2010**, *98*, 2682–2691. [CrossRef] [PubMed]

36. Hansson, T.; Marelius, J.; Åqvist, J. Ligand binding affinity prediction by linear interaction energy methods. *J. Comput. Aid. Mol. Des.* **1998**, *12*, 27–35. [CrossRef]

37. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: Fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **2015**, *43*, W443–W447. [CrossRef]

38. Li, Z.; Li, X.; Huang, Y.Y.; Wu, Y.; Liu, R.; Zhou, L.; Lin, Y.; Wu, D.; Zhang, L.; Liu, H.; et al. Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 27381–27387. [CrossRef]

39. Odhar, H.A.; Ahjel, S.W.; Albeer, A.A.M.A.; Hashim, A.F.; Rayshan, A.M.; Humadi, S.S. Molecular docking and dynamics simulation of FDA approved drugs with the main protease from 2019 novel coronavirus. *Bioinformation* **2020**, *16*, 236. [CrossRef]

40. Krieger, E.; Nielsen, J.E.; Spronk, C.A.; Vriend, G. Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph. Model.* **2006**, *25*, 481–486. [CrossRef]

41. Krieger, E.; Dunbrack, R.L.; Hooft, R.W.; Krieger, B. Assignment of protonation states in proteins and ligands: Combining pK a prediction with hydrogen bonding network optimization. *Methods. Mol. Biol.* **2012**, *819*, 405–421. [PubMed]

42. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef] [PubMed]

43. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef]

44. Jakalian, A.; Jack, D.B.; Bayly, C.I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641. [CrossRef] [PubMed]

45. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725. [CrossRef] [PubMed]

46. Theobald, D.L.; Wuttke, D.S. THESEUS: Maximum likelihood superpositioning and analysis of macromolecular structures. *Bioinformatics* **2006**, *22*, 2171–2172. [CrossRef]

47. Aqvist, J.; Medina, C.; Samuelsson, J.E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385–391. [CrossRef]

48. Klvaňa, M.; Bren, U. Aflatoxin B1-Formamidopyrimidine DNA adducts: Relationships between structures, free energies, and melting temperatures. *Molecules* **2019**, *24*, 150. [CrossRef]

49. Rifai, E.A.; Van Dijk, M.; Vermeulen, N.P.; Yanuar, A.; Geerke, D.P. A Comparative Linear Interaction Energy and MM/PBSA Study on SIRT1–Ligand Binding Free Energy Calculation. *J. Chem. Inf. Model.* **2019**, *59*, 4018–4033. [CrossRef]

Sample Availability: Not available.