12-1-2020

Inhibitors of the SARS-CoV-2 Papain-like Protease

Zachary P. Del Mundo

University Honors Program
Capstone Faculty Approval Page

Capstone Title (print or type)

Inhibitors of the SARS-CoV-2 Papain-like Protease

Student Name (print or type) Zachary Del Mundo

Faculty Supervisor (print or type) Timothy J. Hagen

Faculty Approval Signature Timothy J. Hagen

Department of (print or type) Chemistry and Biochemistry

Date of Approval (print or type) December 1, 2020

Date and Venue of Presentation _____

Check if any of the following apply, and please tell us where and how it was published:

☐ Capstone has been published (Journal/Outlet):

_____

☐ Capstone has been submitted for publication (Journal/Outlet):

_____

Completed Honors Capstone projects may be used for student reference purposes, both electronically and in the Honors Capstone Library (CLB 110).

If you would like to opt out and not have this student's completed capstone used for reference purposes, please initial here: _____ (Faculty Supervisor)

NORTHERN ILLINOIS UNIVERSITY

Inhibitors of the SARS-CoV-2 Papain-like Protease

A Capstone Submitted to the

University Honors Program

In Partial Fulfillment of the

Requirements of the Baccalaureate Degree

With Honors

Department Of

Chemistry and Biochemistry

By

Zachary Del Mundo

DeKalb, Illinois

8 May 2021

Abstract

The papain-like protease (PLpro) of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a target enzyme for treatment of the 2019 coronavirus infectious disease (COVID-19). Multiple lead competitive inhibitors to the protease have been previously discovered, particularly VIR250, HGN1034, and GYX402. Analogs of these ligands were designed using ChemSketch or Chem3D. The ligand-active site interactions of those unique ligands were observed and compared using AutoDock Tools, which also computes a predicted inhibitory constant (K_i). Over the Fall 2020 semester, proficiency in performing several functions of AutoDock Tools was acquired, and a total of thirteen unique and original compounds designed based on the aforementioned leads were tested. Three analogs of HGN1034 showed to have the strongest potential for future research and application: ZDM16 had the closest docking pose to the lead HGN1034 ($K_{i\ ZDM16}$ = 329.04 nM); ZDM13 and ZDM14 had the lowest computed inhibitory constants, lower than that of HGN1034 ($K_{i\ ZDM14}$ = 15.96 nM < $K_{i\ ZDM13}$ = 16.45 nM < $K_{i\ HGN1034}$ = 19.78 nM). The limitations of computational chemistry are considered and rationalized, and prospective follow-up studies and possibilities for future designs of inhibitors of the SARS-CoV-2 PLpro enzyme are discussed.

Introduction

Drug design and discovery for the treatment of patients infected with SARS-CoV-2 is a critical contemporary field of medicinal chemical research, concerning the ubiquitous COVID-19 pandemic.[1-4,6] A potential target enzyme of this coronavirus is the papain-like protease (PLpro), a tetrameric cysteine protease unique to the virus. Essential for SARS-CoV-2 maturation and proliferation, inactivating this enzyme with an inhibitory small molecule would effectively suppress the pathogenic activity of the virus within the human body.[3,4,6]

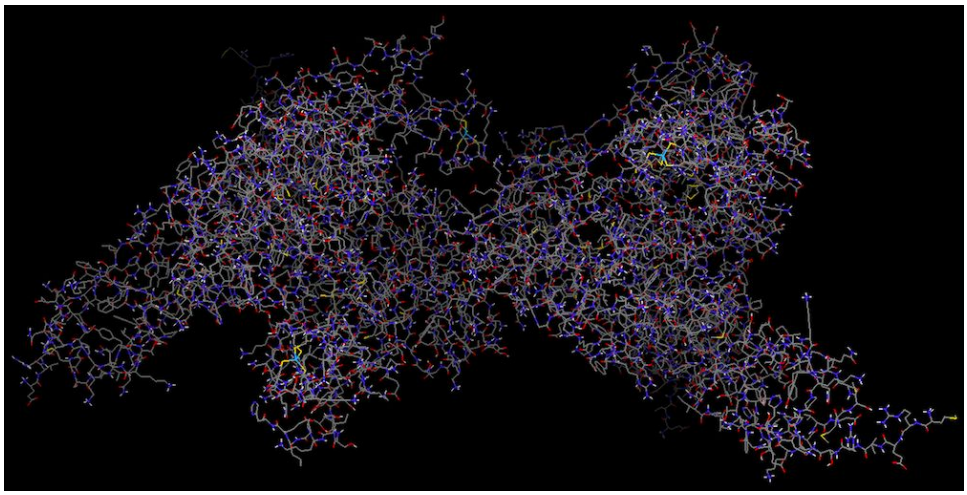


Figure 1. Lines display of PLpro (from AutoDock Tools)

Medicinal chemists have strived to develop an effective drug targeting PLpro to treat COVID-19. Since the beginning of the outbreak, several synthetic small molecules have been designed and tested with PLpro both virtually and in laboratory, and have shown high inhibition to the enzyme.[2,4] These lead compounds are used as the structural basis of designing new drug molecules with potentially higher potency.[3-6] The purpose of this research is to design and discover PLpro inhibitors with potentially higher inhibition than leads previously tested and disclosed to the public (available through the Protein Data Bank).

The computational molecular viewing and modeling software, AutoDock Tools (AutoDock), is used to simulate the docking of an inhibitory ligand in the active site of an enzyme essential to an infectious virus (or bacterial species). For each "docking," AutoDock accounts for the bond lengths, rotatable and non-rotatable bonds, and partial charges of a ligand, generating a detailed visualization of the position of the small molecule with respect to active

site residues. This enables the delineation of specific enzyme-ligand intermolecular interactions. The software also computes various measurements of binding affinity, particularly the inhibitory constant (K_i) of a ligand.[5,6]

The lead PLpro inhibitor most explored in this research was HGN1034, tested by the Hagen Group of the NIU Department of Chemistry and Biochemistry. The structure of this small molecule can be slightly modified to test for changes in inhibition. If the docking results of such an analog suggest similar or even higher inhibition than the lead, it may be produced in an organic laboratory and then tested with a sample of PLpro in real life. Figure 2 below is the skeletal formula of HGN1034.

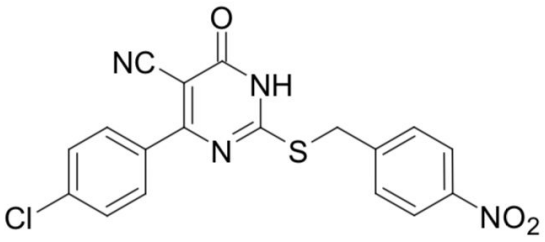


Figure 2. HGN1034

Methods

Ligand design and preparation: Ligands are designed using either of the molecular drawing softwares, ChemSketch or Chem3D. A new molecule can be created from scratch, however an existing file of a lead is usually modified. The MDL Molfiles (.mol) of the sketches are saved from these softwares and then converted to Mol2 format (.mol2), so that they can be opened in AutoDock. This conversion is carried out using a molecular file conversion software called OpenBabel. The final step in preparing a ligand for a docking involves AutoDock; the bond lengths, rotatable and non-rotatable bonds, and partial charges of the ligand are processed and the ligand is saved as a PDBQT file (.pdbqt). PDBQT stands for Protein Data Bank (PDB), partial charge (Q), and atom type (T).

Enzyme preparation: For the enzyme, a PDB file (.pdb) containing the complex three-dimensional structure of the macromolecule is obtained from the PBD itself and opened in AutoDock. Water molecules are removed and all polar hydrogens are made explicit. Additionally, the orignal ligand which comes with the PDB file is deleted to vacate the active

site. The "cleaned" enzyme is then saved as a PDBQT file as well. The four-letter code for the SARS-CoV-2 PLpro file obtained from the PDB is 6WUU.

The main parameter of a docking is the permutation of a "gridbox," which is to encompass an entire active site of the enzyme. The coordinates of a gridbox are specified in all three dimensions, and the gridbox is saved as a grid parameter file (.gpf). Figure 3 below displays the one gridbox used to dock inhibitors into PLpro for this research.

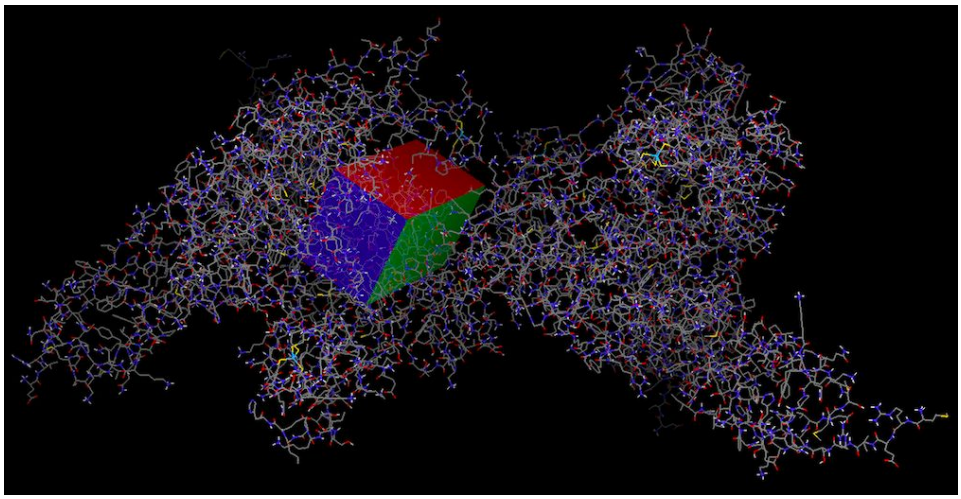


Figure 3. Gridbox encompassing one of the four PLpro active sites

Running AutoDock and Comparative Analysis: With PDBQT files of the ligand and enzyme and a saved gridbox, the docking can then be performed. An algorithm labelled "autogrid4" is first run to account for the possible interactions of the active site residues encapsulated in the gridbox. This is followed by the run of a second algorithm labelled "autodock4," to generate ten "docking poses" of the ligand inside the active site.

The docking poses of a newly designed ligand are analyzed by visually comparing each of its poses to the pose of a lead. This is done by superimposing the docking results over a previously created file of the enzyme with the lead already positioned in the active site. The enzyme is removed from the view, and the positions of the ligands with respect to the same active site are compared. For each tested PLpro inhibitor of this project, the docking pose most closely aligned with the lead was considered the best, and the computed K_i value was recorded and interpreted. Ligands resulting without a single visually good docking pose were deemed inadequate for future experimentation.

The compound VIR250, which comes pre-docked as a covalently bonded irreversible inhibitor in the 6WUU enzyme file of PLpro (the molecule has an alkene group which acts as a Michael acceptor to the active site cysteine residue), was used as the basis for comparison in determining the best docking poses of VIR250 analogs and two other leads, one of which was HGN1034, and the other a more recently discovered inhibitor called GYX402. The docking poses of HGN1034 and GYX402 analogs were then compared with the best poses of their respective leads.

Results

VIR250 and analogs: Three analogs of VIR250 were designed, substituting the group alpha to the carbonyl at the tail-end of the peptidomimetic backbone (Figure 4). These compounds were labelled ZDM1-3, being the first three unique analogs designed for this research. Along with the analogs, the lead VIR250 was docked into PLpro itself, and compared with the pose of the same molecule in the unedited PDB file. None of these four ligands resulted with a similar docking pose, including VIR250 itself as the exact same molecule. Therefore, no K_i values were recorded for these ligands (Table 1).

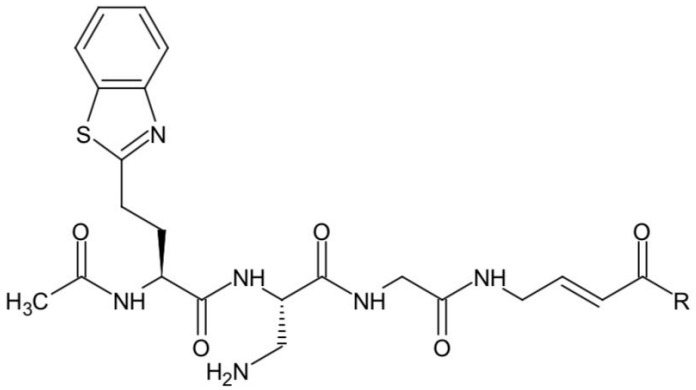


Figure 4. Basic structure of VIR250 and analogs.

T1. SAR table for VIR250 and ZDM1-3

Ligand	Formula	MW	R	K_i
VIR250	$C_{23}H_{30}N_6O_6S$	518.585	— OCH_3	N/A
ZDM1	$C_{20}H_{26}N_6O_5S$	462.522	— H	N/A
ZDM2	$C_{23}H_{27}N_7O_5S_2$	545.634	— C_3H_2NS	N/A
ZDM3	$C_{21}H_{25}F_3N_6O_5S$	530.52	— CF_3	N/A

HGN1034 and analogs: A total of nine analogs of HGN1034 were designed, ZDM 10-18, however ZDM16 and ZDM18 are not displayed in Table 2 since they do not have the basic structure indicated by Figure 5. In Tables 2 and 3, for each ligand with at least one good docking pose, the number of that pose is indicated to the right of the K_i value in parentheses.

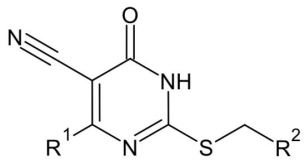


Figure 5. Basic structure of HGN1034 and analogs. Each substituent consists of an aromatic ring bonded at the 4 position unless marked otherwise.

T2. SAR Table for HGN1034 and ZDM10-15 and ZDM17

Ligand	Formula	MW	R1	R2	K_i
HGN1034	$C_{18}H_{11}ClN_4O_3S$	398.822	— C_6H_4Cl	— $C_6H_4NO_2$	19.78 nM (6)
ZDM10	$C_{19}H_{14}N_4O_3S$	378.406	— $C_{19}H14N_4O_3S$	— C_7H_7	N/A
ZDM11	$C_{21}H_{18}N_4O_3S$	406.457	— $C_{21}H_{18}N_4O_3S$	— C_9H_{11}	N/A
ZDM12	$C_{22}H_{20}N_4O_3S$	420.484	— $C_{22}H_{20}N_4O_3S$	— C_4H_9	N/A
ZDM13	$C_{17}H_{11}N_5O_3S$	365.365	— C_5H_4N (4-pyridyl)	— $C_6H_4NO_2$	16.45 nM (1)
ZDM14	$C_{17}H_{11}N_5O_3S$	365.365	— C_5H_4N (3-pyridyl)	— $C_6H_4NO_2$	15.96 nM (8)
ZDM15	$C_{17}H_{11}N_5O_3S$	365.365	— C_5H_4N (2-pyridyl)	— $C_6H_4NO_2$	1.37 μM (10)
ZDM17	$C_{17}H_{11}ClN_4OS$	354.813	— C_6H_4Cl	— C_5H_4N (3-pyridyl)	347.91 nM (6)

For the lead HGN1034, pose 6 was most similar to the pose of VIR250 in the PDB file (Figure 6), with a K_i of 19.78 nM. The longest carbon chain of the molecule shows a similar conformation; the length of the HGN1034 molecule follows the length of VIR250 molecule.

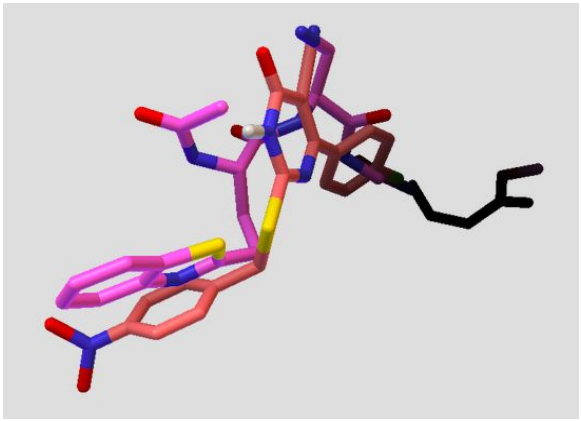


Figure 6. Pose 6 of HGN1034 (coral) against pose of covalently bonded VIR250 (magentain the PLpro active.

Pose 6 of HGN1034 was then used as the basis for comparison in analyzing the poses of ZDM10-18. Between ZDM10-15 and ZDM17, the four compounds, ZDM13, ZDM14, ZDM15, and ZDM17 had good docking poses. ZDM13 (pose 1) and ZDM14 (pose 8) had the lowest K_i values of 16.45 and 15.96 nM respectively, lower than that of the lead HGN1034. Figure 7 below displays the docking poses of both ZDM13 and 14 against pose 6 of HGN1034.

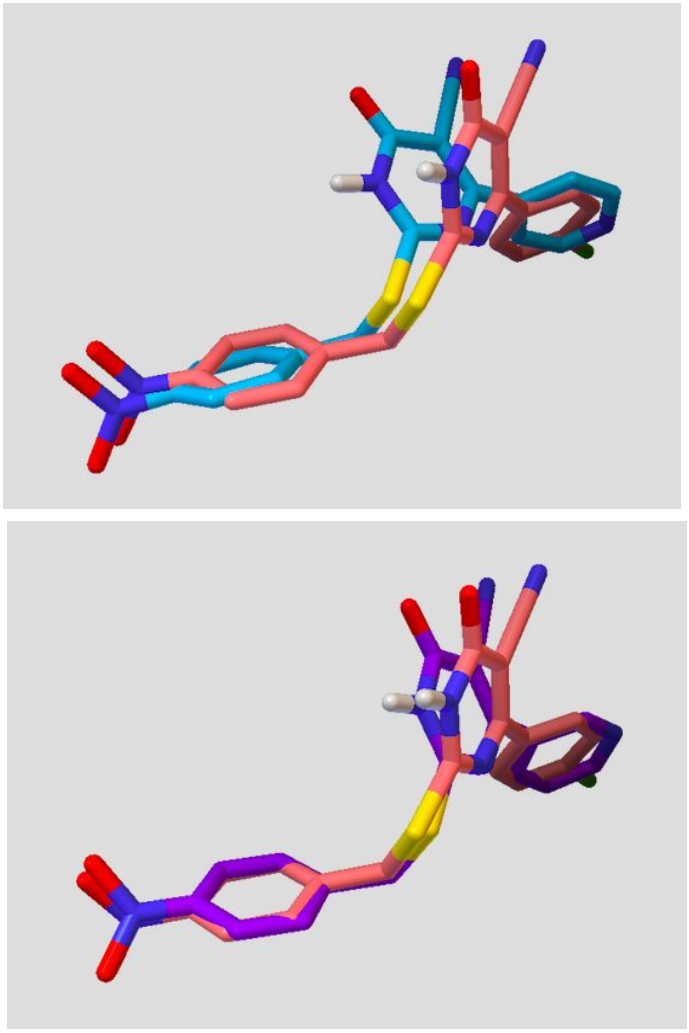
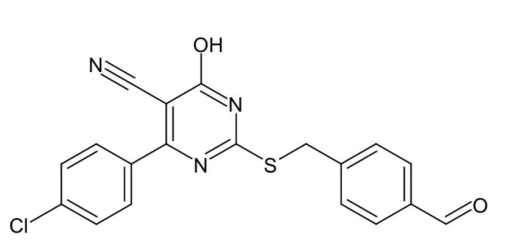


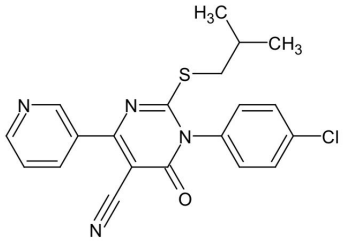
Figure 7. Top: Best pose of ZDM13 (cyan) against pose of HGN1034. Bottom: ZDM14 (violet) against HGN1034 (coral). Both analogs show relatively close alignment to the lead.

ZDM16, ZDM18, GYX402, and ZDM20: The structures and K_i values of these last four compounds are indicated in Figure 8 and Table 3. ZDM16 is not included in Table 2 because unlike ZDM10-15 and ZDM17, the second six-membered ring is aromatic, forming a pyrimidine ring instead of a cyclic amide, therefore bearing a hydroxyl instead of a carbonyl. ZDM18 is not included in Table 2 either since the first and third six-membered rings are para-positioned with respect to each other and the second ring. Notwithstanding the structural differences between these two compounds and those indicated in Table 2, ZDM16 and ZDM18 are still considered HGN1034 analogs, thus their docking poses were compared to pose 6 of HGN1034.

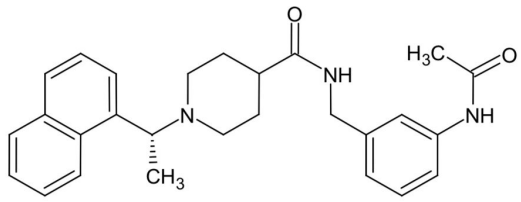
GYX402 is a recently disclosed PLpro inhibitor lead (PDB code: 7D7T) which bears a naphthyl group. As was done for HGN1034, the docking poses of this ligand were compared to the pose of the covalently bonded VIR250 in the original PDB file (6WUU). ZDM20 is an analog of both GYX402 and HGN1034, containing groups from both leads merged together, particularly the chlorophenyl of HGN1034.



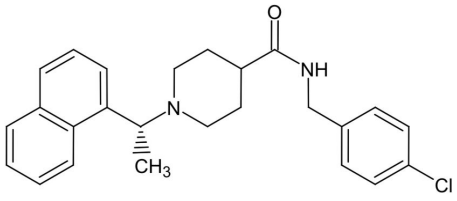
ZDM16



ZDM18



GYX402



ZDM20

Figure 8. Structures of ZDM16, ZDM18, GYX402, and ZDM20.

T3. SAR table for ligands in Figure 8

Ligand	Formula	MW	K_i
ZDM16	$C_{19}H_{12}ClN_3O_2S$	381.835	329.04 nM (4)
ZDM18	$C_{20}H_{17}ClN_4OS$	396.893	270.23 nM (8)
GYX402	$C_{27}H_{31}N_3O_2$	429.553	86.13 nM (7)
ZDM20	$C_{25}H_{27}ClN_2O$	406.947	217.71 nM (5)

Between the docking poses of all the unique PLpro inhibitors (thirteen) designed and tested this semester, pose 4 of ZDM16 was the pose most closely aligned with pose 6 of HGN1034 (Figure 9), showing a very similar confirmation and high superimposability. However, despite this high degree of alignment, its computed K_i value of 329.04 nM is much greater than those of ZDM13 and ZDM14 (16.45 and 15.96 nM).

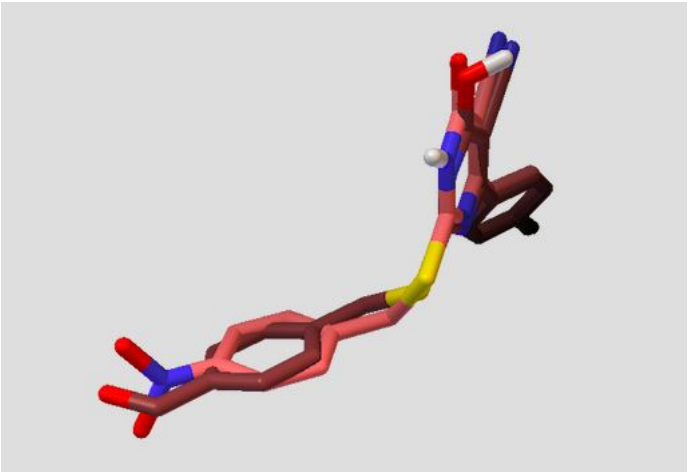


Figure 9. Pose of ZDM16 (sequoia) against pose of HGN1034. Compounds show very close alignment.

The best pose of ZDM18 was pose 8 (Figure 10). The fit is similar to those of ZDM13, ZDM14, and ZDM16, considering the close positions of two of the three six-membered rings, the cyano group, and the thioether. However, the chlorophenyl group is quite protuberant with respect to HGN1034.

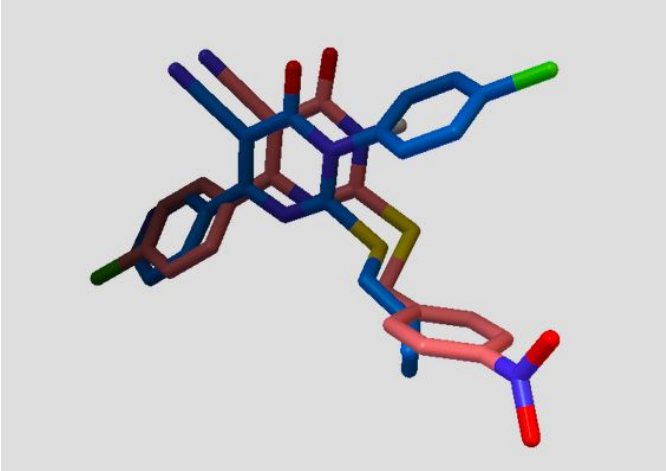


Figure 10. Pose of ZDM18 (blue) against pose of HGN1034. The chlorophenyl group is notably protuberant.

The docking poses of GYX402 were compared to the lead VIR250, rather HGN1034, since it is not part of the family of HGN1034 analogs. The naphthyl group of GYX402 in pose 7 of the ligand was positioned similarly to the benzothiazole group of the covalently bonded VIR250 molecule (Figure 7). When using this pose of GYX402 as the basis of comparison for analog ZDM20 (pose 5), the naphthyl groups of both ligands have similar positions.

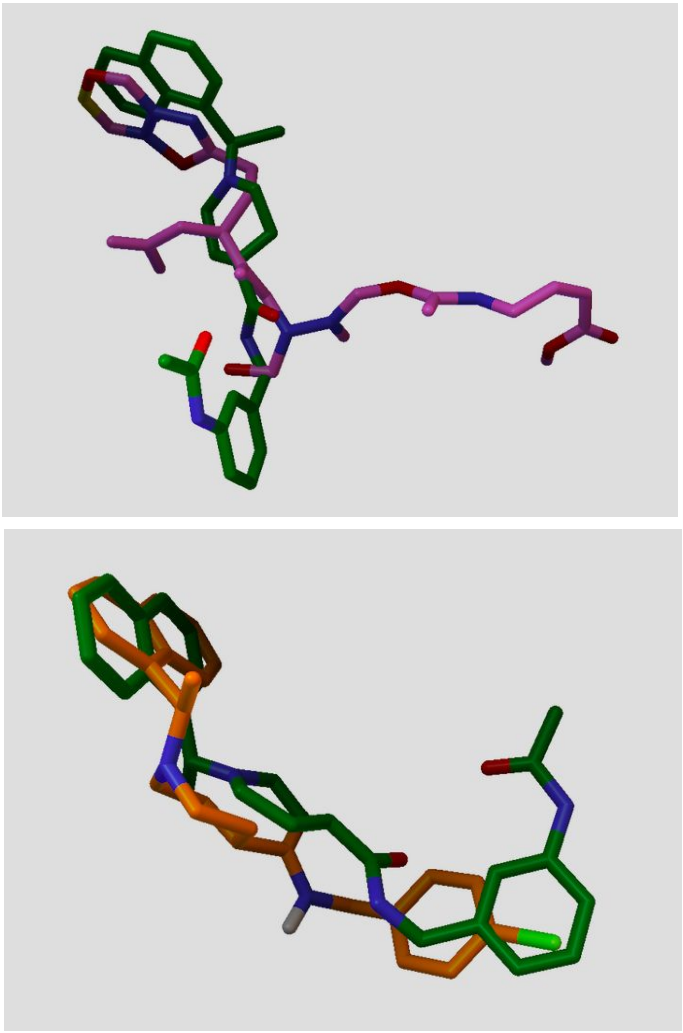


Figure 11. Top: Pose of GYX402 (green) against pose of covalently bonded VIR250. Bottom: Pose of ZDM20 (orange) against pose of GYX402. The fused two-ring systems of each ligand have similar positions.

Discussion

PLpro inhibitors with the strongest potential based on docking results: Of the thirteen unique and original SARS-CoV-2 papain-like protease inhibitors tested this semester (ZDM1-3, 10-18, 20), the ones which showed the greatest potential to serve as leads or usable drugs in future COVID-19 drug research were ZDM16, ZDM13, ZDM14. The remarkably close alignment of pose 4 of ZDM16 with pose 6 of HGN1034 in the PLpro active site (Figure 9) suggests that the two compounds have similar inhibition. On the contrary, however, the computed inhibitory constant for ZDM16 of 329.04 nM is about 17 times greater than that of HGN1034, 19.78 nM, suggesting that the potency of ZDM16 is about 17 times weaker. Compared to ZDM16, the quality of the results for the docking poses and K_i values of ZDM13 and ZDM14 are the converse. The best poses of ZDM13 and ZDM14 (Figure 7) did not align with the lead HGN1034 as well as ZDM16 (although they were close), and the computed K_i values for both ligands are in fact lower that than of HGN1034 ($K_{i\ ZDM14}$ = 15.96 nM < $K_{i\ ZDM13}$ = 16.45 nM < $K_{i\ HGN1034}$ = 19.78 nM).

For these three HGN1034 analogs, the discrepancy between the docking pose and the computed K_i value is rationalized due to the use of molecular modeling software, rather producing and testing the compounds in a laboratory in real life. In general, computational chemistry serves only to predict chemical properties and interactions, including inhibition.

Although the docking poses of ZDM13 and ZDM14 were less similar to HGN1034 than the pose of ZDM16, this does not necessarily indicate that the poses of ZDM13 and ZDM14 are worse. The differences in their poses may actually be due to their higher inhibition (lower K_i values), as computed by AutoDock. It is not absolutely required for the pose of an analog to be as close as possible to the lead.

Prospective analysis for ligands with the best results: Much of the semester was spent learning the highly specific and complex basics of using AutoDock Tools. A significant future area of focus for analyzing the docking results of ZDM13, ZDM14, and ZDM16 would be to measure a vast multitude of binding distances (Å) between atoms of the ligand and atoms of active site residues. Measurements taken for each compound individually docked in the PLpro active site can be compared to rationalize differences in K_i, especially for ZDM16.

Future inhibitor designs: The compounds VIR250, HGN1034, ZDM16, ZDM13, ZDM14, and GYX402 could all be used as leads for the design of new SARS-CoV-2 PLpro inhibitors in the future. Most of the unique and original ligands of this project were designed by substituting a single functional group. As was done with GYX402 and HGN1034 to design ZDM20, groups from two or more different leads can be merged together to combine the key interactions of two or more individual ligands into one new molecule. This would involve a detailed and precise comparative analysis of the positions of certain groups in the docking poses of individual ligands.

Conclusive statements: In light of the current coronavirus pandemic, the design and discovery of inhibitors of the papain-like protease in SARS-CoV-2 is a fast-growing area of organic and biochemical research in which new ideas and discoveries are being made and disclosed rather often. With strong resources, such as the Protein Data Bank and previously collected, and power molecular simulation softwares, such as AutoDock Tools, an extensive understanding of the ligand-active site interactions of a synthetic small molecule with an enzyme can be developed. An equipped background in the chemistry of drug design and prowess with AutoDock Tools is a significant first step in being introduced to the pharmaceutical industry.

References

1) Alamri M. A., et al. (2020). Structure-based virtual screening and molecular dynamics of phytochemicals derived from Saudimedicinal plants to identify potential COVID-19 therapeutics. *Arabian Journal of Chemistry*, (13), 7224-7234.
2) Amin, S. A., et al. (2020). Chemical-informatics approach to COVID-19 drug discovery: Monte Carlo based QSAR, virtual screening and molecular docking study of some in-house molecules as papain-like protease (PLpro) inhibitors. *Journal of Biomolecular Structure and Dynamics*, 1-10.
3) Erlanson, D. A. (2020). Many small steps towards a COVID-19 drug. *Nature Communications*, *11*, 5048.
4) Ghosh, A. K., et al. (2020). Drug Development and Medicinal Chemistry Efforts toward SARS-Coronavirus and Covid-19 Therapeutics. *ChemMedChem*, *15*, 1-27.
5) Helgren, T. R., & Hagen, T. J. (2017). Demonstration of AutoDock as an Educational Tool for Drug Discovery. *Journal of Chemical Education*, (94), 345-349.
6) Mothay, D. & Ramesh, K. V. (2020). Binding site analysis of potential protease inhibitors of COVID-19 using AutoDock. *VirusDisease*, *31*(2), 194-199.