

## INITIAL STUDY ON SARS-COV-2 MAIN PROTEASE INHIBITION MECHANISM OF SOME POTENTIAL DRUGS USING MOLECULAR DOCKING SIMULATION

Pham Minh Quan<sup>1,2,\*</sup>, Le Thi Thuy Huong<sup>1,2</sup>, Tran Quoc Toan<sup>1,2</sup>,  
Ngo Son Tung<sup>3,4</sup>, Nguyen Trong Dan<sup>2,5</sup>, Tran Thi Thu Thuy<sup>1,2</sup>,  
Nguyen Manh Cuong<sup>1,2</sup>, Pham Quoc Long<sup>1,2,\*</sup>

<sup>1</sup>*Institute of Natural Products Chemistry, Vietnam Academy of Science and Technology (VAST),  
18 Hoang Quoc Viet, Ha Noi, Viet Nam*

<sup>2</sup>*Graduate University of Science and Technology, VAST, 18 Hoang Quoc Viet, Ha Noi, Viet Nam*

<sup>3</sup>*Laboratory of Theoretical and Computational Biophysics, Ton Duc Thang University,  
Ho Chi Minh City, Viet Nam*

<sup>4</sup>*Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Viet Nam*

<sup>5</sup>*Vietnam-Russia Tropical Center, Nguyen Van Huyen, Cau Giay, Ha Noi, Viet Nam*

\*Emails: [pham-minh.quan@inpc.vast.vn](mailto:pham-minh.quan@inpc.vast.vn), [mar.biochem@fpt.vn](mailto:mar.biochem@fpt.vn)

Received: 21 March 2020; Accepted for publication: 10 August 2020

**Abstract.** The infection by the new coronavirus SARS-CoV-2 (called as COVID-19 disease) is a worldwide emergency, however, there is no antiviral treatment or vaccine to date. 3C like protease (3CLpro) is the main protease of SARS-CoV-2 that involved in the process of translation of the polypeptide from the genomic RNA to protein components, which are required for virus replication. The crystal structure of this protease has been rapidly resolved and made publicly in the Protein Data Bank recently. Many efforts have been conducted by scientists including the use of several commercial medicines that are known for treatment of HIV and anti-malarial/antibiotic such as arbidol, chloroquine, hydroxychloroquine, azithromycin, darunavir, remdesivir and lopinavir/ritonavir. These drugs exhibited significant efficacy in clinical, however, the understanding at atomic level of how these compounds prevent SARS-CoV-2 protease is still lacking. Therefore, in this context docking protocol was employed to rapidly estimate the binding affinity and binding pose of six drugs on the main protease of SARS-CoV-2 virus. The obtained results might help to shed light on the interaction mechanism of these compounds toward the protein, and thus suggesting an efficient approach to drug discovery and treatments.

**Keywords:** COVID-19, SARS-CoV-2, autodock4, protease inhibitor, molecular docking.

**Classification numbers:** 1.2.1, 1.2.4.

## 1. INTRODUCTION

The severe acute respiratory syndrome CoV (SARS-CoV), Middle East respiratory syndrome CoV (MERS-CoV) and new coronavirus (SARS-CoV-2) belong to *Betacoronavirus* which contain a single positive-stranded RNA from 26 to 32 kb in length and cause a wide array of respiratory, gastrointestinal and neurological diseases in human hosts [1, 2]. In December 2019, the World Health Organisation (WHO) officially announced a cluster of cases detected in Wuhan city, Hubei province of China [3] and today, the infection has spread out to more than 213 countries and territories. The total number of confirmed COVID-19 infections are 15,096,315 cases, and the number of deaths reached 619,520 as of July 22, 2020 [4] which show no sign of decline in the near future. The rapid increase in numbers of infected patients urge scientific community to find vaccines/drugs to cure this disease. However, this process could be time-consuming and take many years to complete as the traditional research pathway and safety test of new developed vaccines/drugs could be a major concern. In the meanwhile, finding an alternate therapy to treat infections and reduce death cases is necessary.

There have been many research efforts conducted by scientists worldwide to find an efficient pathway to treat the disease, especially focusing on potential drug targets. The spike protein of SARS-CoV-2 has been quickly identified as novel target for drug development due to its high affinity in binding to human cell receptor ACE2, thus initiating molecular events that release the viral genome intracellularly [5, 6]. On the other hand, coronaviruses usually encode two or three viral proteases. In the cases of SARS-CoV, the two identified proteases are a papain-like cysteine protease (PL<sup>pro</sup>) [7] and a chymotrypsin-like cysteine protease (3CL<sup>pro</sup>) [8], also known as the main protease. There are multiple domain functions that are active in the replication of the coronavirus and these domains are presented in a protein designated as non-structural protein 3 (nsp3) which is the largest protein in the coronavirus genome [9]. 3C like protease (3CL<sup>pro</sup>) is proven to be involved in the process of translation of the polypeptide from the genomic RNA to protein components that are required structurally or non-structurally for replication and packaging of new generation viruses. Notably, the crystal structure of the COVID-19 main protease in complex with a peptidomimetic inhibitor (PDB ID: 6LU7) was determined recently by the scientists from China and made available on the Protein Data Bank [10]. The RNA genome of the new coronavirus was reported with up to 82 % identity to that of SARS-CoV. Thanks to untiring research efforts, several medicines including arbidol, chloroquine, hydroxychloroquine, azithromycin, darunavir, remdesivir and lopinavir/ritonavir, which are known in the therapy of HIV and anti-malarial/antibiotic infection, have been used in the treatment of SARS-CoV infections and resulted in high efficacy [11 - 13].

Nowadays, computer-aided drug design is usually used to search for new potential drug candidates and one of the popular methods is molecular docking which focus on analyzing the interaction between receptors and compounds [14]. This technique explores the mechanism which inhibit targeted receptors based on binding free energy and amino acids participated in interaction, hence helping the scientists to speed up the time and also save costs on new drugs development. The application of this method has resulted in many publications that predict potential bioactive compounds against SARS-CoV-2 specific biological targets recently [15 - 19]. With the hope to quickly identify candidate drugs for COVID-19 therapy, we conduct molecular docking using Autodock4 tool to simulate the interaction between some potential medicines and main protease of SARS-CoV-2. The obtained results might reveal the mechanism of inhibiting the target protease, thus, giving additional chance to find promising drugs quickly.

## 2. MATERIALS AND METHODS

### 2.1. Protein preparation

Crystal structure of COVID-19 main protease in complex with a peptidomimetic inhibitor N3 (PDB ID: 6LU7) was obtained from Protein Data Bank [10]. Autodock Tools (MGLTools) was utilized to prepare protein for docking simulations [20]. To turn the protein molecule into a free receptor, water molecules and standard inhibitor were removed. Then, the polar hydrogen atoms, default Kollman charges and solvation parameters were allocated to the protein atoms [21]. Obtained atomic coordinates of the protein were then exported into a PDBQT file which will be used for the execution of AutoGrid and AutoDock.

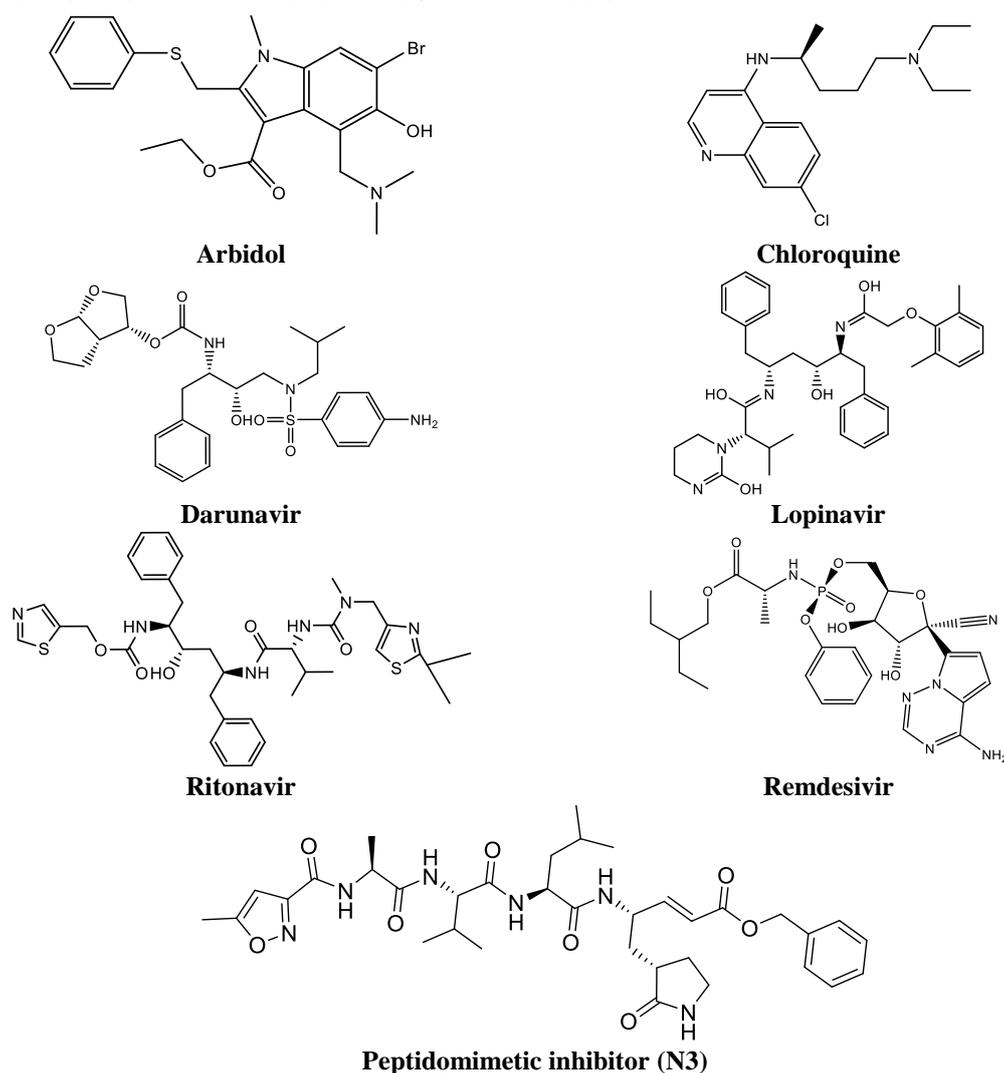


Figure 1. Chemical structure of studied compounds.

### 2.2. Ligand preparation

The structure of 6 studied medicines including arbidol, chloroquine, darunavir, remdesivir and lopinavir/ritonavir were collected from literatures (Figure 1) [13, 22 - 24]. To draw chemical structures of the triterpenes, Marvin software was used. The 3D structure of compounds were built using Pymol 2.2.2 [25]. The energy minimization was carried out using Gabedit 2.5.0 and Chemicalize webserver [26, 27].

### 2.3. Protein structure validation

Before performing molecular docking studies, the COVID-19 main protease crystal structure (PDB ID: 6LU7) was cross-checked with another known structure of SARS-CoV main protease published in 2005 (PDB ID: 2BX4) [28] to clarify whether the new coronavirus induces amino acids mutation within the active site. The active site of SARS-CoV 3CL<sup>PRO</sup> is located in the cleft between domains I and II of the protein, and the binding active site of this protease is composed of essential amino acids such as His41, Met49, Cys145, His163, Glu166, His172 [29]. The validation was carried out using plugins from Chimera 1.13.1 [30] and PyMol [25].

### 2.4. Molecular docking studies using AutoDock4.2.6

Four softwares including PyMOL [25], Discovery Studio Visualizer [31], LigPlus [32] and Maestro [33] were used to analyze the obtained results, which describe distances of hydrogen bonds between the hydrogen and its supposed binding partner. The grid box that encloses amino acids domain involved in the binding active sites, had the dimension of  $50 \times 60 \times 60$  (x  $\times$  y  $\times$  z) with grid spacing of 0.375 Å (Figure 2). AutoGrid and AutoDock were used to calculate the pre-calculated binding affinity of each ligand's atom type and to perform molecular docking simulation, respectively. The parameters of the Lamarckian Genetic Algorithm (LGA) were: 50 runs; elitism of 1; the mutation rate of 0.02; the population size of 300; a crossover rate of 0.80; number of generations of 27,000; the energy evaluations of 50,000,000 and the root-mean-square (RMS) cluster tolerance was set to 2.0 Å in each run. Default parameters were selected for step sizes for translations, quaternions and torsions. The ligand conformation with the lowest free energy of binding, chosen from the most favored cluster, was selected for further analysis.

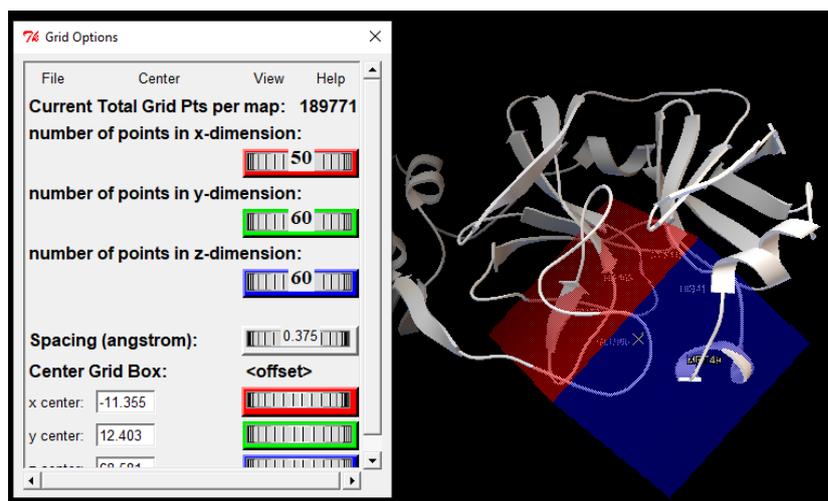


Figure 2. The setup coordinates grid (Grid box) for molecular docking studies on binding site of COVID-19 main protease.



Molecular docking simulation is an important method for understanding various interactions between ligand and protein/enzyme active site, which is helpful in drug discovery in the pharmaceutical industry. In an attempt to find potential compounds for COVID-19 treatments, AutoDock4.2.6, with the Lamarckian Genetic Algorithm, were used to analyze the docking probability of 6 medicines currently used in clinical on the binding pocket of the main protease which play an important role in the propagation of the virus. Obtained results could reveal pivotal information on the inhibition mechanism of the compounds toward target protein, thus, providing helpful suggestion in structure characteristics for further drug development. All the docked compounds were compared to each other and the ranking were sorted from the lowest to the highest binding energy (Table 1). It is observed that peptidomimetic inhibitor N3 has binding energy -13.13 kcal/mol and will be considered as threshold value for this docking study, and thus any ligands whose docking energies are close to this value would be viewed as potential. The standard inhibitor formed 4 H-bonding with residues Cys145, His163, Glu166 and Arg188, in which three of them are considered as essential amino acids for protease inhibiting function.

Table 1. The docking score of studied compounds on SARS-CoV-2 main protease.

No.	Drug	Binding free energy (kcal/mol)	Ligand efficacy	Interacting amino acids	Usage
1	Remdesivir	-14.12	0.32	Glu166; Thr190	Antiviral
2	Lopinavir	-12.92	0.28	Cys145; His164; Glu166; Gln189	HIV protease inhibitor
3	Darunavir	-12.39	0.33	His41; Cys145; His164; Glu166; Thr190	HIV protease inhibitor
4	Ritonavir	-12.15	0.31	Ser144; Cys145	HIV protease inhibitor
5	Chloroquine	-12.96	0.37	Gly143; His164	Anti-malarial agent
6	Arbidol	-11.14	0.32	Arg188	Antiviral
7	<b>Peptidomimetic inhibitor (N3)</b>	-13.13		Cys145; His163; Glu166; Arg188	

Ligand efficiency (LE) is a useful metric for the selection of lead compounds in drug discovery. It is also a measurement of the binding energy of the ligand per atom, which is calculated according to equation 1. It has been estimated that most of the hits or lead compounds can be considered for further structure optimization given that LE value ranging from 0.25 to 0.6 [34].

$$\Delta g = \frac{\Delta G}{N_{\text{non-hydrogen atoms}}} \quad (1^*)$$

\* $\Delta g$ : Ligand efficiency;  $\Delta G$ : Binding energy.

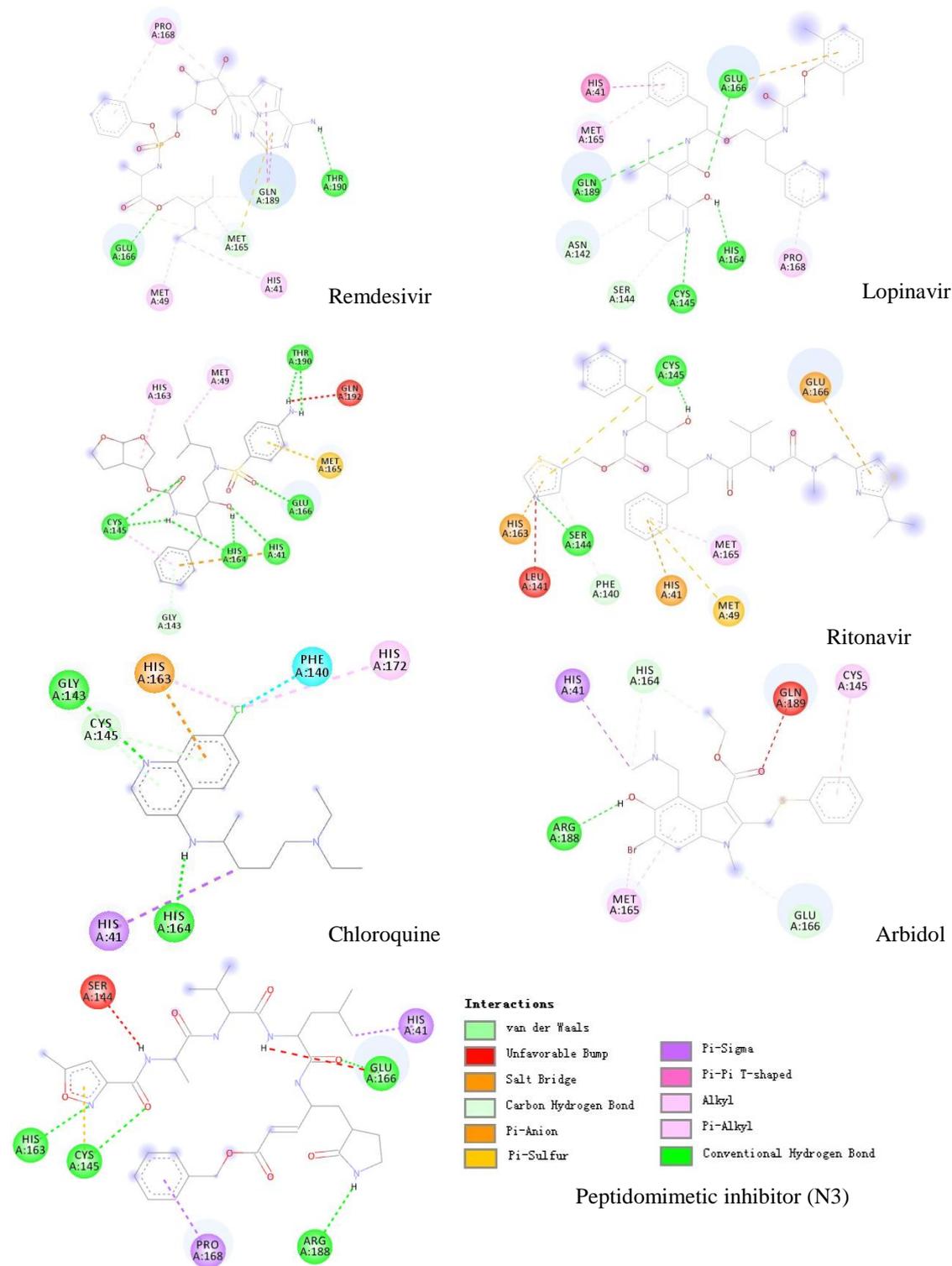


Figure 5. Docking pose of studied medicines to SARS-CoV-2 main protease model.

It is remarkable that remdesivir, an adenosine nucleotide analog, recently utilized in the treatment of COVID-19 and reported as highly effective drug showed the lowest binding free energy (-14.12 kcal/mol). However, the docking conformation of this drug to the active site pocket of protease does not exhibit well fitting. Only 2 hydrogen bonds were formed with amino acids Glu166 and Thr190 and supported by weak interactions with His41, Met49, Met165, Pro168 and Gln189 (Figure 5). This is an interesting result since remdesivir is known to be an inhibitor which focuses on the viral RNA polymerase, thereby causing mistakes in proofreading by viral exoribonuclease. Therefore, this dock pose analysis proves that remdesivir does not specifically interact with main protease target.

Based on the table provided, two HIV protease inhibitor lopinavir-ritonavir showed a minor decrease in binding affinity (-12.92 and -12.15 kcal/mol, respectively). The specificity of two compounds toward targeted protease is demonstrated through docking conformation in Figure 5, Lopinavir formed 4 H-bonding interaction with the main protease including Cys145, His164, Glu166 and Gln189. Meanwhile, ritonavir, known as additional agents that increase plasma Lopinavir concentration, induced only 2 hydrogen bonds with residues Ser144 and Cys145. In addition, the ligand efficacy value of lopinavir and ritonavir were 0.28 and 0.31, respectively which falls within the acceptable range. This result indicates that the structure of these two compounds could be still promising for further optimization in regards to improving bioactivities. Darunavir is another famous HIV-1 protease inhibitor that prevents the formation of mature infectious virus particles. Our results show that darunavir has the potential inhibiting protease with an estimated binding energy in the middle rank (-12.39 kcal/mol) and LE value 0.33. Five H-bonding were created by this drug with His41, Cys145, His164, Glu166 and Thr190 in which three of them are important residues, besides, the interaction was further strengthened by hydrophobic bonds with Met49, Gly143, His163 and Met165. These information suggest that darunavir is a specific inhibitor of SARS-CoV-2 in the main protease.

Chloroquine, a common anti-malarial agent, is effective in COVID-19 treatment as it increases the endosomal pH, an essential factor for virus fusion. The low binding affinity of this drugs (-12.96 kcal/mol) indicates that it does not fit well in the binding site of the protease, thus, the mechanism of this medicine is yet to be investigated. Dock pose analysis further confirms this statement, with chloroquine forming only 2 hydrogen bonds with Gly143 and His164, which are not the essential amino acids in the active site. Arbidol, known as a broad-spectrum anti-viral medicine, showed the least potential to dock with COVID-19 main protease in the active site. The dock score of the drug was -11.14 kcal/mol with only one hydrogen bond created with Arg188 which is located outside the important pocket. This suggests the structure of this candidate is not appropriate to inhibit the protease. On the other hand, the high effectiveness of this medicine in the treatment of new coronavirus-infected patients proves that its mechanism was based on another potential target.

#### 4. CONCLUSIONS

In this study, the chymotrypsin-like cysteine protease (3CL<sup>pro</sup>), known as main protease of COVID-19 has been cross-checked with that of SARS-CoV using sequence alignment method and structure overlay between two models. Obtained results prove that these two models have up to 91.3% of identity and the binding active site or new coronavirus is still conserved, thus, it is considered as potential target for drug development.

With the aim to investigate the inhibition mechanism on SARS-CoV-2 main protease which suggests pathway for studies of potential compounds development, molecular docking

studies have been carried out to investigate the mechanism of action of some medicines currently used in COVID-19 clinical treatment. Obtained results reveal that lopinavir-ritonavir and darunavir are drugs that possess specific structure characteristic that can fit well with the binding site of target protease through interactions with essential amino acids, hence inhibiting the replication function of the virus. These results could be a helpful suggestion for scientists to conduct further research related to structure-based of known drugs in order to develop compounds with desired bioactivity for treatment.

**Acknowledgements.** This research is funded by a basic research project (2020) from Institute of Natural Products Chemistry – Vietnam Academy of Science and Technology.

## REFERENCES

1. Zumla A., Chan J. F. W., Azhar E. I., Hui D. S. C., and Yuen K. Y. - Coronaviruses - drug discovery and therapeutic options, *Nature Reviews Drug Discovery* **15** (5) (2016) 327-347.
2. de Wit E., van Doremalen N., Falzarano D., and Munster V. J. - SARS and MERS: recent insights into emerging coronaviruses, *Nature Reviews Microbiology* **14** (8) (2016) 523-534.
3. <https://www.who.int/westernpacific/emergencies/pneumonia-in-wuhan-china>. Accessed 12th January 2020.
4. [https://www.worldometers.info/coronavirus/?utm\\_campaign=homeAdvegas1](https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1)? COVID-19 CORONAVIRUS PANDEMIC.
5. Lan J., Ge J., Yu J., Shan S., Zhou H., Fan S., Zhang Q., Shi X., Wang Q., Zhang L., and Wang X. - Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, *Nature* **581** (7807) (2020) 215-220.
6. Shang J., Ye G., Shi K., Wan Y., Luo C., Aihara H., Geng Q., Auerbach A., and Li F. - Structural basis of receptor recognition by SARS-CoV-2, *Nature* **581** (7807) (2020) 221-224.
7. Ratia K., Saikatendu K. S., Santarsiero B. D., Barretto N., Baker S. C., Stevens R. C., and Mesecar A. D. - Severe acute respiratory syndrome coronavirus papain-like protease: Structure of a viral deubiquitinating enzyme, *Proceedings of the National Academy of Sciences* **103** (15) (2006) 5717-5722.
8. Chen S., Chen L., Tan J., Chen J., Du L., Sun T., Shen J., Chen K., Jiang H., and Shen X. - Severe Acute Respiratory Syndrome Coronavirus 3C-like Proteinase N Terminus Is Indispensable for Proteolytic Activity but Not for Enzyme Dimerization, *Journal of Biological Chemistry* **280** (1) (2005) 164-173.
9. Chen Y., Savinov S. N., Mielech A. M., Cao T., Baker S. C., and Mesecar A. D. - X-ray Structural and Functional Studies of the Three Tandemly Linked Domains of Non-structural Protein 3 (nsp3) from Murine Hepatitis Virus Reveal Conserved Functions, *Journal of Biological Chemistry* **290** (42) (2015) 25293-25306.
10. Liu X., Zhang B., Jin Z., Yang H., and Rao, Z. - The crystal structure of COVID-19 main protease in complex with an inhibitor N3. To be published. PDB code 6LU7, <https://www.rcsb.org/structure/6LU7>.

11. Colson P., Rolain J. M., Lagier J. C., Brouqui P., and Raoult D. - Chloroquine and hydroxychloroquine as available weapons to fight COVID-19, *International Journal of Antimicrobial Agents* **55** (4) (2020) 105932.
12. Gautret P., Lagier J. C., Parola P., Hoang V. T., Meddeb L., Mailhe M., Doudier B., Courjon J., Giodanengo V., Vieira V. E., Dupont H. T., Honore S., Colson P., Chabriere E., Scola B. L., Rolain J.
13. M., Brouqui P., and Raoult D. - Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial, *International Journal of Antimicrobial Agents* (2020).
14. Wang M., Cao R., Zhang L., Yang X., Liu J., Xu M., Shi Z., Hu Z., Zhong W., and Xiao G. - Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro, *Cell Research* **30** (3) (2020) 269-271.
15. Marshall G. R. - Computer-aided drug design, *Annu Rev Pharmacol Toxicol.* **27** (1987) 193-213.
16. Thuy B. T. P., My T. T. A., Hai N. T. T., Hieu L. T., Hoa T. T., Loan H. T. P., Triet N. T., Anh T. T. V., Quy P. T., Tat P. V., Hue N. V., Quang D. T., Trung N. T., Tung V. T., Huynh L. K., and Nhung N. T. A. - Investigation into SARS-CoV-2 Resistance of Compounds in Garlic Essential Oil, *ACS Omega* **5** (14) (2020) 8312-8320.
17. Peele K.A., Durthi C. P., Srihansa T., Krupanidhi S., Sai A. V., Badu D. J., Indira M., Reddy A. R., and Venkateswarulu T. C. - Molecular docking and dynamic simulations for antiviral compounds against SARS-CoV-2: A computational study, *Informatics in Medicine Unlocked* **19** (2020) 100345.
18. Gupta S., Singh A. K., Kushwaha P. P., Prajapati K. S., Shuaib M., Senapati S., and Kumar S. - Identification of potential natural inhibitors of SARS-CoV2 main protease by molecular docking and simulation studies, *Journal of Biomolecular Structure and Dynamics* (2020) 1-12. <https://doi.org/10.1080/07391102.2020.1776157>.
19. Yu R., Chen L., Lan R., Shen R., and Li P. - Computational screening of antagonists against the SARS-CoV-2 (COVID-19) coronavirus by molecular docking, *International Journal of Antimicrobial Agents* **56** (2) (2020) 106012.
20. Zhang Y., Zheng N., Hao P., Cao Y., and Zhang Y. - A molecular docking model of SARS-CoV S1 protein in complex with its receptor, human ACE2, *Computational Biology and Chemistry* **29** (3) (2005) 254-257.
21. Morris G. M., Huey R., Lindstrom W., Sanner M. F., Belew R. K., Goodsell D. S., and Olson A. J. - AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *Journal of Computational Chemistry* **30** (16) (2009) 2785-2791.
22. Muchtaridi M., Syahidah H. N., Subarnas A., Yusuf M., Bryant S. D., and Langer T. - Molecular Docking and 3D-Pharmacophore Modeling to Study the Interactions of Chalcone Derivatives with Estrogen Receptor Alpha, *Pharmaceuticals* **10** (4) (2017) 81.
23. Deng L., Li C., Zeng Q., Liu X., Li X., Zhang H., Hong Z., and Xia J. - Arbidol combined with LPV/r versus LPV/r alone against Corona Virus Disease 2019: a retrospective cohort study, *Journal of Infection* **81** (1) (2020) e1-e5.
24. Lai C. C., Shih T. P., Ko W. C., Tang H. J., and Hsueh P. R. - Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The

- epidemic and the challenges, *International Journal of Antimicrobial Agents* **55** (3) (2020) 105924.
25. Sahu K. K., Mishra A. K., and Lal A. - Novel coronavirus (2019-nCoV): Update on 3rd Coronavirus Outbreak of 21st Century, *QJM: An International Journal of Medicine*, (2020). <https://doi.org/10.1093/qjmed/hcaa081>.
  26. Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.8, 2015.
  27. <https://chemicalize.com/welcome> Chemicalize.
  28. Allouche A. R. - Gabedit-A graphical user interface for computational chemistry softwares, *Journal of Computational Chemistry* **32** (1) (2011) 174-182.
  29. Tan J., Verchueren K. H. G., Anand K., Shen J., Yang M., Xu Y., Rao Z., Bigalke J., Heisen B., Mesters J. R., Chen K., Shen X., Hualiang J., and Hilgenfeld R. - pH-dependent Conformational Flexibility of the SARS-CoV Main Proteinase (Mpro) Dimer: Molecular Dynamics Simulations and Multiple X-ray Structure Analyses, *Journal of Molecular Biology* **354** (1) (2005) 25-40.
  30. Yang H., Yang M., Ding Y., Liu Y., Lou Z., Zhou Z., Sun L., Mo L., Ye S., Pang H., Gao G. F., Anand K., Bartlam M., Hilgenfeld R., and Rao Z. - The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor, *Proceedings of the National Academy of Sciences* **100** (23) (2003) 13190-13195.
  31. Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., and Ferrin T. E. - UCSF Chimera?A visualization system for exploratory research and analysis, *Journal of Computational Chemistry* **25** (13) (2004) 1605-1612.
  32. Dassault Systèmes BIOVIA, Discovery Studio Visualizer, v17.2.0.16349, San Diego: Dassault Systèmes 2019.
  33. Laskowski R. A., and Swindells M. B. - LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery, *Journal of Chemical Information and Modeling* **51** (10) (2011) 2778-2786.
  34. Maestro S. R. - Schrödinger Release 2019-4: Maestro, Schrödinger, LLC, New York, NY, 2019.
  35. Hopkins A. L., Groom C. R., and Alex A. - Ligand efficiency: a useful metric for lead selection, *Drug Discovery Today* **9** (10) (2004) 430-431.