



Identification of a novel dual-target scaffold for 3CLpro and RdRp proteins of SARS-CoV-2 using 3D-similarity search, molecular docking, molecular dynamics and ADMET evaluation

Adnane Aouidate^a, Adib Ghaleb^b, Samir Chtita^c, Mohammed Aarjane^d, Abdellah Ousaa^a, Hamid Maghat^a, Abdelouahid Sbai^a, M'barek Choukrad^a, Mohammed Bouachrine^{a,e} and Tahar Lakhli^a

^aMolecular Chemistry and Natural Substances Laboratory, Faculty of Science, Moulay Ismail University, Meknes, Morocco; ^bLaboratory of Analytical and Molecular Chemistry (LCAM), Polydisciplinary Faculty, Cadi Ayyad University, Safi, Morocco; ^cLaboratory physical of chemistry of materials, Faculty of sciences Ben M'Sik, Hassan II University of Casablanca, Casablanca, Morocco; ^dLCBAE, Faculty of Science, University Moulay Ismail, Meknes, Morocco; ^eEst Khenifra, Sultan Moulay Sliman University, Beni Mellal, Morocco

Communicated by Ramaswamy H. Sarma.

ABSTRACT

The new SARS-CoV-2 coronavirus is the causative agent of the COVID-19 pandemic outbreak that affected whole the world with more than 6 million confirmed cases and over 370,000 deaths. At present, there are no effective treatments or vaccine for this disease, which constitutes a serious global health crisis. As the pandemic still spreading around the globe, it is of interest to use computational methods to identify potential inhibitors for the virus. The crystallographic structures of 3CLpro (PDB: 6LU7) and RdRp (PDB 6ML7) were used in virtual screening of 50000 chemical compounds obtained from the CAS Antiviral COVID19 database using 3D-similarity search and standard molecular docking followed by ranking and selection of compounds based on their binding affinity, computational techniques for the sake of details on the binding interactions, absorption, distribution, metabolism, excretion, and toxicity prediction; we report three 4-(morpholin-4-yl)-1,3,5-triazin-2-amine derivatives; two compounds (2001083-68-5 and 2001083-69-6) with optimal binding features to the active site of the main protease and one compound (833463-19-7) with optimal binding features to the active site of the polymerase for further consideration to fight COVID-19. The structural stability and dynamics of lead compounds at the active site of 3CLpro and RdRp were examined using molecular dynamics (MD) simulation. Essential dynamics demonstrated that the three complexes remain stable during simulation of 20 ns, which may be suitable candidates for further experimental analysis. As the identified leads share the same scaffold, they may serve as promising leads in the development of dual 3CLpro and RdRp inhibitors against SARS-CoV-2.

ARTICLE HISTORY

Received 14 May 2020
Accepted 31 May 2020

KEYWORDS

COVID-19; 2019-Novel Coronavirus; molecular docking; molecular dynamics; polymerase; main protease

1. Introduction

The recent emerged coronavirus outbreak of severe asymptomatic pneumonia has geographically appeared as a human pathogen in the city of Wuhan, China. This virus was designated as 2019-nCoV at the very beginning of its emergency in China (Wang et al., 2020). Then, it is named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by the *Coronaviridae* Study Group (CSG) of International Committee on Taxonomy of Viruses (ICTV) based on the novelty and comparative genomic analysis. The SARS-CoV-2 belongs to the family *Coronaviridae*, which includes RNA viruses, and it is the third highest deadly human coronavirus reported in the 21st century followed by the SARS and MERS, expressing the highest fatality rate (Wu et al., 2020). The new virus is closely related (89.1% and 60% nucleotide sequence similarity) to SARS and MERS coronaviruses, respectively. On January 31, 2020, the (world Health Organization) WHO declared COVID-19 as a public health emergency of

international concern. We have witnessed the rapid spread of COVID-19; which is the official name of diseases caused by SARS-CoV-2; that has already gripped all the continents of the globe with multiple epicenters, but the death curve is noticeably higher in America and Europe, with the worst hit reported in USA, Spain and Italy. By Juin, 01 2020, COVID-19 the pandemic has affected more than 6,230,000 confirmed cases, and lead to at least 374,000 deaths over 213 countries and territories.

Some studies investigated potential of combinations of human immunodeficiency virus protease inhibitor lopinavir/ritonavir as antiviral therapy to treat COVID-19. Along with the effort of the respected researchers, we put the credit to Liu et al. who have successfully established crystal structure of main protease (Mpro) or chymotrypsin-like protease (3CLpro) from COVID-19, and deposited in the Protein Data Bank (PDB) for public access (<https://www.rcsb.org/structure/6LU7>). This enzyme plays an essential role in processing of translated polyproteins, such as protease inhibitors are

believed to be the right choice to halt the virus life cycle. Xu et al. reported that nelfinavir was identified as the best potential inhibitor against SARS-CoV-2 3CLpro, based on molecular docking studies among 4 drug compounds (nelfinavir, perampanel, praziquantel and pitavastatin). In addition, RNA-dependent RNA polymerase (RdRp), an essential enzyme in viruses replication, which make it, another key target to find therapeutic agents to COVID-19. Literature studies reported that antiviral drug Remdesivir as a potent inhibitor to virus replication.

While certain physical treatment has been shown to assist patients to fight this disease with their own immune systems, till no there is no proven remedies (antiviral drugs or vaccines) for COVID-19, which is aggravating the situation. Thus, identification or discovery of new effective antivirals is urgently needed to fight the worldwide corona crisis.

Computer-aided drug discovery/design methods (CADD) have played a major role in the development of therapeutically important small molecules for over three decade, screening chemical virtual libraries using computational methods as molecular docking can save money and reduce time (Aouidate et al., 2018a), consequently, speed up the identification of potential drug candidates (Aouidate et al., 2018b; Lipinski et al., 1997). Several research groups have come up with interesting strategies such as repurposing existing drugs or natural products to fight against COVID19(Das et al., 2020; Khan et al., 2020; Mittal et al., 2020), identification of peptide like small molecules using virtual screening of large chemical databases (Pant et al., 2020) or using fragment based approach to design new binders of main protease corona virus (Choudhury, 2020). As we are running of time and the virus is spreading quickly, we have screened the CAS COVID-19 Antiviral Compound Dataset, which includes ~50000 chemical compounds against 3CLpro and RNA-dependent RNA polymerase (RdRp) using computational methods and ligand and structure based screening and in this study, we report the identification of different compounds with CAS IDs (**2001083-69-6**, **2001083-68-5**, **63248-75-9**, **264621-13-8**, **1025098-90-1**, **1253912-09-2**) as potent inhibitors of 3CLpro and (**833463-10-8**, **833463-11-9**, **833465-33-1**, **2001083-69-6**, **833463-19-7**, **833464-45-2**) as potent inhibitors RNA-dependent RNA-polymerase (RdRp), most compounds are 4-(morpholin-4-yl)-1,3,5-triazin-2-amine derivatives, the analysis of SARS-CoV-2 main protease and RNA-dependent RNA polymerase binding sites reveals are combinations of hydrophobic, hydrophilic and charged residues holding with hydrogen bonds in excess, therefore, the 1,3,5-triazine that is aligned centrally in both proteins binding pockets could be a good choice to occupy the central part of the molecules to be substituted by different hydrophobic, hydrophilic and charged fragments.

In this work we identified a number of drug-like compounds, which led to the discovery of several chemical entities that provide new scaffolds that could serve as the core of novel 3CLpro and RdRp 2 inhibitor families. All the four compounds were employed for 20ns MDS study. On the basis of various parameters like RMSD, RMSF, Rg and SASA, we report **2001083-68-5**, **2001083-69-6** and **833463-19-7**

(Figure 6) that they present the same scaffold as lead compounds, which could serve as dual inhibitors for 3CLpro and RdRp, respectively. Though, further *in vitro* and/or *in vivo* research is required to validate the CADD results.

2. Material and methods

2.1. Collection and curation of the chemical library

For VS, the chemical structures of 50000 compound include *antiviral* drugs and related *compounds* that are structurally similar to known *antivirals* (CAS COVID-19 Antiviral Compound Dataset) were downloaded as 3D sdf files from the Chemical Abstract Service database (<https://www.cas.org/covid-19-antiviral-compounds-dataset>) and prepared with an *in-house* KNIME workflow (using Rdkit and 3D-e-Chem nodes)(Kooistra et al., 2018; Landrum, n.d.), first duplicated molecules, compounds with Pan Interference Assay Structures (PAINS) (Dahlin et al., 2015) have been deleted and salts have been stripped out, then best tautomer for every molecule has been generated and optimized, next a modified Lipinski's rule(Lipinski, 2004; Lipinski et al., 1997) (with $300 \leq Mw \leq 700$ g/mol, $5 \leq$ Number of rotatable bonds ≤ 12) has been applied to filter out non drug like molecules, which allowed us to narrow down the database to 5378 drug like compound.

This filtered database has been submitted to a simple-pharmacophore based VS (3D-similarity search (Dobi et al., 2014)) workflow as a first step, and obtained molecules from this step have finally used in a VS using molecular docking.

2.2. 3D-similarity virtual screening

A multi-conformational version of the database was generated; 30 conformer per molecule were generated from the 5378 molecule left from the curation and preparation step; this resulted in a virtual library containing a total of 161256 conformers. An *in-house* KNIME workflow was employed as a first step of VS process, using 3D-generated pharmacophores from N3 and Remdesivir (Figure 1) as search terms to screen the database.

As there are just few inhibitors known for the two new studied proteins of SARS-CoV-2, here we suggest a simple approach to obtain focused databases for 3CLpro (N3 Co-crystallized ligand) and RdRp SARS-CoV-2 targets based on single ligand pharmacophore, unlike other ligand-based VS methods, the main feature of our approach; is using similarity search in 3D (Dobi et al., 2014); is based on two fitting methods, which could enhance the accuracy, the first one based on overlapping the pharmacophores generated from the candidate structure (B from a database) and the query (reference molecule A), then rank them based on the Root Mean Squared Deviation (RMSD), The formula of RMSD is given by the Equation (1):

$$RMSD = \sqrt{\frac{\sum_{i=1}^n (x_A - y_B)^2}{n}} \quad (1)$$

The second one is using $T_{A,B}$ score, which is a similarity between the reference structure, A and database aligned

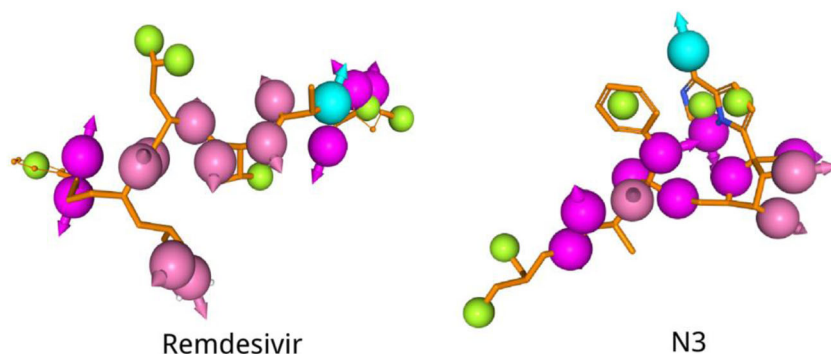


Figure 1. a) 3D-pharmacophores generated from Remdesivir and N3 inhibitors for RNA and 3CLpro respectively. With Green: Aromatic and lipophilic features, Magenta: Hydrogen bond acceptor features, Cyan: Hydrogen bond donor features, Orchid: Hydrogen bond donor and acceptor features.

molecules, B are represented by vector, x , of length n with the i th property having the value x_i (Sheridan, 2007). The formula of Tanimoto coefficient (Dobi et al., 2014) is given by the Equation (2):

$$T_{A,B} = \frac{\sum_{i=1}^n (x_{iA} \cdot y_{iB})}{\sum_{i=1}^n x_{iA}^2 + \sum_{i=1}^n y_{iB}^2 - \sum_{i=1}^n (x_{iA} \cdot y_{iB})} \quad (2)$$

Keeping only the first best 1000 mapped structures from every scoring method, concatenate the results, and finally select the common mapped structures (intersection) between the two databases, which has reduced the database to 205 and 186 molecules for 3CLpro and RdRp, respectively.

2.3. Structure-based virtual screening

The smina software (Koes et al., 2013) was used to carry out molecular docking. The recently resolved 3D structures of SARS-CoV-2 main protease (PDB code: 6LU7 (Chang et al., 2020)) in complex with a the covalent peptide N3 and SARS-CoV-2 RNA dependent RNA polymerase (PDB code: 6LM7) (Gao, 2020) were downloaded from the Protein Data Bank (www.rcsb.org). To prepare the selected proteins for molecular docking, the co-crystallized water and small molecules were removed, protein structures were protonated using reduce tool then non-polar hydrogens were removed using pymol (DeLano, 2002). A grid box with a size of ($x=15.7$, $y=18.2$, $z=21.5$) and center of ($x = -10.782$, $y=15.787$, $z=6.277$) was set to cover the N3 binding site in the 6LU7. While a grid box with a size of ($x=117.9$, $y=116.9$, $z=130.6$) and a center of ($x=21.3$, $y=23.9$, $z=31.1$) was set to cover the binding site in the 6LM7 protein, alongside, the seven conserved motifs (A-G) in the polymerase active site that are involved in a template and nucleotide binding and catalysis (Poch et al., 1989). Validation of the docking process was done by redocking of the co-crystallized ligand (N3) to the 3CLpro binding site and RMS (Root Mean Square) distance between the docked and the experimental co-crystallized binding pose was only 2.73 Å, which is satisfactory in the case of a large peptide as N3.

2.4. Post docking analysis

The results from the structure based VS (for 3CLpro and RdRp) were separately processed using an *in-house* KNIME

workflow (Berthold et al., 2007), firstly molecules were ranked according to the binding energy of their best scoring conformation then the 30 top ranked candidates were inspected visually for their interactions with the hot spot residues, finally selected molecules were submitted for further analysis (ADMET prediction and molecular dynamics). Visual inspection of docking poses and the analysis of protein ligand interactions were performed in Biovia Discovery Studio Visualizer version 2016 (Dassault systems Biovia corp) (Dassault Systèmes BIOVIA, 2016). Visualization images were rendered by PyMOL 2.3 (Schrodinger L.L.C)(DeLano, 2002).

2.5. In silico pharmacokinetics (PK)/pharmacodynamics (PD) evaluation

The success of a drug journey through the body is measured regarding its pharmacokinetics parameters (Adsorption, distribution, metabolism, excretion and toxicity; ADMET). For this purpose, the best-identified *hits* were evaluated for their *in silico* pharmacokinetics parameters using pkCSM (Pires et al., 2015) to prevent the failure of those compounds during clinical studies and enhance their chances to reach the stage of being candidates drugs in future.

2.6. Molecular dynamics

The top hits with best pharmacokinetics properties were shortlisted and further analyzed by molecular dynamic simulations using GROMACS 2018 (Groningen Machine for Chemical Simulations) (Páll et al., 2015; Van Der Spoel et al., 2005). Protonation and structure minimization were performed using the charm36 force field, where hydrogens were added for optimal hydrogen bond network by default. Ligand Topology files were generated using PRODRG server (<http://prodr2.dyndns.org/submit.html>). 3CLpro–ligand and RdRp–ligand complexes were solvated and fully immersed in the center of a cubic box prior using TIP3P water model. The number of water molecules adopted to solvate the complex were as follow: 3CLpro_2001083-69-6 (20,052 molecules); 3CLpro_2001083-68-5 (20,052 molecules); RdRp_833463-10-8 (57,268 molecules); RdRp_833463-19-7 (57,268 molecules). Four and eleven Na⁺ ions were added for the neutralization of the 3CLpro and RdRp systems, respectively, and subsequent energy minimization was performed. The energy

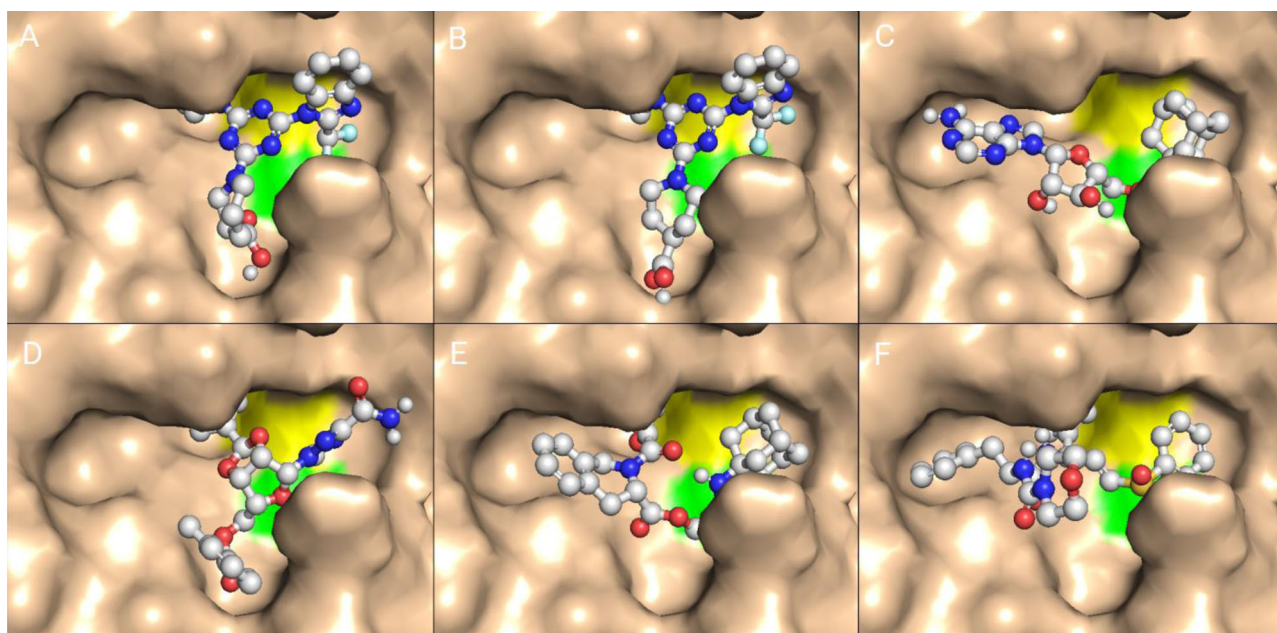


Figure 2. 3D-Binding modes of the six selected candidates in the substrate-binding site of 3CLpro, (A) **2001083-69-6** (-9.064 Kcal/mol), (B) **2001083-68-5** (-8.879 Kcal/mol), (C) **63248-75-9** (-8.857 Kcal/mol), (D) **264621-13-8** (-8.564 Kcal/mol), (E) **1025098-90-1** (-8.558 Kcal/mol) and (F) **1253912-09-2** (-8.400 Kcal/mol). The catalytic dyad Cys145-His41 is colored in yellow and green, respectively.

minimization of systems was performed using the steepest descent minimization of 5,000 steps (maximum number of minimization steps to perform) to remove the weak Van der Waals contacts. Energy minimization was stopped when the maximum force was less than 1.0 kJ/mol. The system was further equilibrated for 50 ps at constant volume and a temperature of 300 K. The molecular dynamic simulations were run for 20,000 ps for each protein–ligand complex, where the coordinates were saved every two ps interval. The calculation of electrostatic interactions were treated by the Particle-Mesh Ewald (PME) (Darden et al., 1993) method. Van der Waals interactions were set at 1.2 nm. LINear Constraint Solver (LINCS) algorithm was applied to constraint the covalent bonds, including heavy atom-H bonds during the molecular dynamics (MD) simulations. The *gmx rms*, *rmsf*, *gyrate*, *gmx hbond*, *gmx sasa* were used for the calculation of RMSD, RMSF, Rg, Hydrogen bonds and SASA respectively.

Lastly, the trajectories were saved for further analysis using the Xmgrace (Turner, n.d.), pymol (DeLano, 2002) and UCSF Chimera (Pettersen et al., 2004).

3. Results and discussion

The aim of the current study is to identify novel potential lead compounds as inhibitors for 3CLpro and RdRp proteins. *In silico* VS of ~50 thousand molecules was performed by establishing a pipeline of PAINS and drug-like filters to assess drug likeness, then N3 and Remdesivir were used as simple-pharmacophores targets for 3D-similarity search to ensure the preparation of focused databases for both proteins, after that compounds that showed a strong binding affinity for 3CLpro and RdRp proteins using molecular docking were selected for further investigations. Further, the ADMET properties were predicted and analyzed to filter top hits and

identify suitable lead candidates, the conformational stability of the best ligands in complex with proteins were studied using MD simulation. The current *in silico* study was undertaken to identify efficient antiviral compounds for COVID19.

3.1. Virtual screening

The prepared chemical database was submitted to a simple 3D-similarity search filter workflow, which was designed by identifying the main features of the N3 and Remdesivir ligands that are involved in interactions with 3CLpro and RdRp (Figure 1), in order to generate focused databases for both proteins. As there is no ligand in complex with the RdRp protein of SARS-CoV-2, which is known to be a target for antiviral drugs. Recommendations of Goa et al.; the authors resolved the structure and in many other studies (Cao et al., 2020; Gao, 2020; Shannon et al., 2020); prompted us to consider Remdesivir as reference molecule for our ligand-based screening company against this target. For 3CLpro the N3 has been considered as reference molecule for our ligand-based VS. The pharmacophore features generated from Remdesivir and N3 are shown in (Figure 1).

Results from the first ligand based VS generated total of 255 and 186 compounds that showed positive result from the two ranking methods RMSD and $T_{A,B}$, so they have been selected as focused databases for 3CLpro and RdRp, respectively.

Smina was used to dock the focused databases from CAS-COVID19-antivirals database. This resulted in a final ranked list, the predicted binding modes of the top 30 molecules were inspected manually for final selection. 6 compounds for 3CLpro (Figure 2) and 6 compounds for RdRp (Figure 4) of this list were chosen for acquisition and ADMET prediction. Table 1 lists their details.

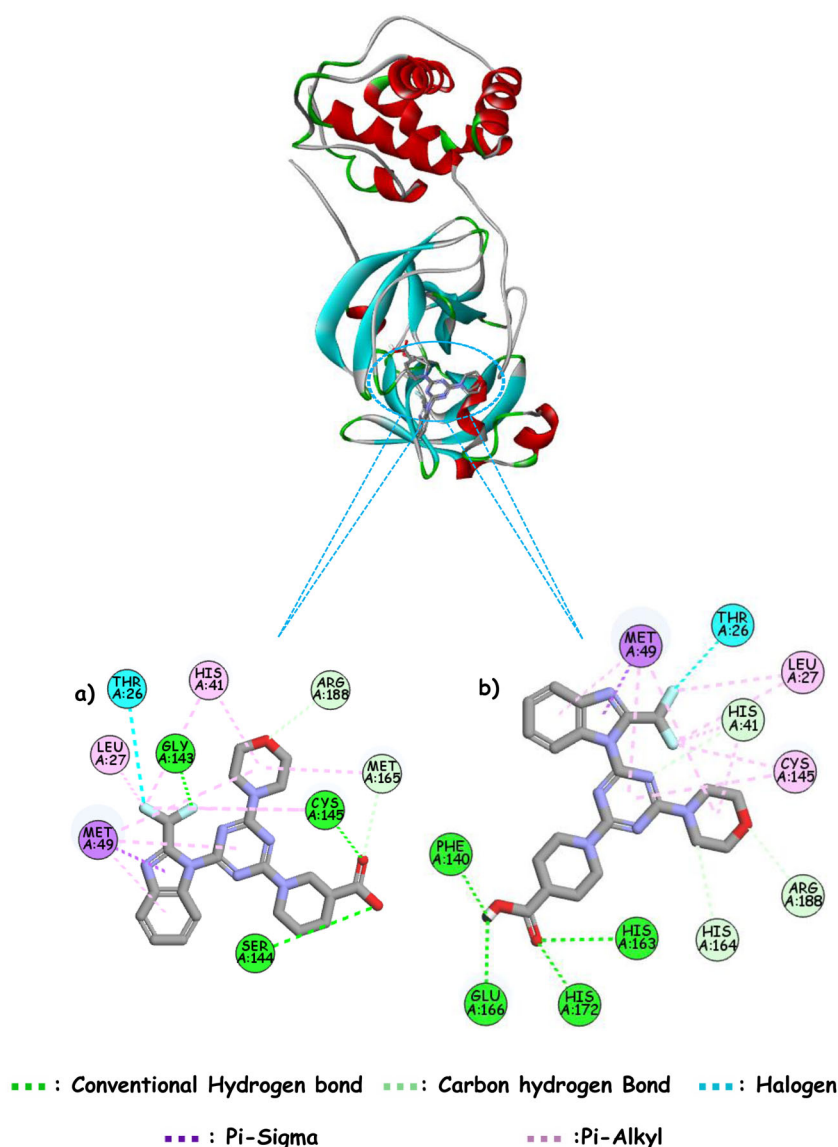


Figure 3. (a) & (b) shows the interaction of top hit molecules Dataset 2001083-69-6 and 2001083-68-5 with main protease of coronavirus for COVID-19.

3.1.1. 3CLpro-ligand complexes

For 3CLpro we observed that most of the hit compounds show common interactions and are involved in H-bond interaction with residues Phe140, Glu143, Asn142, Ser144, Cys145, His163, Glu166, His172, Gln189, Thr190 and hydrophobic interaction with His41, Met49, Met165 which is found consistent with recent studies (Macchiagodena et al., 2020; Ul Qamar et al., 2020).

Among the top 6 hit molecules from CAS COVID19 Dataset, compounds **2001083-69-6** and **2001083-68-5** show the highest affinity -9.064 Kcal/mol and -8.879 Kcal/mol; respectively and also good ADMET analysis (Table 2). Binding interactions of top hit molecules, compounds **2001083-69-6** and **2001083-68-5** with 3CLpro binding pocket are shown in Figure 3(a, b).

Compound **2001083-69-6** forms three strong hydrogen bonds, the first one between the $-NH$ of the Gly143 residue and the Fluorine of the difluoromethyl group ($NH-F$ at distance of 2.73 Å), the second one between Oxygen of the carboxyl group substituent on the piperidinyl moiety and the

$-SH$ of the Cys145 dyad residue ($O-SH$ at distance of 3.65 Å), the third one between the Oxygen of OH in the carboxyl group substituent on the piperidinyl moiety and the hydrogen of the OH group of the Ser144 residue, which can be seen in green dotted lines as shown in Figures 2(a) and 3(a).

Also, 1,3,5-triazine moiety and morpholine ring are aligned centrally to the binding pocket are involved in hydrophobic interaction with dyad residue His41 and Met49 residues.

Compound **2001083-69-5** is stabilized by four strong hydrogen bonds, the two first one are between OH of the carboxyl substituent of the piperidinyl moiety and Oxygen of the carbonyl of the Phe140 and the Glu166 residues ($OH-O$ at distance of 1.89 and 2.48 Å) and the third and fourth ones are between Oxygen of the carbonyl group in the piperidinyl moiety and the $-NH$ of the His163 and His172 residues ($O-NH$ at distance of 2.42 and 3.08 Å), which can be seen in green dotted lines. The molecule is also stabilized by different pi-pi alkyl interactions dyad residues His41 and Cys145 as shown in Figures 2(b) and 3(b).

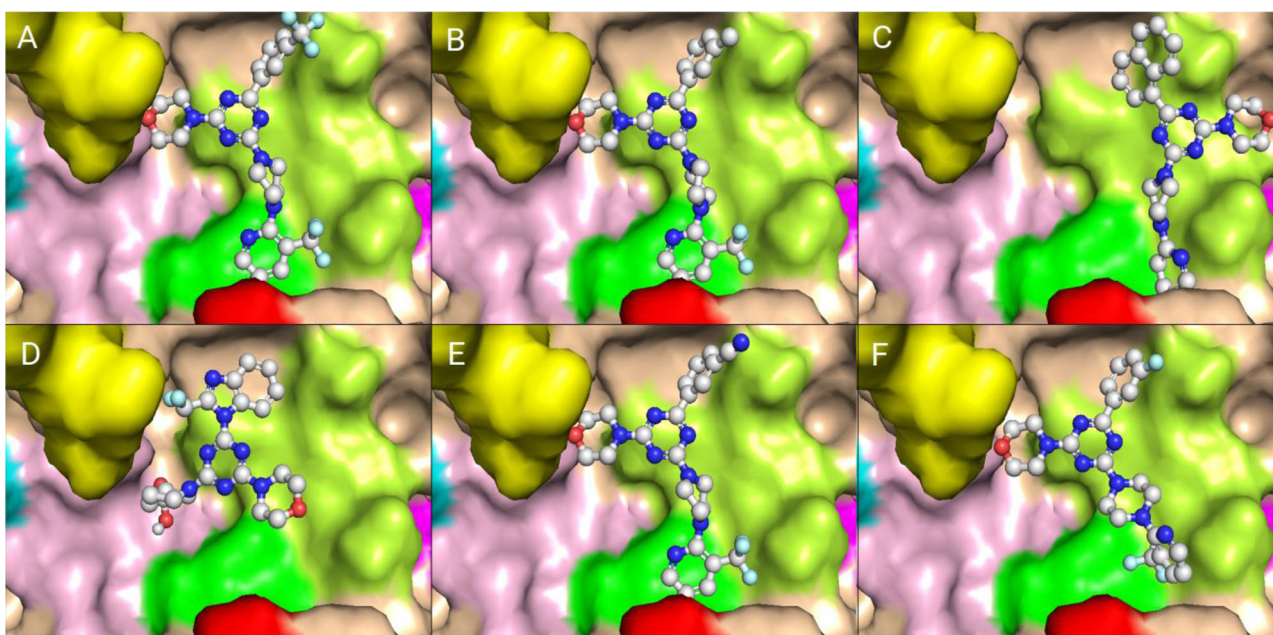


Figure 4. 3D-Binding modes of the six hits compounds in the substrate-binding site of RdRp, (A) **833463-10-86** (−9.529 Kcal/mol), (B) **833463-11-9** (−9.205 Kcal/mol), (C) **833465-33-1** (−9.162 Kcal/mol), (D) **178477-46-8** (−8.895 Kcal/mol), (E) **2001083-69-6** (−8.850 Kcal/mol) and (F) **833463-19-7** (−8.816 Kcal/mol). The motifs (A-G) are colored in A (green), (B) light pink, (C) white green, (D) Magenta, (E) Red, (F) Yellow and (G) Cyan respectively.

Table 1. List of Smina binding affinities for potential COVID19 3CLpro and RdRp.

CAS ID	Hydrogen bonds	Other interactions	Binding affinity
3CLpro inhibitors			
2001083-69-6	Glu143, Cys145, Ser144	His41, Leu27, Arg188, Met165, Thr26, Met49	−9.064
2001083-68-5	Glu166, Phe140, His163, His172	His164, Arg188, Cys145, His41, Leu27, Thr26, Met49	−8.879
63248-75-9	Ser144, Cys145, Asn142, Thr190, Arg188	His41, Leu27, Met165, Gln189	−8.857
264621-13-8	Thr26, Cys145, His163, Glu166	His172, Met49, Met165	−8.564
1025098-90-1	Glu143, Cys145, Ser144, Glu166	Met165, His41, Leu27, Met49	−8.558
1253912-09-2	Glu166, Glu143, Gln189	Cys145, Leu27, Asn142, His41, Met49, Met165, Pro168	−8.400
RdRp inhibitors			
833463-10-8	Tyr455, Arg553, Tyr619	Arg624, Asp623, Lys621, Asp760, Asp618	−9.529
833463-11-9	Arg553, Tyr619	Asp618, Tyr455, Arg624, Asp623, Lys621, Asp760	−9.205
833465-33-1	Lys621	Asp618, Pro620, Tyr455, Arg553, Lys798	−9.162
2001083-69-6	Arg553, Arg624, Thr556, Asn691, Thr687, Thr680	Lys621, Arg555	−8.850
833463-19-7	Arg553, Tyr619, Ser682	Arg624, Lys621, Asp623, Asp618, Asp760	−8.816
833464-45-2	Tyr619	Asp618, Asp760, Arg624, Lys621, Asp623, Arg553	−8.751

Table 2. *In silico* ADMET prediction of identified 3CLpro and RdRp SARS-Cov-2 inhibitors.

CAS ID	Absorption and distribution			Metabolism				Excretion and toxicity		
	Blood-Brain Barrier (LogBB)	Intestinal absorption	Caco-2 permeability	CYP2D6 Substrat	CYP3A4 Substrat	CYP2D6 Inhibitor	CYP3A4 Inhibitor	Clearance	AMES Tox	hERG inhibition
3CLpro inhibitors										
2001083-69-6	−2.06	71.31	0.99	No	No	No	No	0.032	No	No
2001083-68-5	−2.04	70.55	1.012	No	No	No	No	0.068	No	No
63248-75-9	−0.37	86.04	0.699	No	Yes	No	Yes	0.817	No	No
264621-13-8	−1.19	88.25	0.972	No	Yes	No	Yes	0.284	No	No
1025098-90-1	−0.35	94.93	1.285	No	Yes	No	Yes	0.385	No	No
1253912-09-2	−0.88	77.30	0.898	No	Yes	No	Yes	0.696	Yes	No
RdRp inhibitors										
833463-10-8	−1.824	96.449	0.84	No	Yes	No	No	−0.473	No	No
833463-11-9	−1.367	99.29	1.23	No	Yes	No	No	−0.126	No	No
833465-33-1	−0.947	100	1.33	No	Yes	No	Yes	0.26	No	No
2001083-69-6	−2.06	71.31	0.99	No	No	No	No	0.032	No	No
833463-19-7	−1.543	100	1.123	No	Yes	No	No	0.026	No	No
833464-45-2	−1.570	98.730	1.22	No	Yes	No	No	0.333	No	No

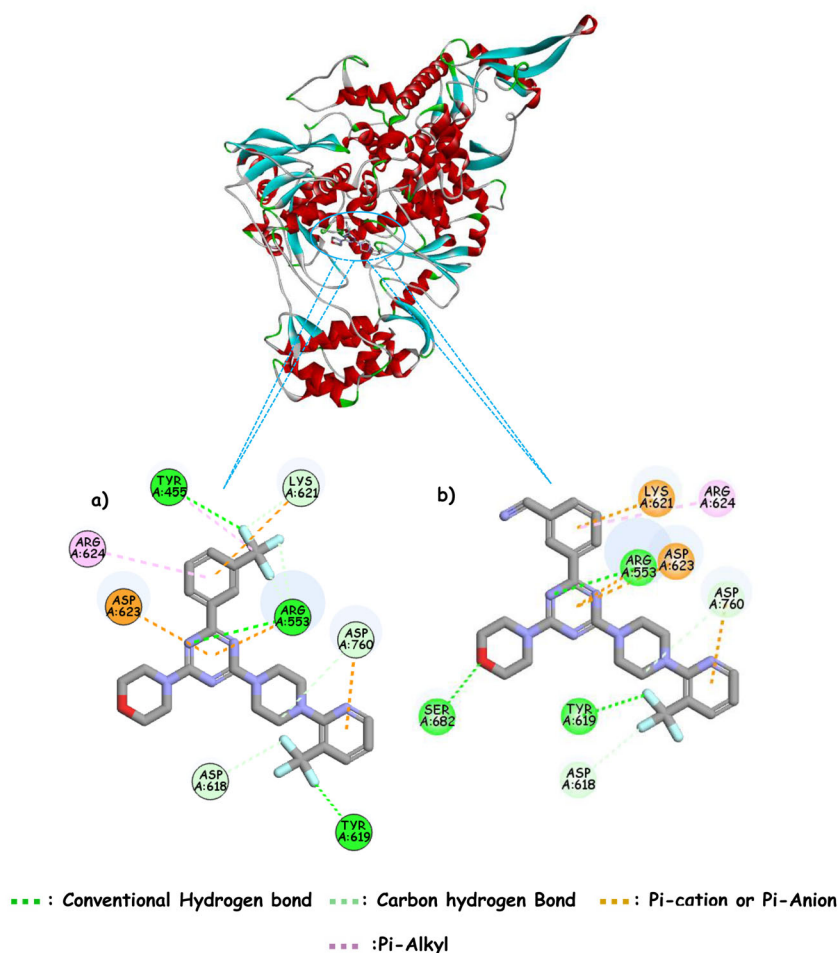


Figure 5. (a) & (b) show the interaction of top hit molecules Dataset **833463-10-8** and **833463-19-7** with RdRp of coronavirus for COVID-19.

3.1.2. RdRp-ligand complexes

For RdRp we observed that most of the hit compounds are involved in interactions with residues functionally important aspartate residue of motif A (Asp623) and motif C (Asp760), together with conserved Arginine residues of motif F (Arg553), which were found highly interacted with the compounds as shown in Table 1. Among these, motif A and C are most strikingly conserved aspartate residues that bind divalent metal ions necessary for catalytic activity.

Based on different binding energies (Table 1) and ADMET properties (Table 2) of different hits, two tops hits **833463-10-8** and **833463-19-7**, which show good ADMET properties in comparison to others were chosen for further investigations.

In the docking analysis, three hydrogen bonds between compound **833463-10-8** and RdRp SARS-CoV-2 binding pocket were formed. The first hydrogen bond 2.84 Å (Tyr455-N-F) was mediated by the Trifluoromethyl group, the second one 2.97 Å (Arg553-NH3 + -N) was formed by the N of the 1,3,5-triazine moiety and the Asp553 of the motif F, and the third 2.36 Å (Tyr 619-NH-F) was formed between the Trifluoromethyl of the pyridine moiety and the NH of the Tyrosine backbone, which can be seen in green dotted lines in Figures 4(a) and 5(a).

The molecule is also stabilized by three carbon hydrogen bonds with Asp618 (motif A), Lys621 (motif A) and Asp760

(motif C), and two Pi-Anion interactions with Asp623 (motif A) and Asp760 (motif C) and two Pi-Cation interactions with Arg553 (motif C) and Lys621 (motif A).

As depicted in Figures 4(f) and 5(b), compound **833463-19-7** is stabilized by three hydrogen bonds between the NH of the backbone Tyr619 (motif A) and Trifluoromethyl group A (Tyr619-NH-F) at distance 2.34 Å and one hydrogen bond with Ser682. In addition 1,3,5-triazine moiety, which is aligned centrally to the binding pocket it is involved in H-bond and Pi-Cation interaction with Arg553 of motif F, one Pi-Cation interaction with Lys621 (motif A), also the molecule is stabilized by a Pi-Cation Asp623(motif A), one hydrogen carbon interaction with Asp760 (motif C).

3.2. ADMET properties

In drug discovery, the prediction of ADMET properties is an important study to escape the failure of drugs in the clinical phases (Kennedy, 1997). Pharmacokinetic and bioavailability predictions are an essential tool in drug discovery process and should be considered to develop a new drug. The absorption, distribution, metabolism, excretion and toxicity predictions results are shown in Table 2, for the pkCSM predictive model, a high caco-2 permeability would translate in predicted values > 0.90, Intestinal absorbance value below 30% indicates poor absorbance. Low value of total clearance

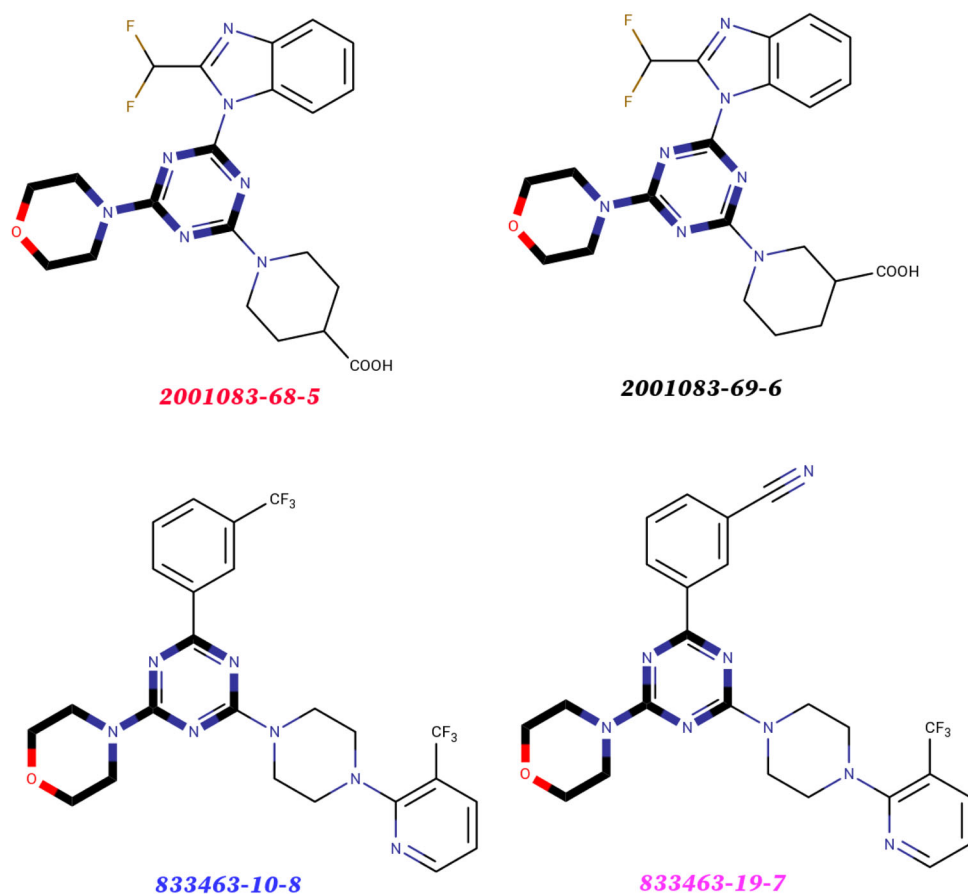


Figure 6. Structure of the top identified lead compounds (Common scaffold is shown in bold).

(logCLtot) means high drug half lifetime. For a given compound a $\log_{BB} < -1$ considered poorly distributed to the brain. Positive result in AMES test suggests that compound could be mutagenic.

For 3CLpro inhibitors, compounds **2001083-69-6**, **2001083-68-5** and **264621-13-8** show excellent absorption and distribution properties, which could be considered as permeable compounds with poor distribution into the brain, also they present good clearance property and show no hERG inhibition nor AMES toxicity. The subtypes of cytochrome P450 CYP2D6 and CYP3A4 indicate that compounds **2001083-69-6** and **2001083-68-5** could not be substrates or inhibitors for the two main subtypes, therefore probably could not be metabolized, thus a low chance of drug-drug interactions. That we suggest them as promising inhibitors for 3CLpro main protease of SARS-CoV-2 and have been selected for MD to investigate their stability of those compounds in RdRp binding site.

Based on the analysis of different ADMET properties of proposed RdRp inhibitors (Table 2), we have found that all six compounds present satisfactory ADMET properties, however, compounds **833463-10-8** and **833463-19-7** present a combination of high intestinal absorption, high Caco-2 permeability and low blood-brain barrier values, which indicate that they show a little chance to cross the blood-brain barrier. As known, cytochrome P450 (CYP) enzymes are heme-proteins that play a key role in the metabolism of drugs; consequently, exploring those parameters may prevent their

failure in clinical studies and have a deep insight on how they react in the body, which allows medicinal chemist to introduce new functional groups on the molecule to prevent the metabolic pathways susceptible to give toxic or very polar compounds that can be eliminated very easily from the body. Accordingly, that can help to synthesize metabolically stable drugs, as well as to avoid drug-drug interactions. As shown in Table 2 most compounds are substrate to CYP3A4 but not inhibitor to CYP2D6 or CYP3A4. Finally, as compounds **833463-10-8** and **833463-19-7** present good absorption, distribution and metabolism properties, also show no AMES mutagenicity or hERG inhibition properties. Thus, they have been chosen as models for MD simulation to investigate their stability in RdRp binding site.

3.3. Conformational dynamics and stability of protein-ligand complex

3.3.1. Root mean squared deviation

The applied molecular docking based virtual screening and the extensive bioavailability analysis lead to the identification of four compounds as models for 3CLpro and RdRp inhibition of coronavirus for COVID19 as shown in Figure 6, 2001083-69-6 and 2001083-68-5 for 3CLpro and **833463-10-8** and **833463-19-7** for RdRp having the most favorable binding interactions, and best pharmacokinetics profiles. Four colors (Black, Red, Blue and Magenta) have been used to make explanation smoother and easier.

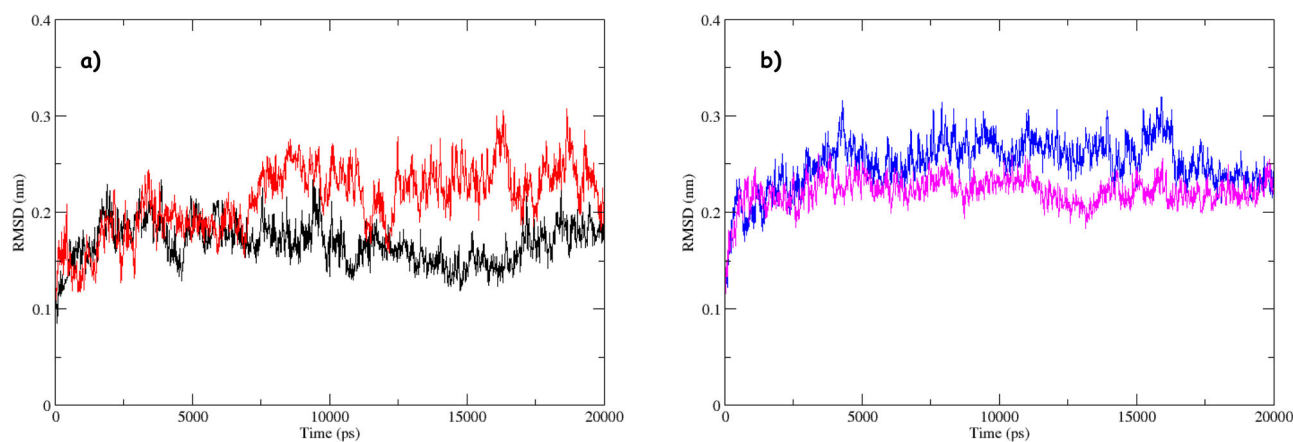


Figure 7. Molecular dynamics simulation. RMSD of the backbone over the 20 ns MDS at 300 K of the complexes systems, the color codes are for (a) 3CLpro-2001083-69-6 (Black) and 3CLpro-2001083-68-5 (Red), (b) RdRp-833463-10-8 (Blue) and RdRp-833463-19-7 (Magenta).

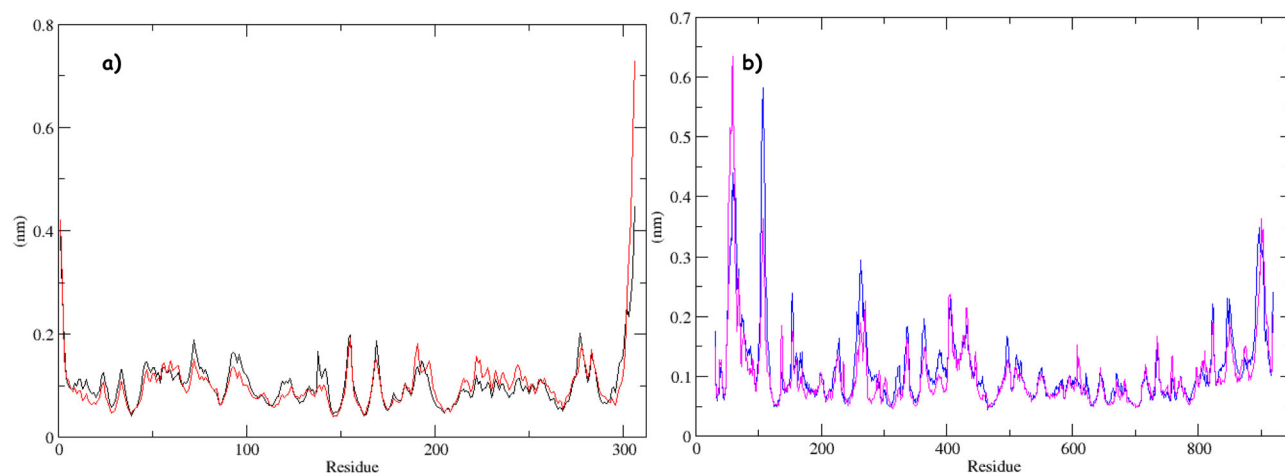


Figure 8. Molecular dynamics simulation. RMSF of the backbone over the 20 ns MDS at 300 K of the complexes systems, the color codes are for (a) 3CLpro-2001083-69-6 (Black) and 3CLpro-2001083-68-5 (Red), (b) RdRp-833463-10-8 (Blue) and RdRp-833463-19-7 (Magenta).

However, the stability of protein-ligand complexes depends on the solvent conditions around the protein and the molecular interactions. It has been observed in several cases, that compounds presenting better docking scores and molecular interactions may fail to bind with protein in experimental results (Azam & Abbasi, 2013). Due to the limitation in computing power, the structural behavior, dynamics and flexibility of 3CLpro-ligand and RdRp-ligand complexes were assessed for 20 ns of MD simulation. The root mean square deviation values of both proteins and ligands were calculated against the initial structure in the protein-ligand complexes and plotted using the Xmgrace software to compare the protein backbone stability.

The backbone of the 3CLpro-2001083-68-5 (Red) complex showed significant fluctuation compared to the 3CLpro-2001083-69-6 (Black) complex as shown in Figure 7(a). The substantial RMSD value can be explained by the mobility of the loops. At the same time, the visual analysis of the trajectory confirms the stability of the secondary structure of the protein in both complexes. The RMSD average value for 3CLpro-2001083-69-6 (Black) (1.6 Å) was lower than that of 3CLpro-2001083-68-5 (Red) (2.12 Å) and it is

interesting to see that the RMSD curve for 3CLpro-2001083-69-6 (Black) is remarkably more stable than this of 3CLpro-2001083-68-5 (Red).

For RdRp-ligand complexes RdRp-833463-10-8 (blue) and RdRp-833463-19-7 (Magenta) shown in Figure 7(b). Overall, the backbone of RdRp-833463-10-8 (blue) shows higher RMSD value compared to RdRp-833463-19-7 (Magenta) and it is interesting to see that the RMSD curves for both protein-ligand complexes after 18 ns present almost the same conformation changes pattern within the acceptable range.

3.3.2. Root mean squared fluctuation

The root-mean-square fluctuation (RMSF) values of backbone C_{α} atoms were calculated by averaging over all the conformations sampled during 20 ns simulation. RMSF values provide insight into structural fluctuations as well as flexibility of different regions of protein, the RMSF fluctuated less, which meant the residues are stable, otherwise the residues are unstable.

The RMSFs of 3CLpro-2001083-69-6 (Black) and 3CLpro-2001083-68-5 (Red) are calculated and shown in Figure 8(a),

most of the proteins' residues in both protein-ligand complexes present RMSF values below 2 Å and the general shape of the fluctuation's curve for both complexes are the same. However, the N-terminal and C-terminal show large conformational changes as seen in [Figure 8\(a\)](#). We found that there are no significant changes between both ligands in the 3CLpro binding pocket.

The RMSFs of RdRp-**833463-10-8** (Blue) and RdRp-**833463-19-7**(Magenta) are calculated and shown in [Figure 8\(b\)](#), the general shape of the fluctuation's curve for both complexes are almost the same. However, the N-terminal and C-terminal show large conformational changes as seen in [Figure 8\(b\)](#). We found that there is no significant change between both ligands in the RdRp binding pocket.

3.3.3. Compactness

The radius of gyration (Rg) factor is used to describe the compactness of the protein-ligand complexes during MD. It is simply a measurement of the distance between the center of mass of the protein atoms with its terminal in a given time frame. Generally, the compact protein or globular protein shows less variation in the gyration value while the expanded form of the structure showed higher Rg value. In this study, variations occurring in Rg values were plotted against time for all the complexes as shown in [Figure 9\(a, b\)](#) for 3CLpro and RdRp, respectively.

For 3CLpro-ligand complexes shown in [Figure 9\(a\)](#), the Rg curves for 3CLpro-**2001083-69-6** (Black) and 3CLpro-**2001083-68-5** (Red) the same type of Rg pattern with a steady behavior during the MD simulation, which indicates that both ligands binding leads the compactness in the 3CLpro protein.

For Rg curves of RdRp-ligand complexes shown in [Figure 9\(b\)](#), after detailed observation and analysis of the Rg plot, we found that after 4 ns, the fluctuation curves of RdRp-**833463-10-8** (Blue) was higher than that of RdRp-**833463-19-7** (Magenta), which indicates the stability of RdRp-**833463-10-8** (Blue) complex in comparison to RdRp-**833463-19-7** (Magenta).

3.3.4. Hydrogen bond analysis

The binding stabilities of the studied top hits compounds in both proteins 3CLpro and RdRp of coronavirus for COVID19 were monitored during the trajectory period of the MD simulations. The stabilities of the protein-ligand complexes were evaluated by calculating the Hydrogen bond profiles using the `g_hbond` tool of Gromacs and the number of hydrogen bonds vs time are plotted in [Figure 10](#).

For 3CLpro-ligand complexes, the analysis revealed that the average numbers of H_bonds observed in **2001083-69-6** (Black) and **2001083-68-5** (Red) in 3CLpro are 1.818 with a maximum of 5 H_bonds and 1.910 with a maximum of 4 H_bonds per timeframe during the MD simulation period ([Figure 10\(a, b\)](#)), respectively. From the above analysis, we conclude that both complexes present good stability.

For RdRp-ligand complexes, the analysis revealed that the average numbers of H_bonds observed in **833463-10-8**

(Blue) and **833463-19-7** (Magenta) in RdRp are 0.645 with a maximum of 4 H_bonds and 0.241 with a maximum of 3 H_bonds per timeframe during the MD simulation period ([Figure 10\(c, d\)](#)), respectively. From the above analysis, we conclude that **833463-10-8** (Blue) forms more hydrogen bonds with RdRp in comparison to **833463-19-7** (Magenta), but other properties and MD trajectories suggest that **833463-10-8** (Blue) leaves the binding pocket while to **833463-19-7** (Magenta) remains inside the binding pocket of RdRp and the hydrogens bonds are mainly produced right there.

3.3.5. RMSD, radius gyration and SASA of ligands

As the RMSF and Radius gyration of the proteins present similar patterns, In addition to H_bond analysis, we have tried to examine three properties to illustrate the stabilities of the selected ligands in the binding pocket of 3CLpro and RdRp proteins during the simulation of 20 ns as shown in [Figure 11\(a-f\)](#) respectively: (1) Ligand Root mean square deviation of a ligand with respect to the reference conformation (the first frame at time = 0 is used as the reference) (RMSD); (2) Radius of Gyration (Rg) - It is used to measure the 'extendedness' of a ligand, and is equivalent to its principal moment of inertia; (3) Solvent Accessible Surface Area (SASA) - the surface area of a molecule accessible by a H₂O molecule; which can help to understand more the behavior of ligands inside binding pockets.

Properties of ligands in 3CLpro are shown in [Figure 11\(a, c, e\)](#):

RMSD values of both ligands **2001083-69-6** (black) and **2001083-68-5** (Red) in the initial stage show fluctuation from 0 ns to 5 ns but later maintains constant during the simulation. The overall RMSD of these ligands were below 2 Å. However, the RMSD value and fluctuation of **2001083-68-5** were higher than **2001083-69-6**. This analysis indicates that compound **2001083-69-6** (Black) in 3CLpro maintained constant RMSD during the simulation after 5 ns compared to **2001083-68-5**(Red).

Also, Rg-time variation shown in [Figure 11\(c\)](#) for both ligands **2001083-69-6** (Black) and **2001083-68-5** (Red) in the receptor binding pocket of 3CLpro protein were observed almost constants, the **2001083-69-6** (Black) compound Rg average is 4.2 Å, which is lower than this of **2001083-68-5** (Red) (Rg average 4.45 Å), However, the Rg-time diagrams suggested steady conformation changes for both ligands.

After detailed observation, for SASA plots shown in [Figure 11\(e\)](#) for both ligands **2001083-69-6** (Black) and **2001083-68-5** (Red), we found that there is no significant change between them in the 3CLpro binding pocket, because they are both present the same type pattern.

Properties of ligands in RdRp are shown in [Figure 11\(b, d, f\)](#):

RMSD values of both ligands **833463-10-8** (Blue) and **833463-19-7**(Magenta) are shown in [Figure 11\(b\)](#). The fluctuations curves of **833463-10-8** (Blue) are higher than that of **833463-19-7** (Magenta) and it is of interest to see that the

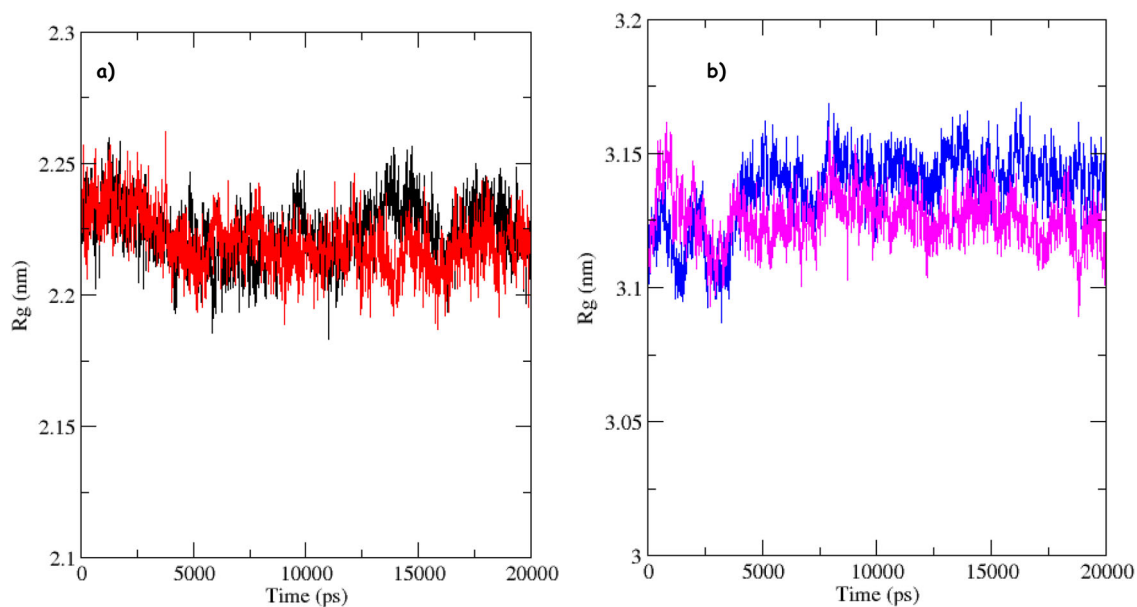


Figure 9. Molecular dynamics simulation. Radius gyrate of the backbone over the 20 ns MDS at 300 K of the complexes systems, the color codes are for (a) 3CLpro-2001083-69-6 (Black) and 3CLpro-2001083-68-5 (Red), (b) RdRp-833463-10-8 (Blue) and RdRp-833463-19-7(Magenta).

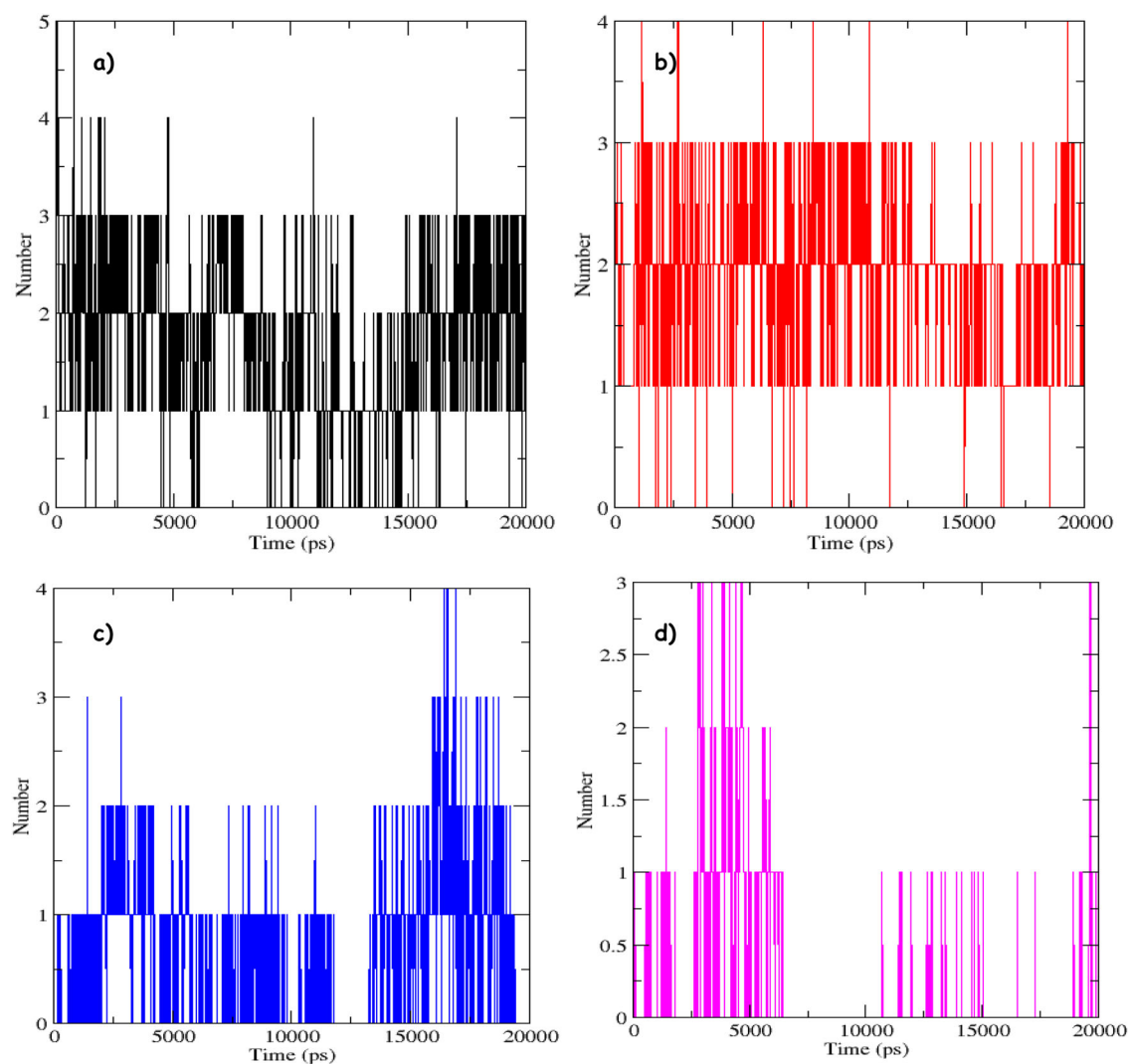


Figure 10. Number of hydrogen bonds of the complex of 3CLpro of coronavirus for COVID-19 with (a) 2001083-69-6 (Black) and (b) 2001083-68-5 (Red) and RdRp of coronavirus for COVID-19 with (c) 833463-10-8 and 833463-19-7(Magenta) as a function of time are retained or broken during simulation time scale 20 ns.

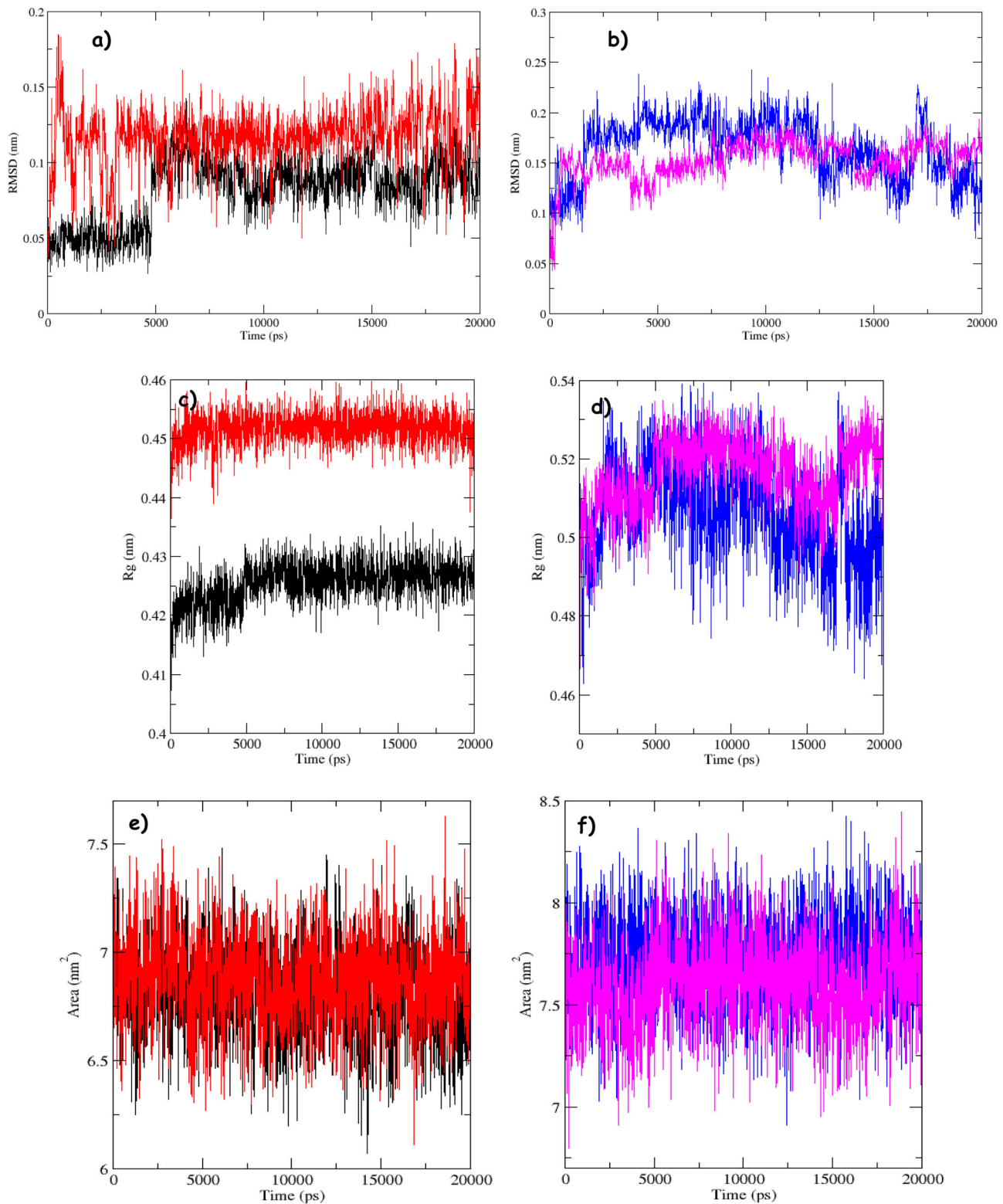


Figure 11. Molecular dynamics simulation. Variation in the ligands properties during the course of 20 ns MDS at 300 K, the color codes are for **2001083-69-6** (Black) and **2001083-68-5** (Red) in 3CLpro; **833463-10-8** (Blue) and **833463-19-7**(Magenta) in RdRp.

RMSD curve for **6833463-19-7**(Magenta) is remarkably stable than that of **833463-10-8** (Blue).

For Rg variation for both ligands **833463-10-8** (blue) and **833463-19-7**(Magenta) in the binding pocket of RdRp protein are shown in Figure 11(d). Rg variation of **833463-10-8** (blue) shows higher fluctuations during the 20 ns simulation when compared to **833463-19-7**(Magenta). The Rg-time

plots suggested a steadier conformation of **833463-19-7**(Magenta) than **833463-10-8** (blue).

Also, SASA plot of both ligands **833463-10-8** (blue) and **833463-19-7**(Magenta) shown in Figure 11(d), indicates that variation of **833463-10-8** (blue) shows higher fluctuations during the 20 ns simulation when compared to **833463-19-7**(Magenta). For the ligand **833463-10-8** (blue), the average

value of SASA is 7.724nm^2 during throughout the 20 ns simulation, while the ligand **833463-19-7**(Magenta) shows SASA of 7.603nm^2 during the simulation. In overall **833463-19-7**(Magenta) presents less value of SASA compared to **833463-10-8** (blue) because of the consistency of the **833463-19-7** (Magenta) in the binding pocket of the protein.

MD simulation results indicate that, both compounds **2001083-68-5** and **2001083-69-6** form stable complexes with 3CLpro, with higher stability of 3CLpro-**2001083-69-6** complex. Also **833463-19-7** forms a stable complex with RdRp, the three compounds remain close to their initial docking positions even in unrestrained simulations, hence suggesting formation of stable complexes. All of these indicating that the complexes of 3CLpro-**2001083-68-5**, 3CLpro-**2001083-69-6** and RdRp-**833463-10-8** are very likely to have an inhibition activity of SARS-CoV-2, especially they present similar scaffold (4-(morpholin-4-yl)-1,3,5-triazin-2-amine) as depicted in in bold (Figure 6), which could serve as dual inhibition of both proteins 3CLpro and RdRp of coronavirus COVID19.

4. Conclusions

In view of the fact that 3CLpro and RdRp inhibitors play a significant role for the treatment of COVID19, which is an international crisis, we have conducted a computer aided drug design study to rapidly and efficiently screen optimal ligands from the 50000 antiviral compounds in (CAS COVID-19 Antiviral Compound Chemical Database). Based on binding affinity and interactions analysis, six compounds were shortlisted for each protein most of them with a 4-(morpholin-4-yl)-1,3,5-triazin-2-amine scaffold. Further, based on ADMET analysis, among the shortlisted, four molecules with best ADMET properties and least toxic behavior (**2001083-69-6** and **2001083-68-5** in 3CLpro) and (**833463-10-8** and **833463-19-7** in RdRp) were selected as models for MD simulation to illustrate the ligand-proteins stability. MD simulation and analysis of four lead compounds, two in complex with 3CLpro revealed that both of the them formed stable complexes in the ligand-binding pocket of 3CLpro, while for the two compounds in complex with RdRp, only one ligand formed a stable during 20-ns simulation.

It is expected that the new dual 3CLpro and RdRp inhibitors for coronavirus SARS-CoV-2(COVID19 disease) may become potential drug candidates. Or at least, they may stimulate a new way for developing novel inhibitors against COVID19. However, further *in vitro* studies are required to validate the findings

Disclosure statement

Authors declare no conflict of interest.

Funding

Cet article entre dans le cadre du programme de CNRST (Maroc): PROGRAMME DE SOUTIEN A LA RECHERCHE SCIENTIFIQUE ET TECHNOLOGIQUE EN LIEN AVEC LE "COVID-19". Intitulé du projet: Recherche de nouvelles molécules candidates pour des médicaments

anti-Covid-19 à base de plantes médicinales marocaines ou d'autres sources.

References

- Aouidate, A., Ghaleb, A., Ghamali, M., Chtita, S., Ousaa, A., Choukrad, M., Sbai, A., Bouachrine, M., & Lakhlifi, T. (2018a). Computer aided drug design based on 3D - QSAR and molecular derivatives as PIM2 inhibitors: A proposal to chemists. *In Silico Pharmacology*, 6(1), 5. <https://doi.org/10.1007/s40203-018-0043-7>
- Aouidate, A., Ghaleb, A., Ghamali, M., Chtita, S., Ousaa, A., Choukrad, M., Sbai, A., Bouachrine, M., & Lakhlifi, T. (2018b). QSAR study and rustic ligand-based virtual screening in a search for aminooxadiazole derivatives as PIM1 inhibitors. *Chemistry Central Journal*, 12(1), 1–12. <https://doi.org/10.1186/s13065-018-0401-x>
- Azam, S. S., & Abbasi, S. W. (2013). Molecular docking studies for the identification of novel melatoninergic inhibitors for acetylserotonin-O-methyltransferase using different docking routines. *Theoretical Biology & Medical Modelling*, 10(1), 63. <https://doi.org/10.1186/1742-4682-10-63>
- Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinel, T., Ohl, P., Sieb, C., Thiel, K., & Wiswedel, B. (2007). KNIME: The Konstanz information miner. In *Studies in classification, data analysis, and knowledge organization* (GfKL 2007). Springer.
- Cao, Y., Deng, Q., & Dai, S. (2020). Remdesivir for severe acute respiratory syndrome coronavirus 2 causing COVID-19: An evaluation of the evidence. *Travel Medicine and Infectious Disease*, 101647. <https://doi.org/10.1016/j.tmaid.2020.101647> (p. 101647).
- Chang, Y., Tung, Y., Lee, K., Chen, T., Hsiao, Y., Chang, C., Hsieh, T., Su, C., Wang, S., Yu, J., Lin, Y., Lin, Y., Tu, Y. E., & Tung, C. (2020). Potential therapeutic agents for COVID-19 based on the analysis of protease and RNA polymerase docking. *Preprints 2020, 2020020242. Virology Journal*, 1–7. <https://doi.org/10.20944/preprints202002.0242.v1>
- Choudhury, C. (2020). Fragment tailoring strategy to design novel chemical entities as potential binders of novel corona virus main protease. *Journal of Biomolecular Structure and Dynamics*, 38, 1–14. <https://doi.org/10.1080/07391102.2020.1771424>
- Dahlin, J. L., Nissink, J. W. M., Strasser, J. M., Francis, S., Higgins, L., Zhou, H., Zhang, Z., & Walters, M. A. (2015). PAINS in the assay: Chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a Sulfhydryl-scavenging HTS. *Journal of Medicinal Chemistry*, 58(5), 2091–2113. <https://doi.org/10.1021/jm5019093>
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089–10092. <https://doi.org/10.1063/1.464397>
- Das, S., Sarmah, S., Lyndem, S., & Singha Roy, A. (2020). An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *Journal of Biomolecular Structure and Dynamics*, 38, 1–11. <https://doi.org/10.1080/07391102.2020.1763201>
- Dassault Systèmes BIOVIA. (2016). *Discovery studio modeling environment, release 2017*. Dassault Systèmes.
- DeLano, W. (2002). *The PyMOL molecular graphics system*. DeLano Scientific.
- Dobi, K., Hajdú, I., Flachner, B., Fabó, G., Szaszko, M., Bognár, M., Magyar, C., Simon, I., Szisz, D., Lórinicz, Z., Cseh, S., & Dormán, G. (2014). Combination of 2D/3D ligand-based similarity search in rapid virtual screening from multimillion compound repositories. selection and biological evaluation of potential PDE4 and PDE5 inhibitors. *Molecules (Basel, Switzerland)*, 19(6), 7008–7039. <https://doi.org/10.3390/molecules19067008>
- Gao, Y. (2020). *RCSB PDB - 6M71: SARS-Cov-2 RNA-dependent RNA polymerase in complex with cofactors*. <https://www.rcsb.org/structure/6M71>
- Kennedy, T. (1997). Managing the drug discovery/development interface. *Drug Discovery Today*, 2(10), 436–444. [https://doi.org/10.1016/S1359-6446\(97\)01099-4](https://doi.org/10.1016/S1359-6446(97)01099-4)

- Khan, R. J., Jha, R. K., Amera, G. M., Jain, M., Singh, E., Pathak, A., Singh, R. P., Muthukumar, J., & Singh, A. K. (2020). Targeting SARS-CoV-2: A systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2'-O-ribose methyltransferase. *Journal of Biomolecular Structure and Dynamics*, 38, 1–14. <https://doi.org/10.1080/07391102.2020.1753577>
- Koes, D. R., Baumgartner, M. P., & Camacho, C. J. (2013). Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. *Journal of Chemical Information and Modeling*, 53(8), 1893–1904. <https://doi.org/10.1021/ci300604z>
- Kooistra, A. J., Vass, M., McGuire, R., Leurs, R., de Esch, I. J. P., Vriend, G., Verhoeven, S., & de Graaf, C. (2018). 3D-e-Chem: Structural cheminformatics workflows for computer-aided drug discovery. *ChemMedChem*, 13(6), 614–626. <https://doi.org/10.1002/cmdc.201700754>
- Landrum, G. (n.d.). *RDKit: Open-source cheminformatics*.
- Lipinski, C. A. (2004). Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discovery Today. Technologies*, 1(4), 337–341. <https://doi.org/10.1016/j.ddtec.2004.11.007>
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 23(1–3), 3–25. [https://doi.org/10.1016/S0169-409X\(96\)00423-1](https://doi.org/10.1016/S0169-409X(96)00423-1)
- Macchiagodena, M., Pagliai, M., & Procacci, P. (2020). Identification of potential binders of the main protease 3CLpro of the COVID-19 via structure-based ligand design and molecular modeling. *Chemical Physics Letters*, 750, 137489. <https://doi.org/10.1016/j.cplett.2020.137489>
- Mittal, L., Kumari, A., Srivastava, M., Singh, M., & Asthana, S. (2020). Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach. *Journal of Biomolecular Structure and Dynamics*, 38, 1–19. <https://doi.org/10.1080/07391102.2020.1768151>
- Páll, S., Abraham, M. J., Kutzner, C., Hess, B., & Lindahl, E. (2015). *Tackling exascale software challenges in molecular dynamics simulations with GROMACS Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8759, 3–27. https://doi.org/10.1007/978-3-319-15976-8_1
- Pant, S., Singh, M., Ravichandiran, V., Murty, U. S. N., & Srivastava, H. K. (2020). Peptide-like and small-molecule inhibitors against Covid-19. *Journal of Biomolecular Structure and Dynamics*, 38, 1–10. <https://doi.org/10.1080/07391102.2020.1757510>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, 58(9), 4066–4072. <https://doi.org/10.1021/acs.jmedchem.5b00104>
- Poch, O., Sauvaget, I., Delarue, M., & Tordo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *The EMBO Journal*, 8(12), 3867–3874. <https://doi.org/10.1002/j.1460-2075.1989.tb08565.x>
- Shannon, A., Le, N. T.-T., Selisko, B., Eydoux, C., Alvarez, K., Guillemot, J.-C., Decroly, E., Peersen, O., Ferron, F., & Canard, B. (2020). Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral Research*, 178, 104793. <https://doi.org/10.1016/j.antiviral.2020.104793>
- Sheridan, R. P. (2007). Chemical similarity searches: When is complexity justified? *Expert Opinion on Drug Discovery*, 2(4), 423–430. <https://doi.org/10.1517/17460441.2.4.423>
- Turner, P. J. (n.d.). *XMGRACE (5.1.25)*. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology.
- Ul Qamar, M. T., Alqahtani, S. M., Alamri, M. A., & Chen, L.-L. (2020). Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants [published online ahead of print, 2020 Mar 26]. *Journal of Pharmaceutical Analysis*, 1–7. <https://doi.org/10.1016/j.jpha.2020.03.009>
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16), 1701–1718. <https://doi.org/10.1002/jcc.20291>
- Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *The Lancet*, 395(10223), 470–473. [https://doi.org/10.1016/S0140-6736\(20\)30185-9](https://doi.org/10.1016/S0140-6736(20)30185-9)
- 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., Yuan, M.-L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E. C., & Zhang, Y.-Z. (2020). A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798), 265–269. <https://doi.org/10.1038/s41586-020-2008-3>