

A Virtual Exploration of D-peptides as Effective Inhibitors of IgA1 Proteases

Daniel D. Pearce

Research Advisor: Dr. Paul H. Mueller

Hampden-Sydney College Chemistry Department

Introduction:

Ever since the inception of antibiotics, they have been used to treat various ailments caused by bacteria. However, in the middle of the 20th century, bacteria began to develop resistance to antibiotics due to widespread overuse and inappropriate prescribing (Ventola). As a result, new antibiotics were developed to combat the inefficiency of the old antibiotics, but this process was unable to stop bacteria from once again developing immunity. Today, the same dilemma with antibiotics is present. While some progress has been made in order to slow the antibiotic resistance crisis, antibiotic-resistant bacteria still pose a dangerous threat worldwide. One of the most dangerous types of bacteria are those that secrete IgA1 proteases. Inhibition of these bacteria by anti-virulence strategies would help to alleviate the antibiotic resistance crisis.

In the developed world, many bacterial infections are caused by bacterium that secrete an IgA1 protease (Wood et al). The protease attacks the hinge region of the IgA1 to split it into two fragments. IgA1 serves as the first line of defense against pathogens. When IgA1 is inhibited, the immunogenic response is thwarted and allows for bacteria to establish a foothold in the body. These bacteria include the genera *Neisseria*, *Haemophilus*, and *Streptococcus*. Currently, *N. Meningitidis* remains a prevalent cause of bacterial meningitis in the United States. Furthermore, *S. Pneumoniae* has been shown to be a major cause of community-acquired pneumonia and meningitis worldwide. The most at-risk groups for both species of bacteria are infants, children, and teenagers.

Despite their prevalence, bacteria that secrete IgA1 protease have no known universal inhibitor. The most likely reason is because the protease is unique to each species of these bacteria. Specifically, the area where the protease binds to the IgA1 hinge region is highly selective for each strain of bacteria. However, L-peptides with sequences based on the hinge

region of IgA1 have been shown to be competitive inhibitors of the IgA1 protease. While these L-peptides are fair inhibitors, they are still not that potent nor universal. In fact, the most potent L-peptides have been shown to be tetrapeptides.

Small peptides are not the only inhibitors of IgA1 proteases. Garner et al discovered natural products, such as remerine and berberine, to be effective inhibitors of *Streptococcus Pneumoniae* IgA1 protease (Garner et al). Shehaj et al found a benzoic ester to have an IC_{50} value of roughly $5.9\mu\text{M}$ as an inhibitor of *H. Influenzae* (Shehaj et al). Nonetheless, peptides likely get most of the attention because they make up the IgA1 hinge region and their characteristics make them good drug candidates.

L-peptides are short chains of L-amino acids. When facing down the hydrogen-carbon bond with the hydrogen in the foreground, L-amino acids have the carbonyl group, R-group, and amine group going in a clockwise direction in that specific order. Since R-groups can be varied, L-amino acids can have either a "S" or "R" configuration. When facing down the hydrogen-carbon bond with the carbon in the foreground, if the priority of the other functional groups goes clockwise, then the amino acid has a "R" configuration. If the priority of the functional groups goes counterclockwise, then the amino acid has a "S" configuration. However, since L-amino acids in L-peptides are connected usually by methylene groups, most L-amino acids will have an "S" configuration (The Merck Index).

D-peptides are short chains of D-amino acids. D-amino acids and L-amino acids are enantiomers of each other. The same is true for L-peptides and D-peptides. Therefore, a D-amino acid is a mirror image of its corresponding L-amino acid. When facing down the hydrogen-carbon bond with the hydrogen in the foreground, D-amino acids have the carbonyl group, amine group, and R-group going in a clockwise direction in that specific order. Again, D-amino acids

can have a “S” or “R” configuration depending on the R-group. Furthermore, since D-amino acids in D-peptides are usually connected by a methylene group, most D-amino acids will have an “R” configuration (The Merck Index).

While structurally similar to D-amino acids, L-amino acids have been observed to comprise the majority of naturally occurring amino acids in the human body. Specifically, only L-amino acids make up the building blocks of proteins. Furthermore, most amino acids that are essential to good health are L-amino acids (Rose). However, D-amino acids have attractive properties that make them suitable drug candidates. D-peptides are known to be less likely to trigger an immunogenic response (Welch) and more efficiently absorbed in the digestive tract (Pappenheimer). D-peptides also cost less to produce compared to L-peptides (Feng). Perhaps the most important reason is that D-peptides are highly resistant to proteolytic degradation. L-peptides are very susceptible to proteolytic degradation because proteases are designed to break down proteins and peptides which are comprised of L-amino acids (Feng).

As a result, a D-peptide that can mimic the therapeutic properties of an L-peptide could hold promise as a potential drug. For example, Dieter Willbold *et al* found that D-peptides could hold promise in blocking a crucial step in the development of Alzheimer’s disease. A common method to create a D-peptidomimetic is the retro-inverso analogue of L-peptides. This method involves reversing the amino acid sequence of the L-peptide and inverting the chirality of the central carbon of each amino acid involved, hence the name (Rai).

Due to the limitations imposed by COVID-19 restrictions, we could only conduct *in silico* “experiments” using docking software. While these “experiments” are not ideal, they still serve a purpose. Docking software is a relatively cheap method to “screen” for compounds but can be a time-consuming process, especially screening for thousands of compounds. Docking

programs differ in their approach for “screening” potential compounds, but significantly base their approach on the structure of the potential compounds. The reason for this basis is because structure and composition of compounds are intricately linked to the physical and chemical properties possessed by them. In the case of this research, docking programs will be used to “screen” for compounds that have structural properties similar to that of known inhibitors of IgA1 proteases.

It is also necessary for docking programs to account for chemical bonds, charge of the atoms, and the chemical potential energy attributed to the potential molecules being “screened” for research. In molecular modeling, these properties are accounted by “force fields” which model the forces between atoms within a molecule and that between molecules. The computational parameters for the “force fields” can differ, but are derived from quantum mechanics, experiments in chemistry and physics, or both. Nonetheless, “force fields” utilize interatomic potential functions to describe the energy landscape. “Force field” parameters also include information about different types of atoms, chemical bonds, three-dimensional interactions, nonbond interactions, and other data deemed relevant to the program.

Docking programs must account for the previously mentioned characteristics of molecules to be fairly accurate; nevertheless, docking parameters should include more empirical data to increase their accuracy. However, as more parameters are added to a docking program, the more computational power is needed to run the program. This conundrum results in the most accurate docking programs requiring an expensive computer with high computing power while the less accurate ones can run on computers with less computing power. Therefore, a tradeoff between accuracy and cost is usually the best option for selecting a docking program.

The only programs used for docking simulations were Autodock 4.2.6 and AutoDockTools (ADT). Like the parameters of a docking program, the exact method by which the programs model the receptor-ligand binding differs from one program to the next. AutoDock 4.2.6 uses a combination of the lock-and-key and the Induced fit models. The lock-and-key analogy states that the receptor is a rigid lock with a specific key that fits it perfectly, the ligand. The induced fit model expresses both the receptor and the ligand as more flexible. In addition, the structure of the receptor-ligand complex is not the same as the two molecules fitted rigidly together, rather the structure is more malleable and new bonds are created during the process. In the case of AutoDock 4.2.6, the receptor is a rigid lock and the ligand is a flexible key.

For AutoDock to enact this enzyme-ligand binding, the program uses PDB and PDBQT files. PDB files describe the three-dimensional structure of a molecule, whereas PDBQT files are PDB files that also state the partial charges on atoms and describe the AutoDock 4 atom types. The macromolecule or receptor is saved as a PDBQT file during the process of running AutoDock but must first be loaded in as a PDB file first. The ligand file can be saved the same way as the macromolecule or an existing PDBQT file of the ligand can be loaded into the program as well. Nonetheless, PDBQT files of both the receptor and ligand are needed to run AutoDock 4.2.6.

The “force field” used by AutoDock 4.2.6 in this research is the Assisted Model Building with Energy Refinement (AMBER) force field (Morris et al). This force field was designed to encompass a wide variety of interactions and bonds for biomolecules. Hence, the AMBER force field is great for simulating a docking between a protein and peptide. Furthermore, the program is based on high-computational data. According to Hornak et al, AMBER force fields were designed to be “suitable for protein simulations in the condensed phase.”

The main goal for this summer will be to design a D-peptide that is a potent inhibitor of *H. Influenzae* IgA1 protease. *H. Influenzae* was determined to be the target macromolecule because it is the only IgA1 protease present on the PDB. The sequence of the D-peptide will be based on some of the most potent L-peptides with sequences based on the hinge region of IgA1. Utilizing ChemSketch, I can build models of several proposed D-peptides and examine their properties, then use AutoDock to dock the proposed D-peptides to the IgA1 protease. Additional goals are to become well-acquainted with AutoDock and knowledgeable of the mechanism by which the *H. Influenzae* protease binds to IgA1.

Experimental:

All docking simulations began by opening AutoDock Tools (ADT). All PDB files and PDBQT files were either obtained from The Protein Data Bank or ChemSketch. OpenBabel was used to convert files if necessary. ChemSketch was used mostly for creating the ligand molecules as PDB files and then saving them as PDBQT files in ADT. Nonetheless, all receptor molecules were loaded into the program as PDB files and the ligand files as PDBQT files. Once the receptor or macromolecule was loaded into ADT, water was deleted from the molecule and non-polar hydrogens were merged into the molecule. Furthermore, Gasteiger charges were added to the macromolecule and the macromolecule was then saved as a PDBQT file. Next, the ligand was loaded into the program and the number of active torsions on it were set. Both the receptor and ligand were ready for AutoGrid.

AutoGrid is a feature that sets the energy landscape, determines the type of bonding that will occur, and allows the user to set the area of the binding site between the two molecules. The user sets up the covalent maps and the grid box as the parameters for a GPF file. AutoGrid

requires the user to manually input which molecule is the receptor and which is the ligand.

AutoGrid runs the GPF file in conjunction with a GLG file to prepare the receptor and ligand for AutoDock.

After AutoGrid is done running, the receptor and ligand can be prepared for AutoDock. Once again, the user must input which molecule is the receptor and which molecule is the ligand. Genetic algorithm parameters were used as the DPF output file used to run AutoDock. As with AutoGrid, a conjunction file will be used with the DPF file, namely a DLG file. AutoDock is ready to run. After AutoDock is done running, the dockings can be analyzed by the most-likely configurations ranked by energy.

Results:

The first docking simulation was to test the AutoDock program and confirm the known results. The receptor used was the COX-2 enzyme and Naproxen. After docking, it was clear that many of the conformations ranked by energy were located at the known active site of the COX-2 enzyme. Naproxen forms Van der Waals contacts within the binding site and forms hydrogen bonds to the side chains of Tyr355 and Arg120 (Duggan). The docking results from AutoDock were quite different from the known results. After some reevaluation, it was determined that an IZP ligand was also present in the complex which was skewing the data. Knowing this mistake and in the interest of time, research was resumed, and it was assumed that AutoDock would yield the right results if redone.

The rest of the docking simulations were docking small peptide inhibitors of IgA1 proteases to the *H. Influenzae* IgA1 protease. The first docking was the tetrapeptide Ser-Pro-Ser-Thr, one of Burton's best inhibitors, to the protease. The grid box encompassed the entire *H.*

Influenzae protease to determine the most likely binding site. The best binding site occurred around a large arm or side chain protruding from the main portion of the molecule. Specifically, there is a pocket close to the base of the side chain that seems to fit the tetrapeptide very well. Various residues of residues 570-630 of the protease seem to encompass this binding pocket. The *H. Influenzae* IgA1 protease is known to cleave the hinge region of IgA1 at a Proline and either a Serine or Threonine (Burton et al), although it is not known where the active site is located on the *H. Influenzae* protease. To help confirm the active site of the first docking was indeed on the sidechain arm, I conducted the same docking with a Pro-Ser-Pro-Ser-Thr-Pro peptide with the grid box encompassing the entire protease. While the dockings differed slightly, both had many of them at the same active site on the *H. Influenzae* protease. The last docking was of Acetyl-Ser-Thr-Pro-Pro to the protease. Once again, the peptide bonded at the same active site of the last two peptides but with multiple different conformations.

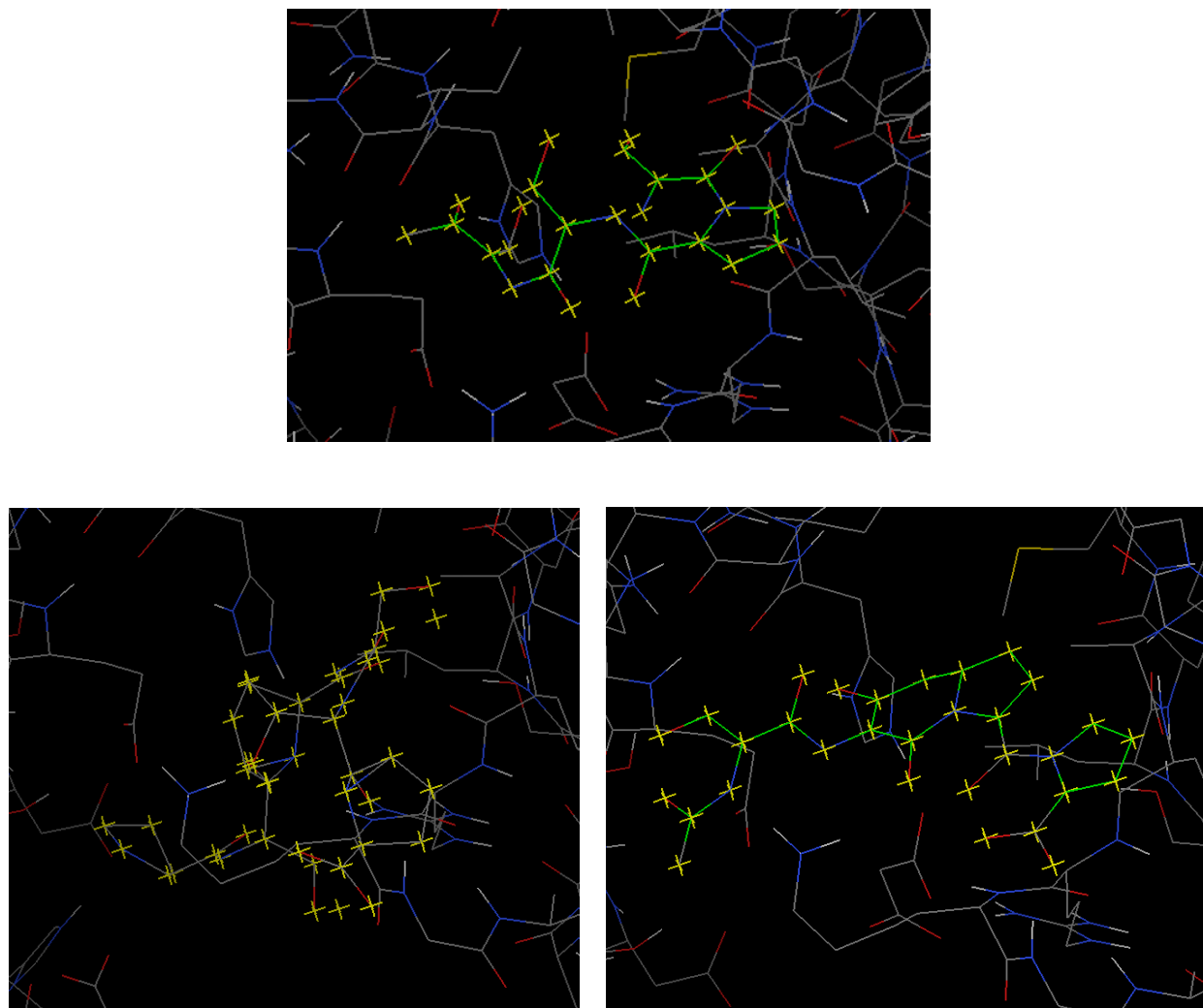


Figure 1. The most likely conformations ranked by energy of each of the three highlighted peptides. All conformations were in the same binding pocket with each having a slightly different pose. Counterclockwise from top: Ser-Pro-Ser-Thr, Pro-Ser-Pro-Ser-Thr-Pro, and Acetyl-Ser-Thr-Pro-Pro.

Conclusions:

While I was not able to complete the main goal of creating a D-peptide inhibitor of *H. Influenzae* IgA1 protease or understand the entire mechanism of the IgA1 protease, I did become well-acquainted with AutoDock and located a possible active site on the IgA1 protease. Based on the docking conformations ranked by energy, it seems that the active site of the *H. Influenzae* protease is a binding pocket between residues 570-630. Nonetheless, more dockings need to be conducted to confidently confirm the location of the active site on the *H. Influenzae* protease. More dockings will also help to determine an accurate representation of the bonds formed during the formation of the complex and which residues on the protease are binding to the ligands. It is also worth noting that near the active site is another ligand, malonic acid. While this ligand was present through all three dockings, it was quite small and seemed to contribute very little to the active site. Nonetheless, more accurate results will be obtained with that ligand removed from the protease. Future work would also include screening for peptides that could be potential inhibitors of IgA1 proteases.

References

- ACD/ChemSketch, version 2020.1.1, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2020.
- Burton, J.; Wood, S.G.; Lynch, M.; Plaut, A.G. Substrate analogue inhibitors of the IgA1 proteinases from *Neisseria gonorrhoeae*. *J. Med. Chem.* 1988, 31, 1647–1651. Retrieved from: <https://doi.org/10.1021/jm00403a027>
- Chi, Y. C., Rahkola, J. T., Kendrick, A. A., Holliday, M. J., Paukovich, N., Roberts, T. S., Janoff, E. N., & Eisenmesser, E. Z. (2017). *Streptococcus pneumoniae* IgA1 protease: A metalloprotease that can catalyze in a split manner in vitro. *Protein science: a publication of the Protein Society*, 26(3), 600–610. Retrieved from: <https://doi.org/10.1002/pro.3110>
- Garner, A. L., Fullagar, J. L., Day, J. A., Cohen, S. M., & Janda, K. D. (2013). Development of a high-throughput screen and its use in the discovery of *Streptococcus pneumoniae* immunoglobulin A1 protease inhibitors. *Journal of the American Chemical Society*, 135(27), 10014–10017. Retrieved from: <https://doi.org/10.1021/ja404180x>
- Halgren, T.A. (1996), Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.*, 17: 490-519. Retrieved from: [https://doi.org/10.1002/\(SICI\)1096-987X\(199604\)17:5/6<490::AID-JCC1>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P)
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, 65(3), 712–725. <https://doi.org/10.1002/prot.21123>
- Johnson, T. A., Qiu, J., Plaut, A. G., & Holyoak, T. (2009). Active-site gating regulates substrate selectivity in a chymotrypsin-like serine protease the structure of *haemophilus influenzae*

- immunoglobulin A1 protease. *Journal of molecular biology*, 389(3), 559–574. Retrieved from: <https://doi.org/10.1016/j.jmb.2009.04.041>
- Klein, A. N., Ziehm, T., Tusche, M., Buitenhuis, J., Bartnik, D., Boeddrich, A., Wiglenda, T., Wanker, E., Funke, S. A., Brener, O., Gremer, L., Kutzsche, J., & Willbold, D. (2016). Optimization of the All-D Peptide D3 for A β Oligomer Elimination. *PloS one*, 11(4), e0153035. Retrieved from: <https://doi.org/10.1371/journal.pone.0153035>
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) [Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility](https://doi.org/10.1002/anie.200901765). *J. Computational Chemistry* 2009, 16: 2785-91.
- Pappenheimer JR, Dahl CE, Karnovsky ML, Maggio JE. Intestinal absorption and excretion of octapeptides composed of D amino acids. *Proc Natl Acad Sci U S A*. 1994;91(5):1942-1945. Retrieved from: <https://doi:10.1073/pnas.91.5.1942>
- Rai, J. Peptide and protein mimetics by retro and retroinverso analogs. *Chem Biol Drug Des*. 2019; 93: 724– 736. Retrieved from: <https://doi.org/10.1111/cbdd.13472>
- Rose A. J. (2019). Amino Acid Nutrition and Metabolism in Health and Disease. *Nutrients*, 11(11), 2623. Retrieved from: <https://doi.org/10.3390/nu11112623>
- Shehaj, L.; Choudary, S.K.; Makwana, K.M.; Gallo, M.C.; Murphy, T.F.; Kritzer, J.A. Small-molecule inhibitors of Haemophilus influenzae IgA1 protease. *ACS Infect. Dis*. 2019, 5, 1129–1138. Retrieved from: <https://pubs.acs.org/doi/abs/10.1021/acsinfecdis.9b00004>
- The Merck Index, Merck & Co. Inc., Nahway, N.J., 11(1989); CRC Handbook of Chem.& Phys., Cleveland, Ohio, 58(1977)
- Ventola C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *P & T: a peer-reviewed journal for formulary management*, 40(4), 277–283.

Welch, B. D., VanDemark, A. P., Heroux, A., Hill, C. P., & Kay, M. S. (2007). Potent D-peptide inhibitors of HIV-1 entry. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 16828–16833. Retrieved from:

<https://doi.org/10.1073/pnas.0708109104>

Wood, S. G., Lynch, M., Plaut, A. G., & Burton, J. (1989). Tetrapeptide inhibitors of the IgA1 proteinases from type I *Neisseria gonorrhoeae*. *Journal of medicinal chemistry*, *32*(10), 2407–2411. Retrieved from: <https://doi.org/10.1021/jm00130a030>