Binding of Hispidol A to C3-like Protease of SARS-CoV-2 (COVID-19)

Nusratun Nahar[1*], Tohmina Afroze Bondhon[2],

Abstract

COVID-19 viral infection is caused by a corona virus (SARS-CoV-2) and which has created a global pandemic. As of July 5, 2020, the disease has infected 11,425,130 persons and caused deaths of 534,487 individuals throughout most of the countries of the world. The statistics for Bangladesh are 162,417 infected and 2,052 deaths. Thus far no effective therapeutics in the form of vaccines and drugs against the disease has been discovered. Since in silico studies have formed a basis for discovery of new drugs, we started screening phytochemicals from various plants to evaluate their binding energies to the main protease of COVID-19, otherwise known as C3-like protease or 3CLpro, which is essential for viral attachment, fusion and edntry. Molecular docking (blind) studies were done with the help of AutodockVina. In our screening, we observed that a triterpenoid, hispidol A, found in the plant Picrasma javanica, gave good binding result to the C3-like protease with a binding energy ΔG of -7.6 kcal/mol. Since the binding of this compound was to the active site of the protease, a region to which a synthetic inhibitor N3 also binds, hispidol A can be potentially an inhibitory drug for COVID-19.

Keywords: *COVID-19, SARS-CoV-2, C3-like protease, Picrasma javanica*

Introduction

Respiratory diseases in humans including cold, pneumonia, and bronchitis are associated with corona viruses (hCoVs) belonging to the family Coronaviridae and so named because the viruses contain a spike protein (S protein) on their surface resembling a corona. Seven HCoVs have been identified so far, namely HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1, severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2 or COVID-19, the last emerging in late December 2019 in Wuhan, China and is the cause of the present pandemic. Four HCoVs (HCoV-229E, HCoV-NL63, HCoV-OC43 and HCoV-HKU1) are commonly circulated in the human population and contribute to approximately one-third of common cold infections in humans every year (Van Der Hoek, 2007). Coronaviruses infect humans by attaching to specific receptors on the human cell through their spike (S) protein. For SARS-CoV and SARS-CoV-2, the receptor is the angiotensin-converting enzyme (ACE)-2. Nucleocapsid is released into the cytoplasm of the host cell following receptor binding (Lim et al, 2016). The genome of both SARS and SARS-CoV-2 encodes two large polyproteins, pp1a and 5 pp1ab. These polyproteins are cleaved to form mature non-structural proteins (NSPs) by the two proteases 3CLpro (3C-like protease or chymotrypsin-like protease) and PLpro (Papain Like Protease) encoded by the open reading frame 1. The 3C-like proteases of SARS and SARS-CoV-2 have a high percentage of sequence identity and differ by only 12 amino acids;

1. Department of Pharmacy, Southeast University, Dhaka
2. Department of Biotechnology & Genetic Engineering, University of Development Alternative, Lalmatia, Dhaka
*Corresponding author: Nusratun Nahar, Assistant Professor, Department of Pharmacy, Southeast University, Banani, Dhaka-1213, Bangladesh

the substrate binding pockets exhibit a strikingly high level of alignment of the key residues involved in substrate binding, including the CYS145-HIS41 dyad, and HIS163/HIS172/GLU166 (Yang et al, 2003). Since the 3CLpro is necessary for viral attachment and entry, researchers believe that it can be a key target for development of therapeutics in the form of inhibitor molecule(s) and efforts are being carried out to find a suitable inhibitor to this protease.

6LU7 is the monomer of the C3-like protease in the active state with the inhibitor N3 attached. The main protease monomer contains three domains of which domains 1 and 2 (residues 8-101 and residues 102-184) form a chymotrypsin-like fold and are responsible for catalysis (Macchiagodena et al, 2020). Thus binding to these two domains by a phytochemical inhibitor can stop the catalytic activity of the protease and eventually viral replication.

As of July 5, 2020, the disease has infected 11,425,130 persons and caused deaths of 534,487 individuals throughout most of the countries of the world. The statistics for Bangladesh are 162,417 infected and 2,052 deaths. The disease still has not shown any symptoms of abating; in fact, the disease is on the rise in many countries of the world. As a result, a drug or a vaccine against this disease is desperately needed but none seems to be forthcoming within a short time period.

Picrasma javanica Blume belongs to the family Simaroubaceae and has ethnic uses as an antiviral agent. We therefore focused on the various chemical compounds present in the plant with the objective of performing in silico studies of bonding between these phytochemicals and the C3-like protease of COVID-19 or SARS-CoV-2. The protease will also be referred to in this manuscript as $3CL^{pro.}$

Methods

Three-dimensional Structure of COVID-19 Major Pprotease (3C-like Protease)

The pdb file (6LU7) of the main protease of SARS-CoV-2 3C-like protease or SARS-CoV-2 3CLpro published before (Liu et al, 2020) was used in the present study following removal of inhibitor N3. The active residues of SARS-CoV-2 3C-like protease are His41 and Cys145. Monomeric form of protein was used for molecular docking.

Compounds used in Docking Studies

We have studied various classes of phytochemicals known to occur in *Picrasma javanica*. However, only that of hispidol A, a triterpenoid will be presented here. Ligand molecules were downloaded from Pubchem (Ihlenfeldt, 2018) in sdf format. Optimization was done with the force field type MMFF94 using Openbable softwares and saved as pdbqt format.

Ligand Molecular Docking Studies

Molecular docking (blind) was conducted using AutoDock Vina (Trott et al, 2010). We report ΔG values as an average of six values from the docking program. We show the pose of phytochemicals bound to SARS-CoV-2 main protease in our figures as obtained from PyMOL and displayed in Discovery Studio (Systemes, D. 2015).

Phytochemicals

The structure of hispidol A is shown in Figure 1.

Results and Discussion

Hispidol A (Figure 1) has a molecular formula of C30H52O4 and a molecular mass of 476.732. It is a triterpenoid by nature and besides *Picrasma javanica*, it has also been reported from *Homalolepis suffruticosa* Engl. (Simaroubaceae). The interaction of hispidol A with the various amino acids of the protease in molecular docking studies is shown in Figure 2. The binding energy ΔG was -7.6 kcal/mol.

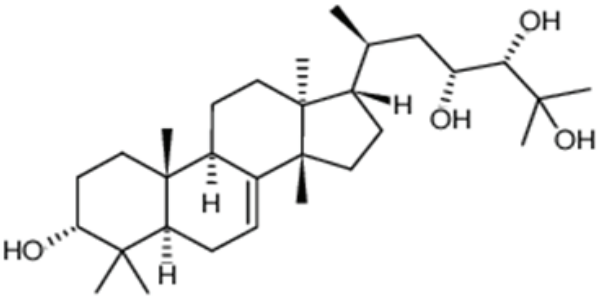

Figure 1. Hispidol A

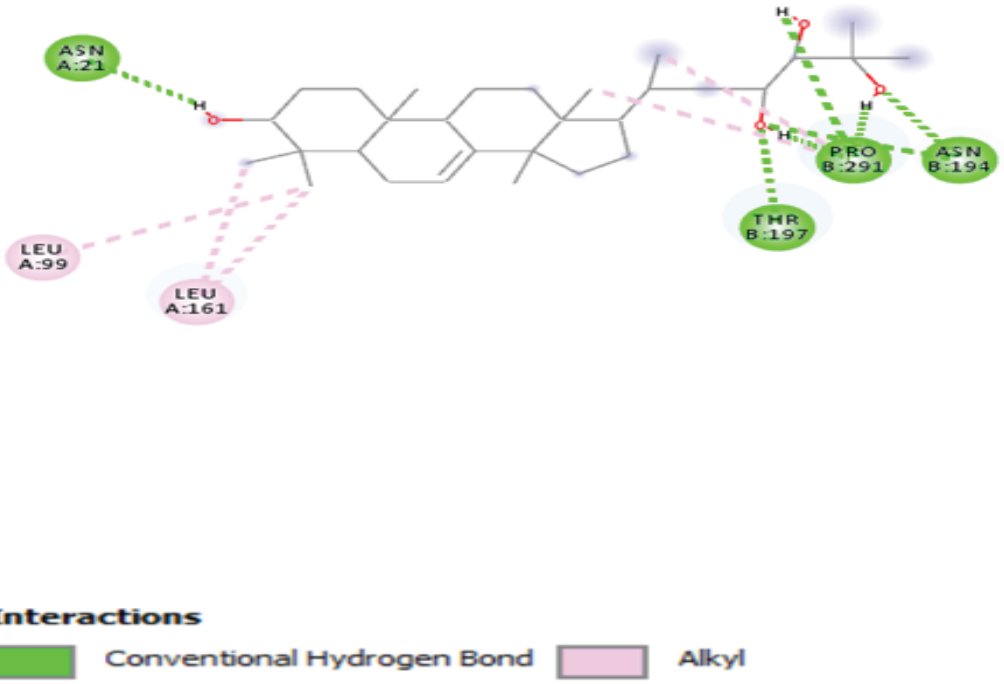

Figure 2. Interaction of hispidol A with C3-like protease of COVID-19.

The interaction of Leu99 and Leu161 with the methyl groups of hispidol A would make for strong hydrophobic interactions and which is augmented with a polar interaction between the –OH group and Asn21. The docking studies indicate that the major interactions of hispidol A are with domain 1 of the C3-like protease. Hispidol also interacts with Asn194, Thr197 and Pro291 but these sites are somewhat distal to domain 2 of the protease. However, these distal interactions may also contribute to the inhibition of the protease.

Conclusion

A triterpenoid, hispidol A, found in the plant Picrasma javanica, in in silico studies was found to bind to the C3-like protease of COVID-19 with a binding energy G of -7.6 kcal/mol. Although actual anti-viral laboratory tests are needed before coming to a definite conclusion, hispidol A appears to be a promising compound for COVID-19 inhibition.

Acknowledgements

This work was supported by funding from all the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

Ihlenfeldt, W. D. (2018). PubChem. A*pplied Chemoinformatics: Achievements and Future Opportunities,* 245-258.

Lim, Y. X., Ng, Y. L., Tam, J. P., & Liu, D. X. (2016). Human coronaviruses: a review of virus–host interactions. *Diseases, 4*(3), 26.

Liu, X., Zhang, B., Jin, Z., Yang, H., & Rao, Z. (2020). The crystal structure of COVID-19 main protease in complex with an inhibitor N3. *Protein Data Bank.*

Macchiagodena, M., Pagliai, M., & Procacci, P. (2020). Inhibition of the main protease 3cl-pro of the coronavirus disease 19 via structure-based ligand design and molecular modeling. *ar Xiv preprint arXiv:2002.09937.*

Systemes, D. (2015). BIOVIA, discovery studio modeling environment. Release 4.5. *Dassault Systemes: San Diego, CA.*

Trott, O., & Olson, A. J. (2009). Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function. *Efficient Optimization, and Multithreading.*

Van Der Hoek, L. (2007). Human coronaviruses: what do they cause?. *Antiviral therapy, 12*(4 B).

Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., ... & Gao, G. F. (2003). 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), 13190-13195.