Vietnam Journal of Science and Technology ***60*** *(6) (2022) 929-947*

doi:10.15625/2525-2518/16459

REVIEW: AUTODOCK 4.2.6 AS AN EFFECTIVE TOOL FOR MOLECULAR DOCKING STUDIES AGAINST SARS-COV-2 MAIN PROTEASE: A TUTORIAL USING MGLTOOLS

Pham Minh Quan[1, 2, *], Le Thi Thuy Huong[1, 2], ,Pham Thi Hong Minh[1, 2], Tran Quoc Toan[1, 2], Do Tien Lam[1, 2], Vu Thi Thu Le[3], Pham Quoc Long[1, 2, *]

[1]*Institute of Natural Products Chemistry, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Ha Noi, Viet Nam*

[2]*Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Ha Noi, Viet Nam*

[3]*Thai Nguyen University of Agriculture and Forestry, Quyet Thang, Thai Nguyen City, Viet Nam*

[*]Email: *minhquanaries@gmail.com; mar.biochem@fpt.vn*

Received: 5 August 2021; Accepted for Publication: 26 November 2021

Abstract. The worldwide pandemic caused by coronavirus SARS-CoV-2 (so called as COVID-19 disease) has affected 219 countries and territories, leading to numerous deaths and global financial crisis. The main protease (Mpro) of SARS-CoV-2 plays an important role in mediating the transcription and replication of virus, thus, one of the main therapeutic measures is to find compounds capable of inhibiting these enzymes as soon as possible. Nowadays, computer-aided drug design plays an important role in the field of drug discovery. In particular, molecular docking is one of the initial steps that effectively screen a numerous number of compounds for their interaction and binding affinity toward targeted enzyme, therefrom, suggesting a short list of potential inhibitors for further drug development processes. As part of our ongoing program to provide simple guideline for scientific community to utilize different docking tools for research purposes, in this article, a complete manual guideline of Autodock 4.2.6 is presented to demonstrate the simulation of interaction between compound PF-07321332 and the main protease of SARS-CoV-2, thereby suggesting an effective tool for scientists to conduct reseach on this disease.

Keywords: molecular docking, autodock 4.2.6, autodock tools, drug design, COVID-19

Classification numbers: 4.8.5, 4.10.4, 1.2.1, 1.2.4.

1. INTRODUCTION

Nowadays, the whole world is facing the most devastating and dangerous pandemic in history namely COVID-19. Detected for the first time since December 2019 in Wuhan city, Hubei province of China, SARS-CoV-2 is a new strain of corona virus that has never been recognized before. Previously, corora virus strains were also known for causing many dangerous pandemics such as Severe Acute Respiratory Syndrome (SARS-CoV) in 2002 or Middle East respiratory syndrome (MERS-CoV) in 2012. Until now, the number of confirmed COVID-19 infections worldwide are 203.533.164 cases, and the number of deaths reached 4.310.066 as of August 8th, 2021 [1].

SARS-CoV-2 main protease (Mpro) is one of the most important proteins 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 [2]. Thus, the search for a potent compound that inhibits the function of Mpro is considered as one of the novel therapies for treatment. Additionally, the structure of the SARS-CoV-2 main protease in complex with a peptidomimetic inhibitor (PDB ID: 6LU7) was reported with up to 82 % identity to that of SARS-CoV [3, 4].

Nowadays, computer-aided drug design (CADD) is increasingly utilized in the field of drug discovery and development. The main advantage of this method is known to help scientists promptly identify compound with potential biological activities, thus, save time and cost for drug development process [5, 6]. In CADD, molecular docking is known as an effective method which simulates the binding orientation of a molecule with an enzyme and calculating their binding affinity. It has been indicated as a vigorous method for predicting potential compounds with expected inhibition activities toward targeted enzyme, thus, suggesting candidates for further drug development studies [7 - 12]. Amongst various docking softwares, AutoDock 4 is an excellent non-commercial docking program that is widely used with more than 10,000 citations during the last ten years [13 - 15]. Despite the fact that the lastest version (AutoDock 4.2.6) was released in 2014, the advantage of its algorithm and scoring function keeps resulting in more accurate identification of potential compounds compared with other docking tools in terms of binding free energy [15]. Up to date, there have been several tutorials on using Autodock 4 for virtual screening purpose previously, however, it is still complicated and requires the use of various softwares [16, 17]. Also, due to the urgent need of developing an effective drug for the treatment of COVID-19 epidemic, a simple protocol which can be easily manipulated by the scientific community is needed in order to accelerate the drug discovery process. In this study, a step by step protocol for the use of Autodock 4.2.6 (AD4) software was given to simulate the interaction between Mpro of SARS-CoV-2 and PF-07321332. The obtained results might help to shed light on the interaction mechanism of the compound and hopefully help scientists to effectively implement virtual screening study on this disease.

2. METHODS

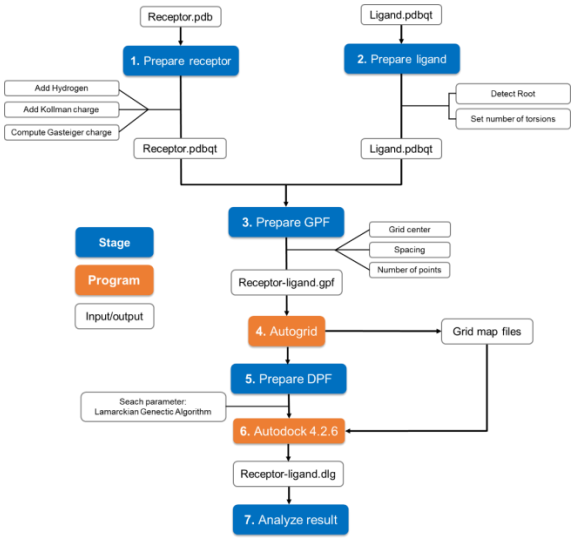


Figure 1. AutoDock 4.2.6 workflow using MGLTools.

A summary of AutoDock 4.2.6 workflow is presented below (Figure 1), the details of the docking process are discussed in the next section.

3. EXPERIMENTAL AND RESULTS

3.1. Requirements

3.1.1 Discovery Studio Visualizer

https://discover.3ds.com/discovery-studio-visualizer-download

3.1.2. AutoDock 4.2.6

Download two files autodock4.exe and autogrid4.exe to the working directory (for example in this study: C:\Users\Admin\Desktop\6lu7)

http://autodock.scripps.edu/downloads/autodock-registration/autodock-4-2-download-page/

3.1.3. MGLTools

http://mgltools.scripps.edu/downloads

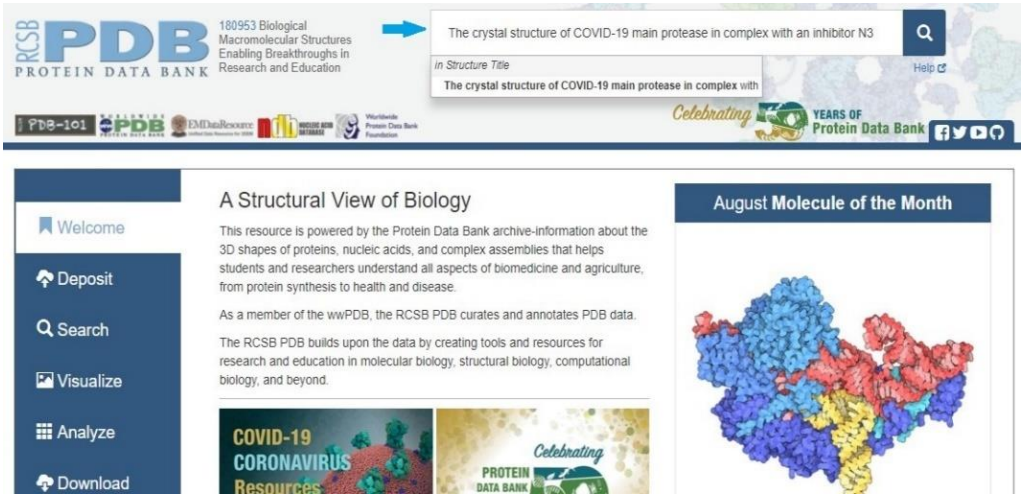
3.1.4. Java

http://www.java.com/en/download/index.jsp

3.2. Prepare input files for docking simulation

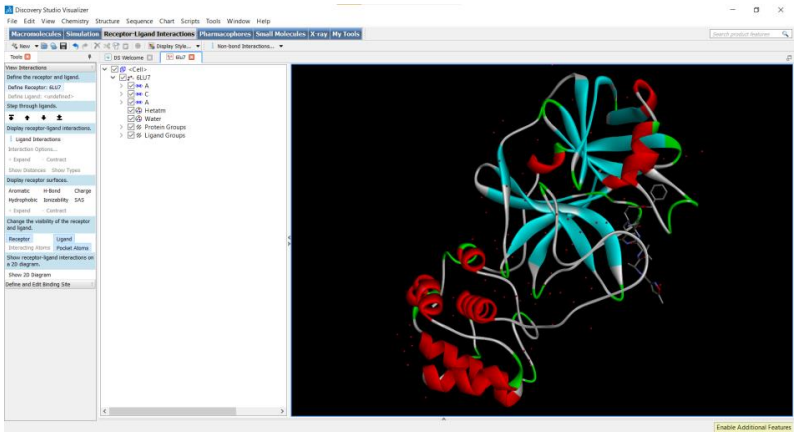
3.2.1. Retrieving protein.pdb file from protein database

- The three-dimensional crytal structure of SARS-CoV-2 Mpro (PDB ID: 6LU7) is downloaded from Protein Data Bank: https://www.rcsb.org/
  - Type the query protein or enzyme (The crystal structure of COVID-19 main protease in complex with an inhibitor N3)
  - Select protein (PDB ID: 6LU7)

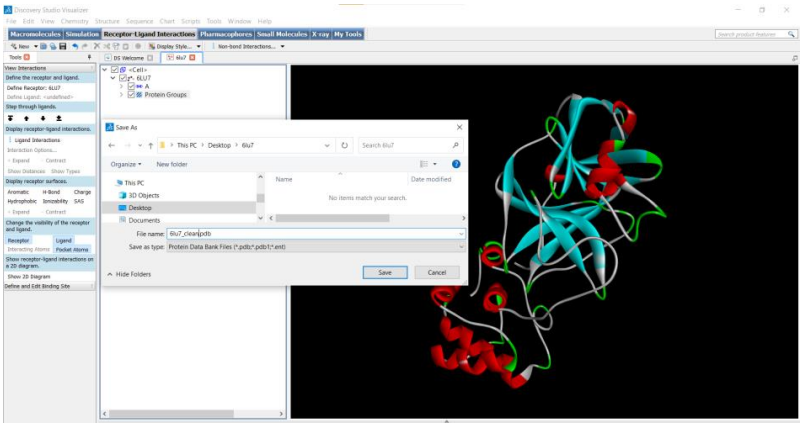


  - Select Download files
  - Click PDB format and download it

- Prepare input data for protein by Discovery studio:
  - Open 6LU7.pdb in Discovery Studio Visualiser
  - Press ctrl + H
  - Select Hetatm, right click and select remove group
  - Select Water and Delete
  - Select Ligand Group and Delete

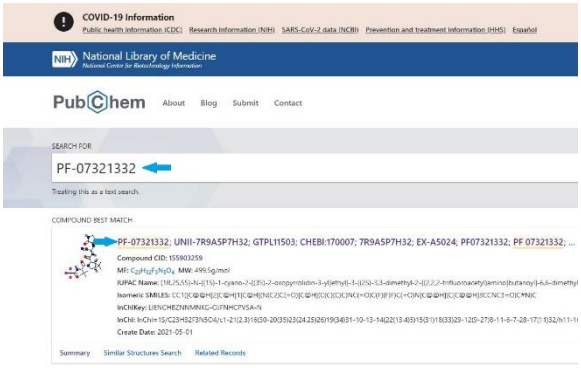


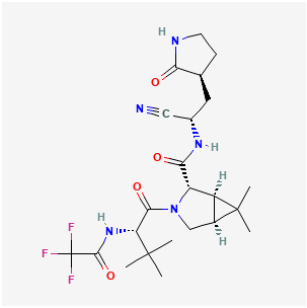
  - Save as 6lu7_clean.pdb



3.2.2. Retrieving ligand.pdb file from major ligand database

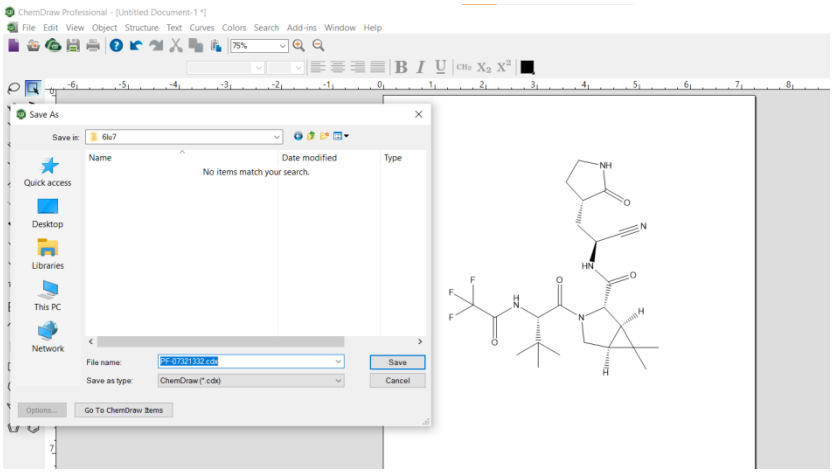
- Obtain ligand structure from Pubchem database (https://pubchem.ncbi.nlm.nih.gov/). In this study, we choose PF-07321332 as reference ligand since it was proved as drug candidate with SARS-CoV-2 Mpro inhibition activity [18].



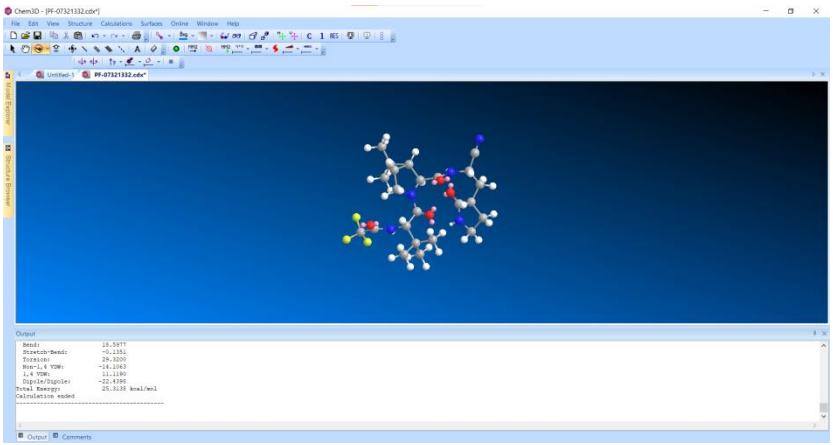


PF-07321332

- Create ligand.pdb file (PF-07321332.pdb)
  - Prepare PF-07321332.cdx file using Chemdraw

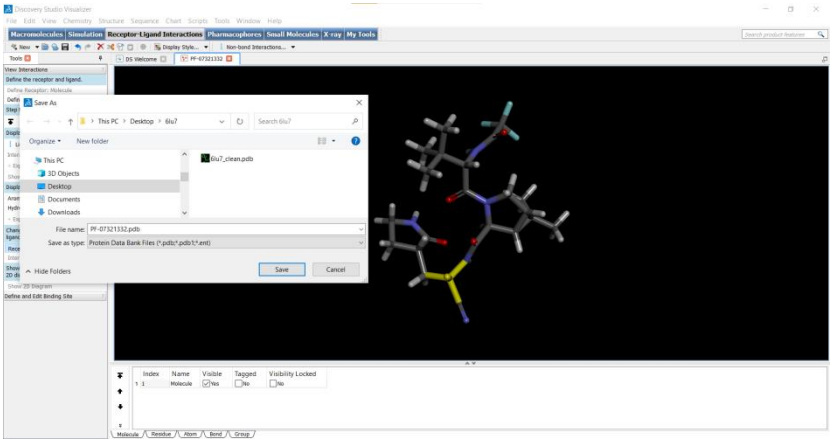


  - Open PF-07321332.cdx file in Chem3D > Ctrl+M (energy minimization the three-dimensional structure of ligand using MM2 algorithm)



  - Save as PF-07321332.mol2 file

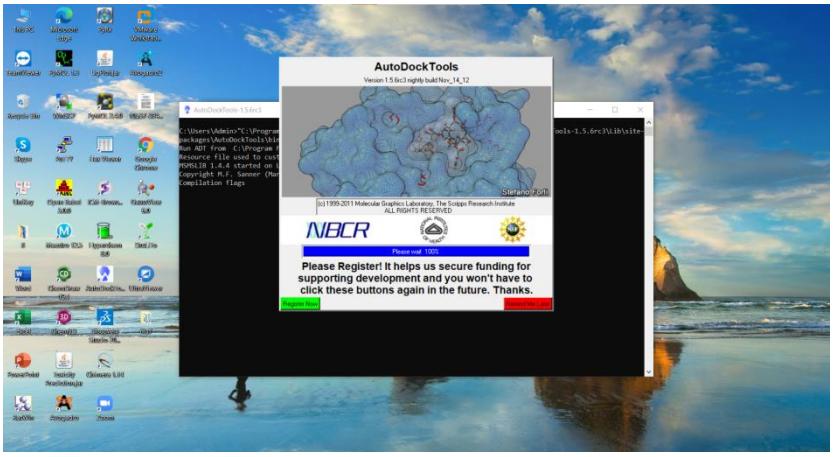
- Open PF-07321332.mol2 in Discovery Studio Visualiser
- Click File to Save as PF-07321332.pdb



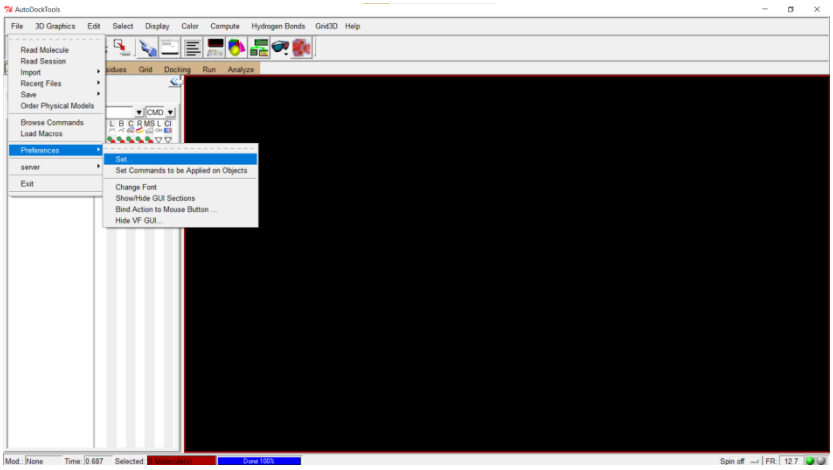
3.3. Molecular docking using AutoDock 4.2.6

3.3.1. Create protein.pdbqt file

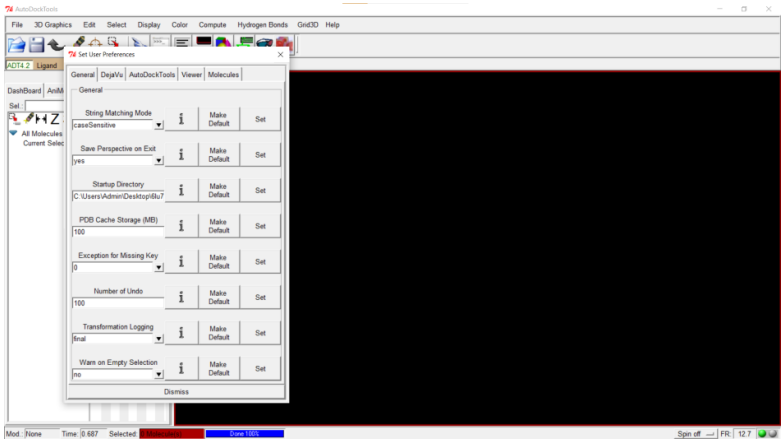
- Open AutoDockTools



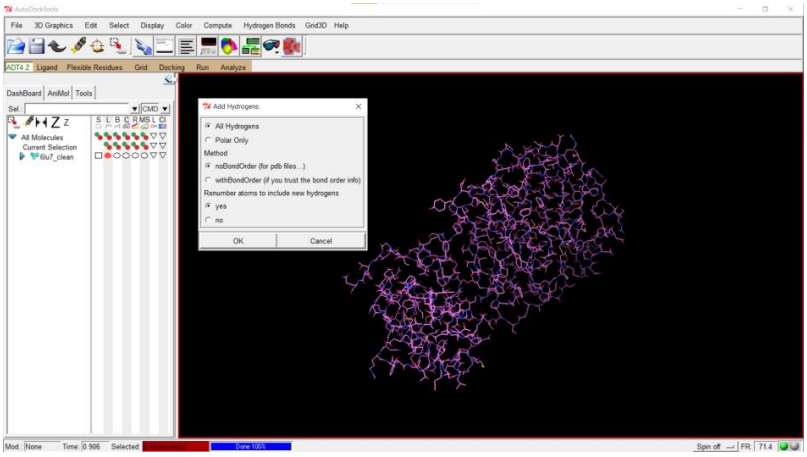
- From the File menu > Preferences > Set



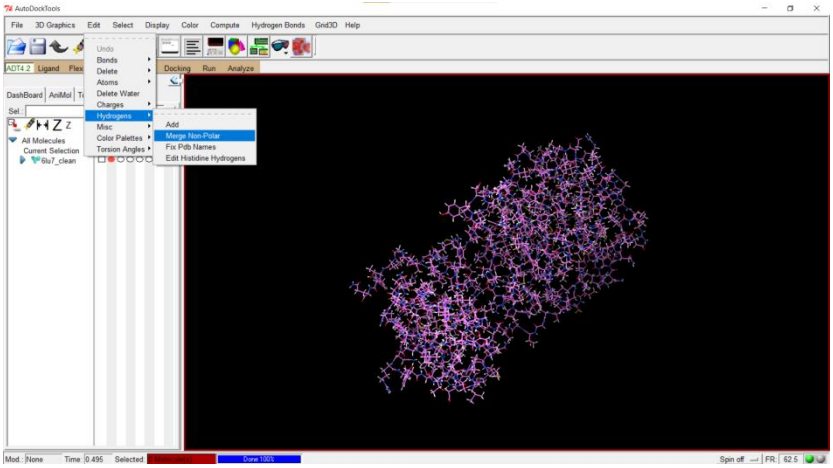
- Direct to the working directory in the Startup Directory (C:\Users\Admin\Desktop\6lu7) and click Set > Dismiss



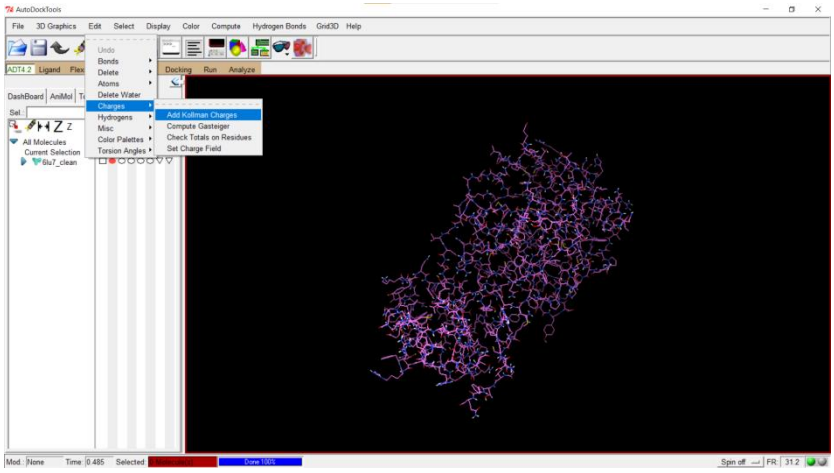
- Open file > Read Molecule > Select and open 6lu7_clean.pdb
- Click on Edit > Hydrogens > Add
- Choose All Hydrogens > noBondOrder (for pdb files…) > Yes > Ok



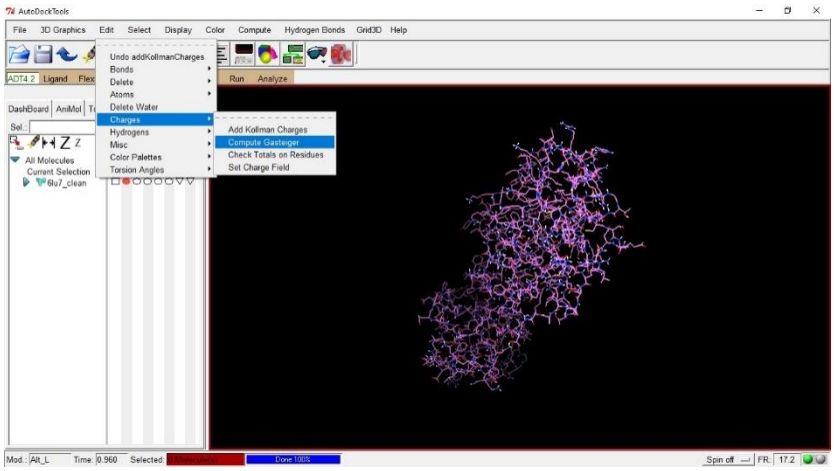
- Click on Edit > Hydrogens > Merge Non-polar



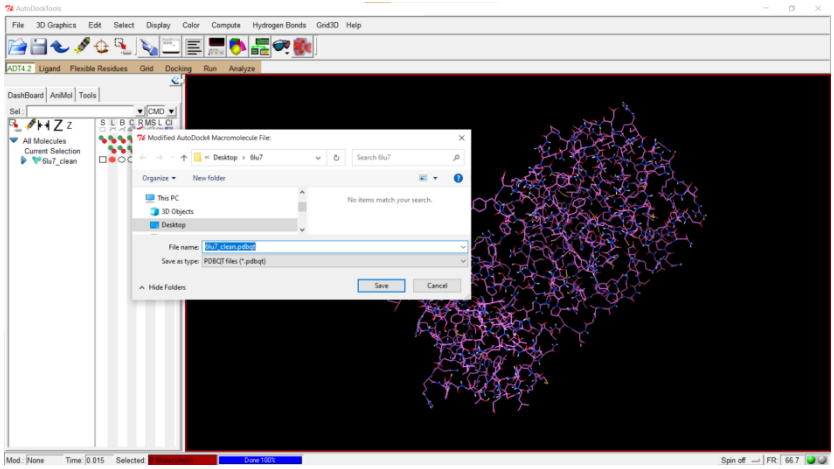
- Click again Edit > Charges > Add Kollman Charges



- Click again Edit > Charges > Compute Gasteiger

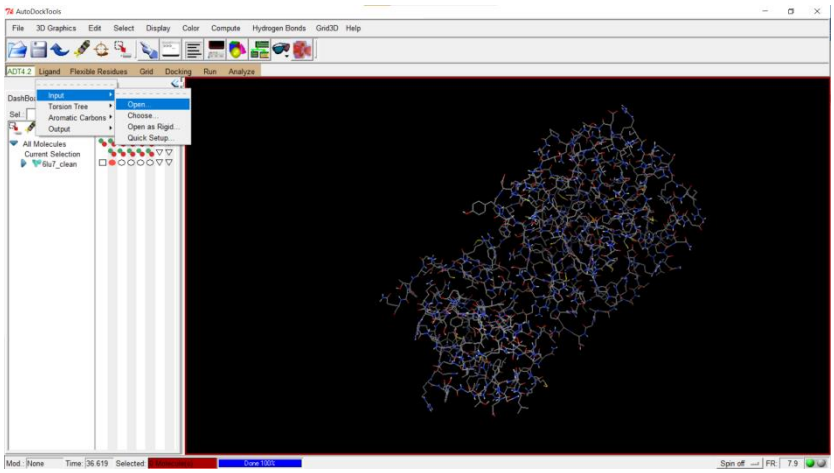


- Click Grid > Macromolecules > Choose > 6lu7_clean > Select Molecule > Ok
- Save 6lu7_clean.pdbqt in the working directory C:\Users\Admin\Desktop\6lu7

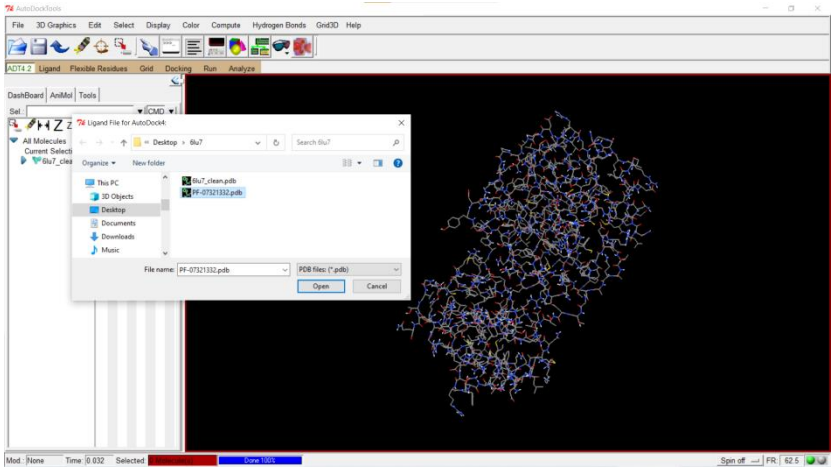


3.3.2. Create ligand.pdbqt file

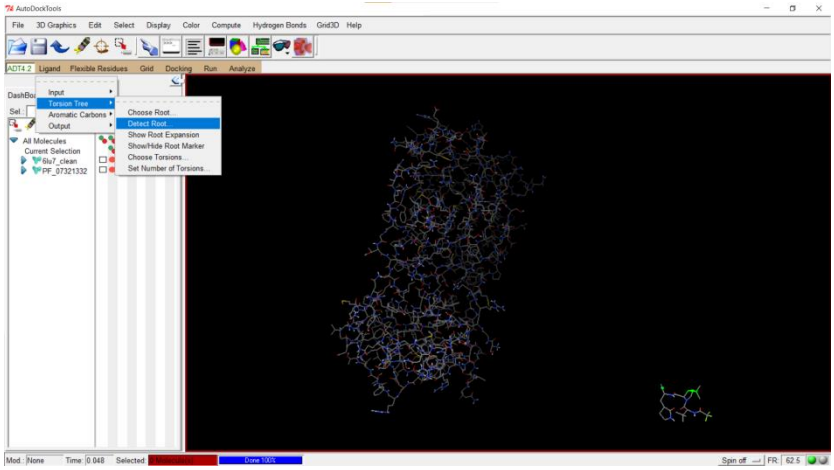
- Select Ligand > Input > Open



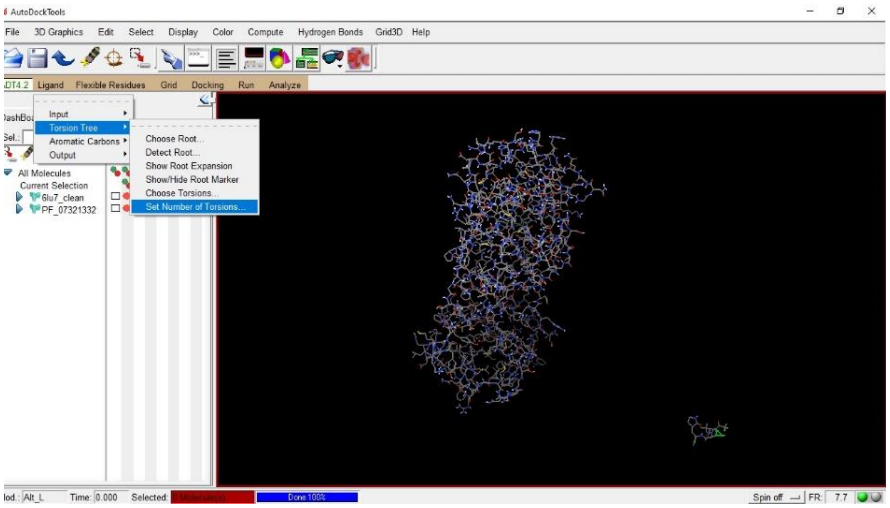
- Choose PF-07321332.pdb > Open



- Click again Ligand > Torsion Tree > Detect Root



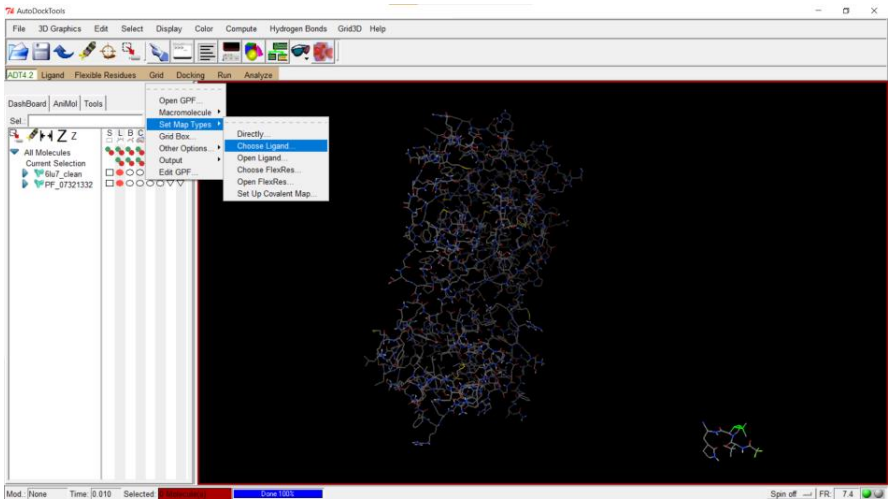
- Click again Ligand > Torsion Tree > Set number of Torsions



- Again Ligand > Output > Save as PDBQT (PF-07321332.pdbqt) in the working directory C:\Users\Admin\Desktop\6lu7

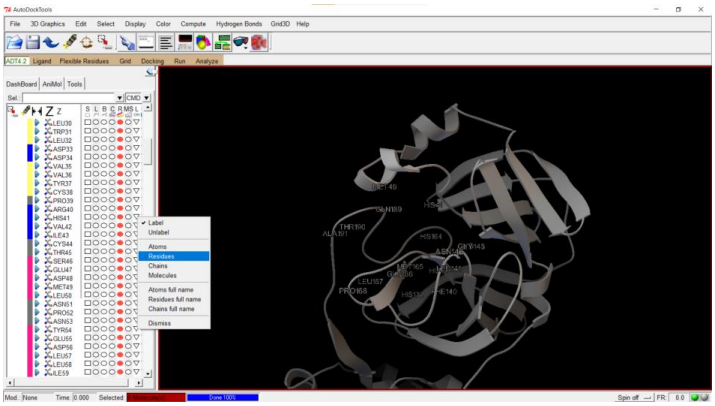
3.3.3. Create Grid parameter file (box.gpf)

- Open Grid >Set Map Types > Choose Ligand > click PF-07321332 > Select Ligand

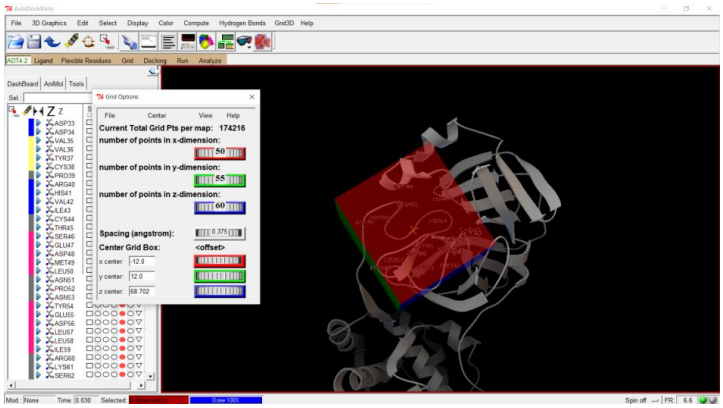


Center the x, y, z atomic coordinates: Normally, the grid box is centered according to the reference inhibitor, however, in other cases, it could be identified according to the active site of protein. In this stuty, we selected coordinates of grid box based on the known ligand binding active site of SARS-CoV-2 main protease. As previous studies have shown, the active region of SARS is composed of important residues: His 41, Met 49, Phe 140, Leu 141, Asn 142, Gly 143, Cys 145, His 163, His 164, Met 165, Glu 166, Lru 167, Pro 168, His 172, Gln 189, Thr 190, and Ala 191 [19].

- Show the important residues to create the box: choose>▽right click > Residues



- Click Grid > Grid Box



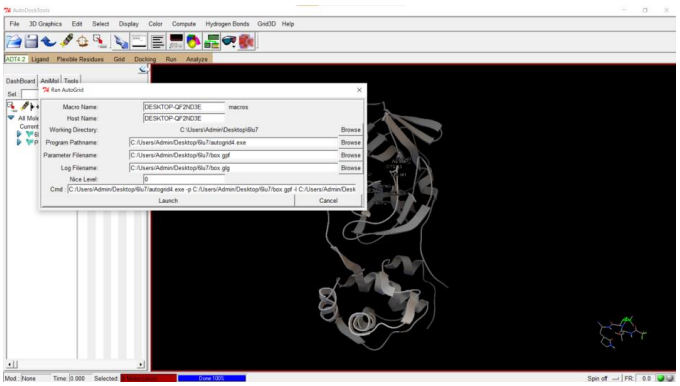
Here we choose the size and coordinates so that the box covers all the essential residues. The grid point was set as 50 × 55 × 60 with x, y, z coordinates of -12.000; 12.000; 68.702.

- Click File > Close saving current
- Click Grid again > Output > Save GPF (name as box.gpf) in the working directory C:\Users\Admin\Desktop\6lu7

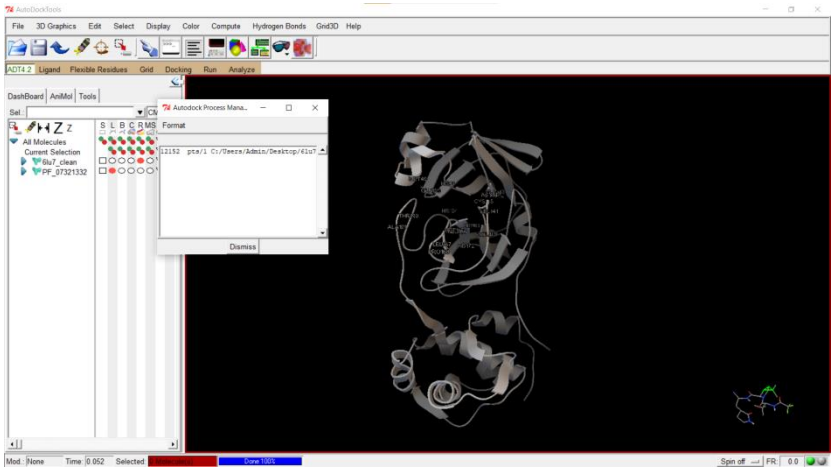
3.3.4. Run AutoGrid4

- Click Run > Run AutoGrid
- Choose link as shown below and click Launch

(*Copy file autogrid4.exe on directory 6lu7 before running AutoGrid4)

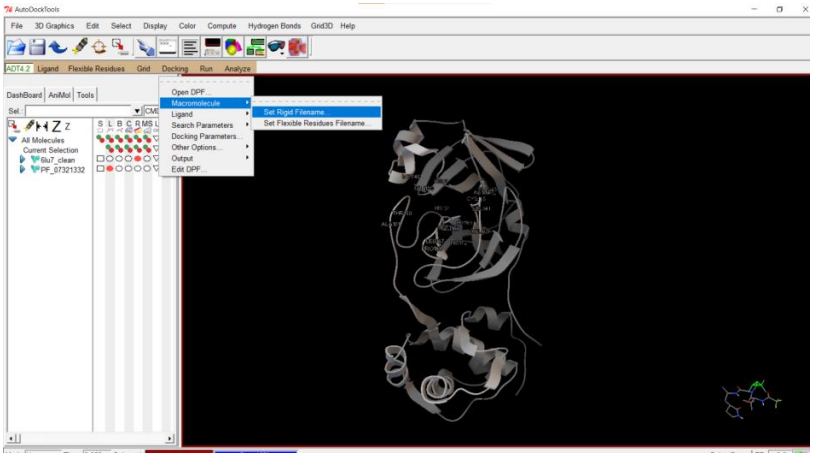


- When running successfully, a job tab will appear on the screen

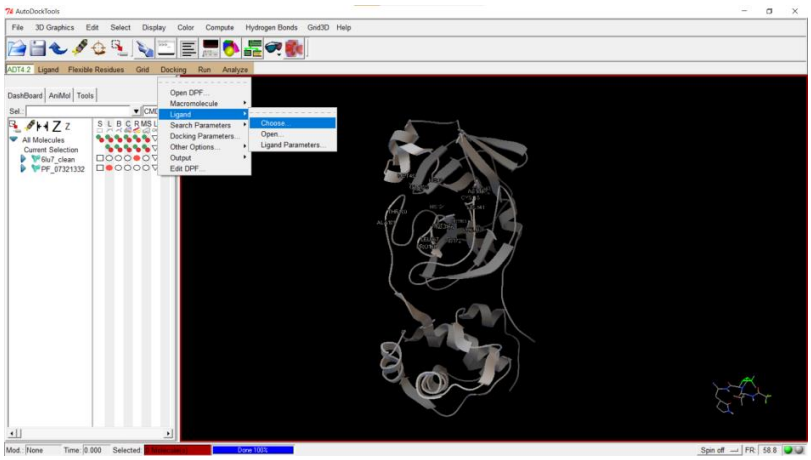


3.3.5. Run Autodock 4

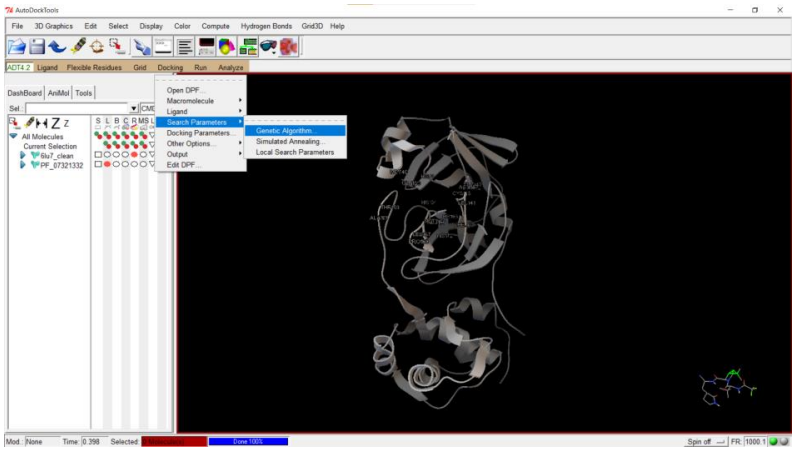
- Click Docking > Macromolecule > Set Rigid Filename > click 6lu7_clean.pdbqt



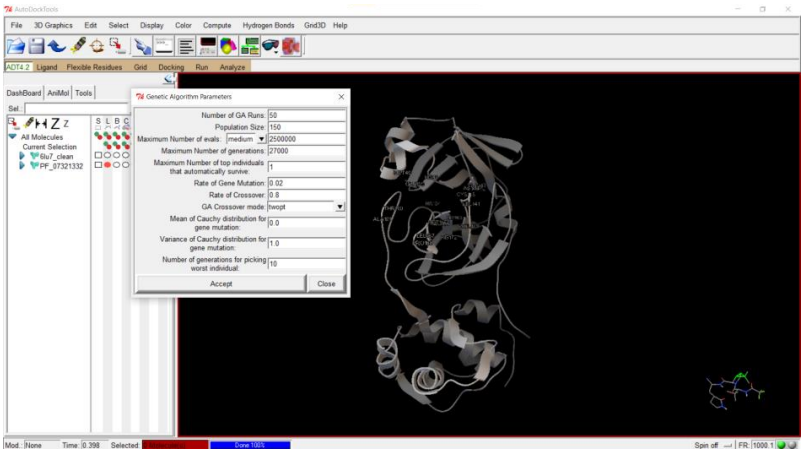
- Click Docking > Ligand > Choose > click PF-07321332 > Select Ligand > accept



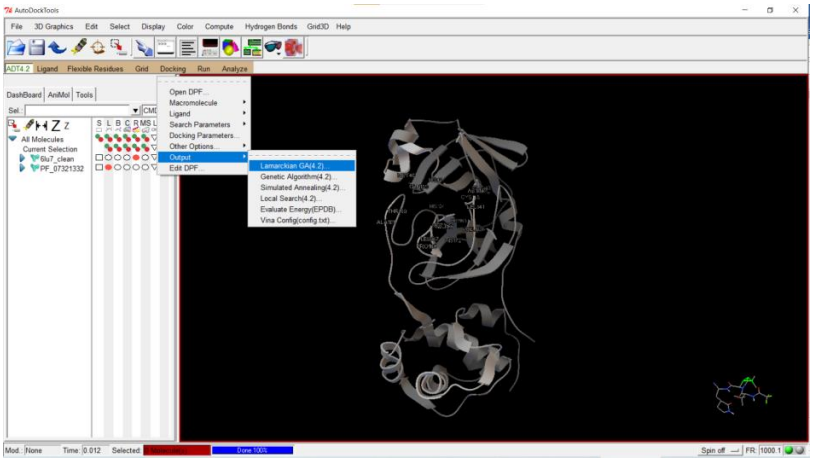
- Click Docking > Search Parameters > Genetic Algorithm



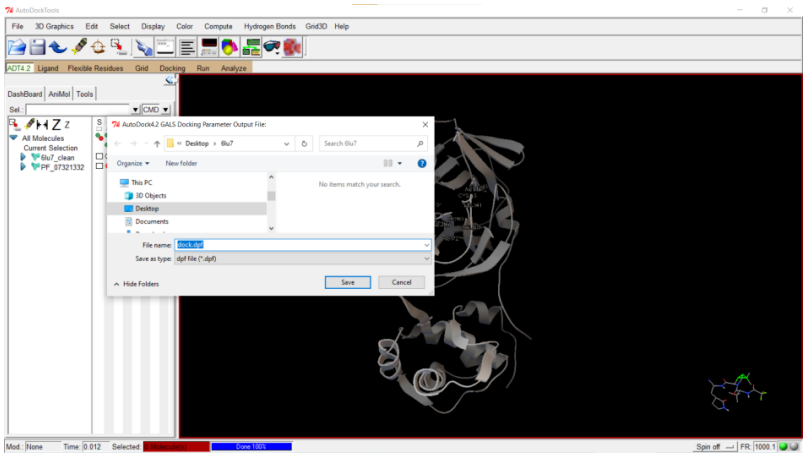
- Set up as shown and choose Accept



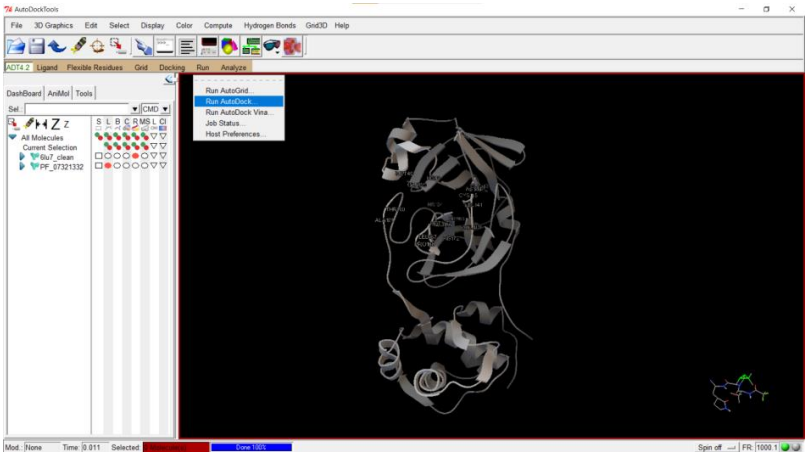
- Click Docking again > Output > Lamarckial (GA 4.2)



- Save DPF file (name as dock.dpf) in the working directory C:\Users\Admin\Desktop\6lu7

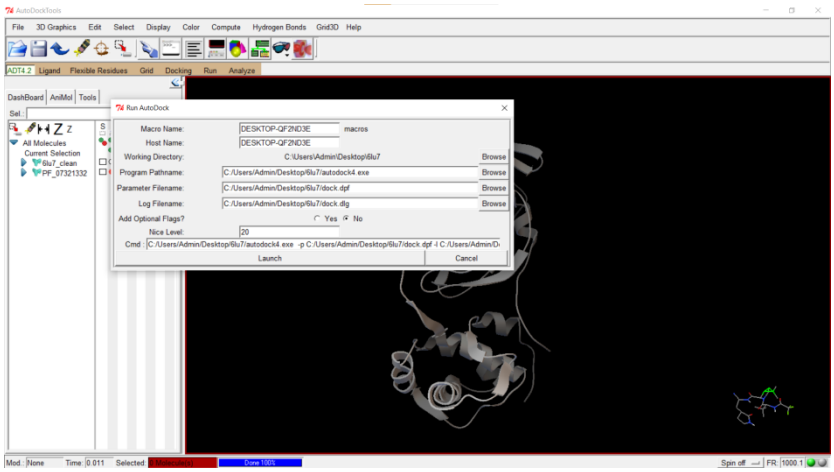


➢ Click Run > Run AutoDock

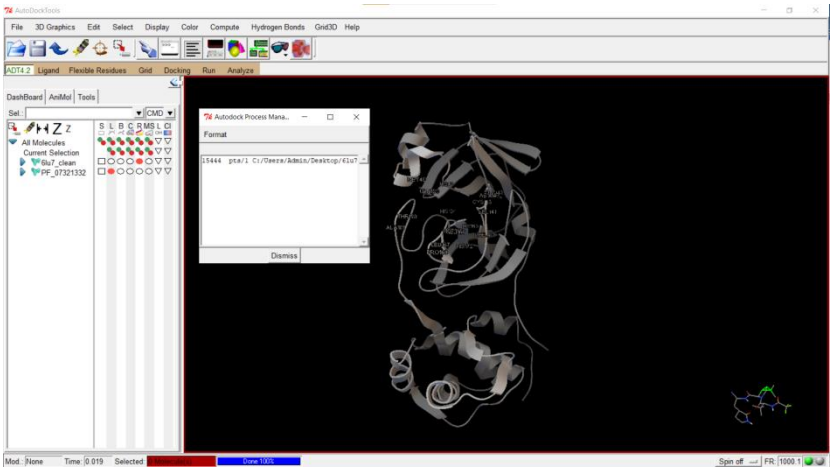


➢ Choose link as shown below and click Launch

(*Copy file autodock4.exe on directory 6lu7 before running AutoDock4)



➢ When running successfully, a job tab will appear on the screen



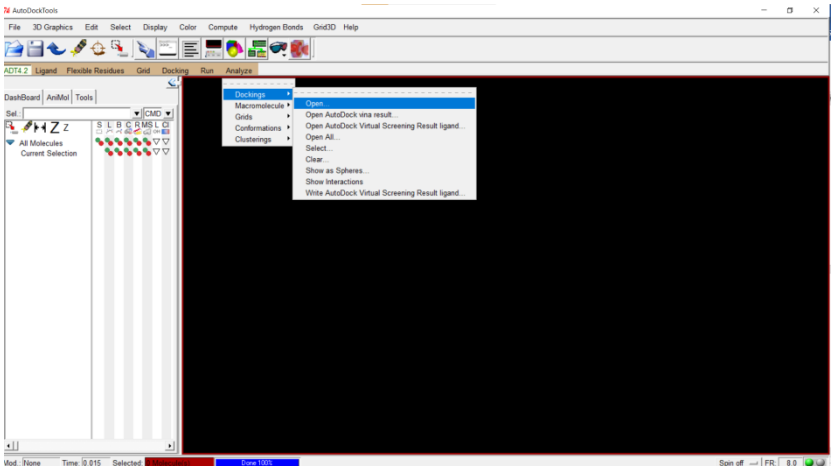
3.3.6. Analyzing result

- Check the energy rankings of the dock poses
  - Read the output result file (dock.dlg), we can see the list of binding free energy of different dock poses as shown below.

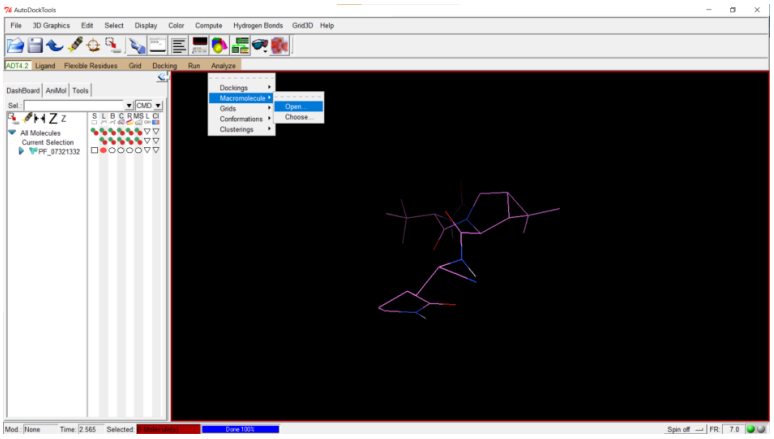
```
CLUSTERING HISTOGRAM
____________________
```

Clus-ter Rank	Lowest Binding Energy	Run	Mean Binding Energy	Num in Clus	Histogram 5 10 15 20 25 30 35
1	-9.24	47	-8.54	4	####
2	-9.11	2	-8.69	2	##
3	-8.84	35	-8.75	2	##
4	-8.69	44	-8.12	3	###
5	-8.62	7	-7.92	5	#####
6	-8.25	38	-7.89	2	##
7	-8.18	23	-8.18	1	#
8	-8.11	20	-7.73	2	##

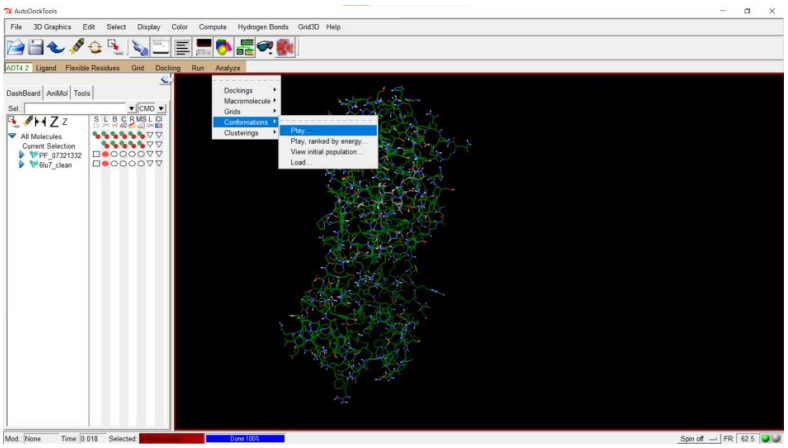
  - Open AutoDockTools > click Analyze > Docking > Open > Select dock.dlg > Ok



  - Again Analyze > Macromolecule > Open

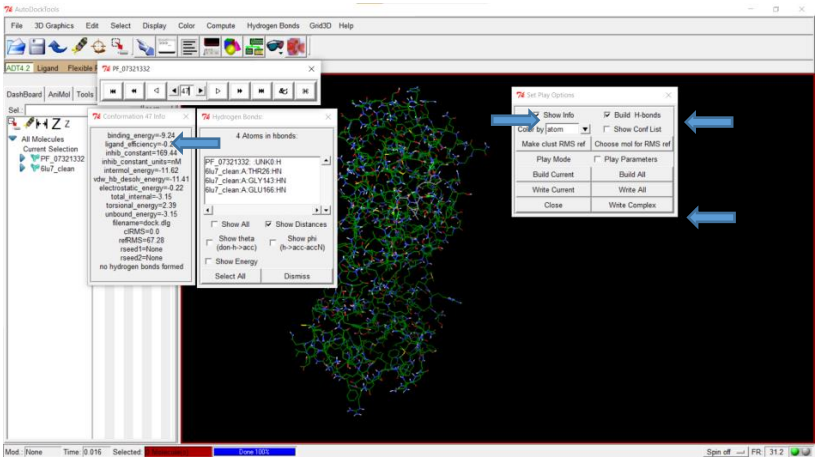


- Again Analyze > Comformations > Play



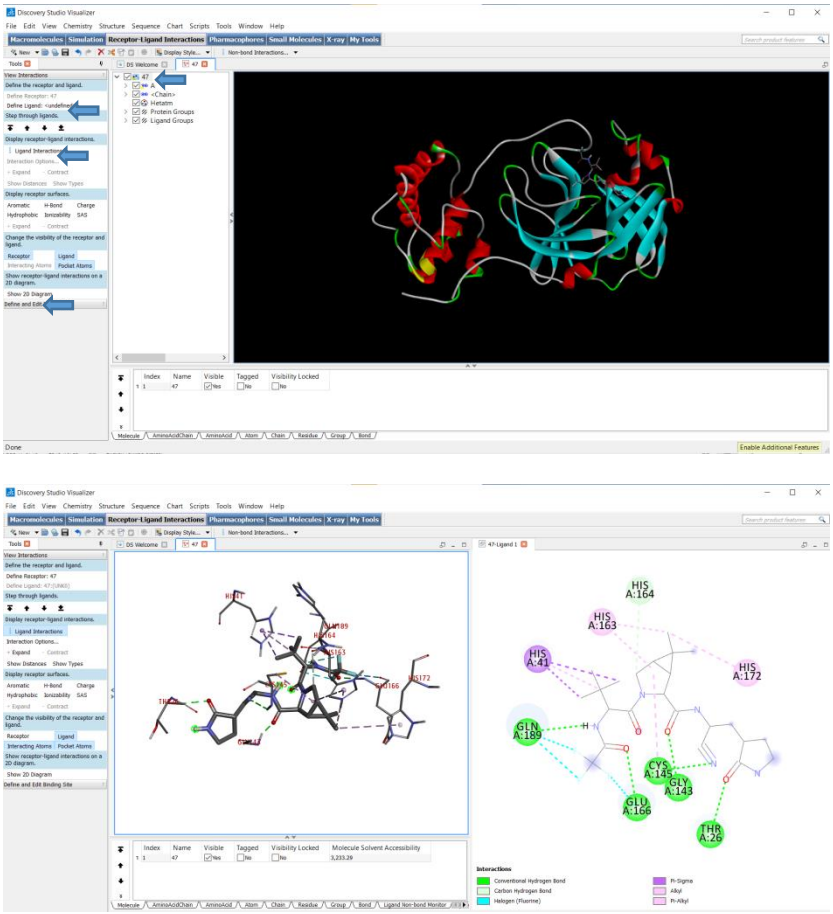
- Click this sign  > click on Show Info and Build H-bonds
- Click Write Complex and save name Dock_47.pdbqt

  (*In our case conformation number 47 was indicated as the best dock pose with binding free energy ΔG = -9.24 kcal/mol)



- Check the constitutive bonds of the cases by Discovery studio Visualiser

- Open file dock_47.pdbqt in Discovery Studio Visualiser
- Click dock47 > Define Ligand > Ligand Interactions > Show 2G Diagram



The results showed the lowest binding free energy of PF-07321332 (-9.24 kcal/mol). Four H-bondings were created by this compound with Thr 26, Gly 143, Cys 145, Glu 166 and Gln189, four of them are essential residues in the active binding region of SARS-CoV-2 Mpro. Besides, the interaction was further strengthened by hydrophobic bonds with His 163 and His 172. In general, these information indicate that PF-07321332 could be a specific inhibitor of SARS-CoV-2 main protease. The obtained results was analyzed in more detail by comparing with two other known SARS-CoV-2 main protease inhibitors (lopinavir and darunavir) using the same docking protocol (Table 1).

Table 1. Docking results of PF-07321332, Lopinavir and Darunavir using AutoDock 4.2.6.

No.	Compound	Binding free energy (kcal/mol)	H-bond interacting residues
1	PF-07321332	-9.24	Thr26, Gly143, Cys145, Glu166, Gln189
2	lopinavir	-8.95	Cys145, His164, Glu166, Gln189
3	darunavir	-8.62	His41, Cys145, His164, Glu166, Thr190

According to the ranking criteria of AutoDock, the more negative the value of docking score, the better the binding affinity of the compound towards targeted receptor. Binding free energy values in Table 1 indicate that PF-07321332 binds toward SARS-CoV-2 main protease with the highest binding affinity, thus explaining why this compound is being considered as the most potential drug candidate at the current state. Dock pose analysis of lopinavir and darunavir also exhibits that they formed hydrogen bond interaction with common essential residues in the active site of targeted protein which suggests the liability of the docking procedure.

4. CONCLUSIONS

This article provided a simple and user-friendly manual of Autodock 4.2.6 to demonstrate the simulation of interaction between compound PF-07321332 and SARS-CoV-2 Mpro (PDB ID: 6LU7). The obtained results show that PF-07321332 is a potential drug capable of binding to essential residues of the active site of targeted protein, hence potentially inhibiting the replication function of the virus.

CRediT authorship contribution statement. Pham Minh Quan, Le Thi Thuy Huong: Investigation, methodology and editing manuscript. Pham Thi Hong Minh, Tran Quoc Toan, Do Tien Lam, Vu Thi Thu Le: analyze data and drafting manuscript. Pham Quoc Long : supervision and revising manuscript.

Declaration of competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. https://www.worldometers.info/coronavirus/.
2. Zhou P., Yang X.L., Wang X. G., Hu B., Zhang L., Zhang W., Si H. R., Zhu Y., Li B., Huang C. L., Chen H. D., Chen J., Luo Y., Guo H., Jiang R. D., Liu M. Q., Chen Y., Shen X. R., Wang X., Zheng X. S., Zhao K., Chen Q. J., Deng F., Liu L. L., Yan B., Zhan F. X., Wang Y. Y., Xiao G. F., and Shi Z. L. - A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature **579** (7798) (2020) 270-273. https://doi.org/10.1038/s41586-020-2012-7.
3. Schoeman D. and Fielding B. C. - Coronavirus envelope protein: current knowledge, Virol J. **16** (1) (2019). https://doi.org/10.1186/s12985-019-1182-0.
4. Fauquet C. M. and Fargette D. - International Committee on Taxonomy of Viruses and the 3,142 unassigned species, Virol J. **2** (1) (2005). https://doi.org/10.1186/1743-422x-2-64.
5. Yu W. and MacKerell A. D. - Computer-Aided Drug Design Methods, In: Antibiotics, Chapter Chapter 5, 2017, pp. 85-106.
6. Marshall G. R. - Computer-Aided Drug Design, Annu. Rev. Pharmacol Toxicol **27** (1) (1987) 193-213. https://doi.org/10.1146/annurev.pa.27.040187.001205.
7. Pham M. Q., Tran T. H. V., Pham Q. L., and Gairin J. E. - In silico analysis of the binding properties of solasonine to mortalin and p53, and in vitro pharmacological studies of its apoptotic and cytotoxic effects on human HepG2 and Hep3b hepatocellular carcinoma cells, Fundam. Clin. Pharmacol **33** (4) (2019) 385-396. https://doi.org/10.1111/fcp.12447.

8. Tao X., Huang Y., Wang C., Chen F., Yang L., Ling L., Che Z. and Chen X. - Recent developments in molecular docking technology applied in food science: a review, Int. J. Food Sci. Technol **55** (1) (2019) 33-45. https://doi.org/10.1111/ijfs.14325.
9. Lavecchia A. and Giovanni C. - Virtual Screening Strategies in Drug Discovery: A Critical Review, Curr. Med. Chem. **20** (23) (2013) 2839-2860. https://doi.org/10.2174/09298673113209990001.
10. Lill M. Virtual Screening in Drug Design, In: In Silico Models for Drug Discovery, Chapter Chapter 1, 2013, pp. 1-12.
11. Maia E. H. B., Assis L. C., de Oliveira T. A., da Silva A. M. and Taranto A. G. - Structure-Based Virtual Screening: From Classical to Artificial Intelligence, Frontiers in Chemistry **8** (2020). https://doi.org/10.3389/fchem.2020.00343.
12. Kontoyianni M. Docking and Virtual Screening in Drug Discovery, In: Proteomics for Drug Discovery, Chapter Chapter 18, 2017, pp. 255-266.
13. Dan N. T., Quang H. D., Van Truong V., Huu Nghi D., Cuong N. M., Cuong T. D., Toan T. Q., Bach L. G., Anh N. H. T., Mai N. T., Lan N. T., Van Chinh L., and Quan P. M. - Design, synthesis, structure, in vitro cytotoxic activity evaluation and docking studies on target enzyme GSK-3β of new indirubin-3'-oxime derivatives, Sci. Rep. **10** (1) (2020). https://doi.org/10.1038/s41598-020-68134-8.
14. Ngo S. T., Tam N. M., Pham M. Q. and Nguyen T. H. - Benchmark of Popular Free Energy Approaches Revealing the Inhibitors Binding to SARS-CoV-2 Mpro, J. Chem. Inf. Model **61** (5) (2021) 2302-2312. https://doi.org/10.1021/acs.jcim.1c00159.
15. Nguyen N. T., Nguyen T. H., Pham T. N. H., Huy N. T., Bay M. V., Pham M. Q., Nam P. C., Vu V. V., and Ngo S. T. - Autodock Vina Adopts More Accurate Binding Poses but Autodock4 Forms Better Binding Affinity, J. Chem. Inf. Model. **60** (1) (2019) 204-211. https://doi.org/10.1021/acs.jcim.9b00778.
16. Mario S. V. T., Mario E. V. T., Pedro A. V., and Ernesto M. - AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4, Biol. Direct. **15** (12) (2020) 1-12. https://doi.org/10.1186/s13062-020-00267-2.
17. Syed M. D. R., Shazi S. and Mohd H. - A simple click by click protocol to perform docking: AutoDock 4.2 made easy for non-bioinformaticians, EXCLI J. **12** (2013) 831-857. https://doi.org/10.17877/DE290R-11534.
18. Jonsson C. B., Golden J. E. and Meibohm B. - Time to 'Mind the Gap' in novel small molecule drug discovery for direct-acting antivirals for SARS-CoV-2, Curr. Opin. Virol. 50 (2021) 1-7. https://doi.org/10.1016/j.coviro.2021.06.008.
19. Jin Z., Du X., Xu Y., Deng Y., Liu M., Zhao Y., Zhang B., Li X., Zhang L., Peng C., Duan Y., Yu J., Wang L., Yang K., Liu F., Jiang R., Yang X., You T., Liu X., Yang X., Bai F., Liu H., Liu X., Guddat L. W., Xu W., Xiao G., Qin C., Shi Z., Jiang H., Rao Z., and Yang H. - Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors, Nature **582** (7811) (2020) 289-293. https://doi.org/10.1038/s41586-020-2223-y.