Monaf Awwa

Targeting quorum sensing transcription factor LasR with natural products

Introduction:

The increasing prevalence of antibiotic resistant bacteria such as *Pseudomonas Aeruginosa* have created an urgent demand for alternative anti-pathogenic drug targets [1]. Recent advances in the understanding of bacterial gene expression have elucidated

a key mechanism, referred to as quorum sensing, by which bacteria coordinate phenotypes such as biofilm formation and fluorescence[2]. The concentration of particular molecules serves as the determinant of whether a particular phenotype will be expressed at the individual level[3]. Due to quorum sensing's intimate connection with pathogenic gene expression, it is a novel source of drug targets in pathogenic bacteria[4]. One validated target is the quorum sensing transcription factor Las Regulator (LasR) protein which upregulates several virulence factors[5]. In order to discover molecules with potential as anti-quorum sensing agents, computational studies can be applied to accelerate the discovery of effective *in vivo* inhibitors.

Pseudomonas Aeruginosa typically exists as a commensal organism in the human body [6]. It is an opportunistic pathogen, causing infections in situations such as flesh wounds or immunocompromised individuals[7]. With the turn of the century, classical treatment of the infection with existing antibiotics was largely nullified by genotypes bearing resistance. This resistance has been attributed to multiple mechanisms, including enzymes that digest antibiotics, biofilms that share resistance plasmids, and mutations to proteins that are existing drug targets[8]. Quorum sensing inhibition seems to be a highly promising solution to antibiotic resistance[9]. It has been postulated that resistance to quorum sensing inhibitors is unlikely due to its role as gene expression regulator rather than direct metabolic role. This remains to be seen, however mathematical models demonstrate that for robust treatment methods inhibitors at multiple quorum sensing loci are required[10]. Nevertheless, if quorum sensing is to be effectively manipulated for antibacterial purposes, a thorough understanding of its genetic control must be elucidated. The figure below shows how the major virulence factors regulate each other in a network schematic[11].

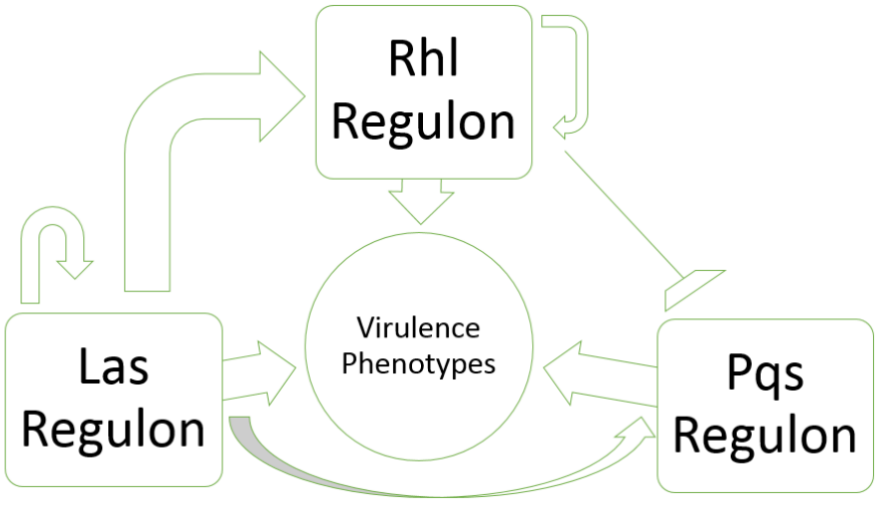


Figure 1. Schematic of upregulation

From this diagram it is clear that there is a sophisticated interplay of the Pqs, Rhl, and Las operons that activates pathogenic gene expression.

The chemical agonist of the las and Rhl operons is an acylhomoserine lactone, depicted below, whereas the agonist of the pqs operon is a 4-quinolone derivative [1].

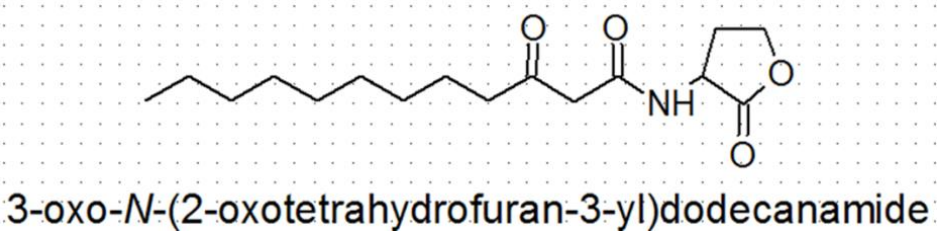


Figure 1. Natural agonist of the las transcription factor LasR

The sophisticated nature of the quorum sensing network enables *P. Aeruginosa* to apply its virulence for purposes such as competing with organisms for iron or phosphate uptake, as well as opportunistic and notorious infections it causes in human lung tissue[12]. The rationale is if there are enough bacterial cells in the vicinity their virulence factors can overwhelm innate host defenses. It should be noted that the rhl operon, which is regulated by the las operon, encodes the factors that lead to long-term biofilm maintenance. Rhl negatively regulates itself, a clever design of natural selection to ensure cell density remains high for biofilm maintenance[13]. The las operon loci is a validated anti-virulence target based on work with several compounds known to bind with it[14]. Previously, high throughput screens of inhibitors of the las synthase have elucidated multiple molecules with potential. Unfortunately, issues with cytotoxicity of inhibitors are a recurring theme in discovering anti-quorum sensing molecules. An interesting question remains regarding whether an inhibitor of the lasR would also directly interfere with downstream operon expression. The pqs operon has also been validated as a loci for anti-virulence agents, however it has a different native agonist than the acyl homoserine lactone of the Las and Rhl operons. The differences in network connections between the quorum sensing hierarchy opens the debate for the optimal drug targets, however for the current experiment the goal is to design an inhibitor of LasR.

Several natural synthesized molecules which have anti-quorum sensing properties have been elucidated. The structures bear a wide degree of variety. By examining the natural agonist, this feature can be rationalized by the distinct hydrophobic tail and amphiphilic ring of the acyl homoserine lactone shown in figure 1. A plethora of molecules have these two features, including cinnamaldehyde, vanillin, and zingerone [15] [16]. Other existing structures with known anti-quorum sensing properties include the diketopiperazine class, brominated furanones produced by algae, as well as other non-specifically cytotoxic compounds.

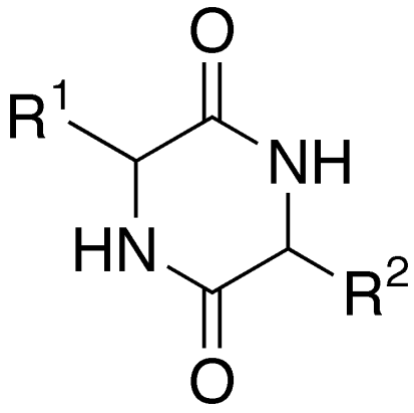
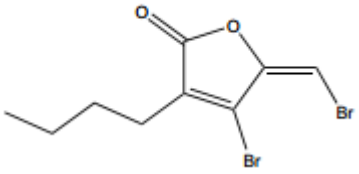
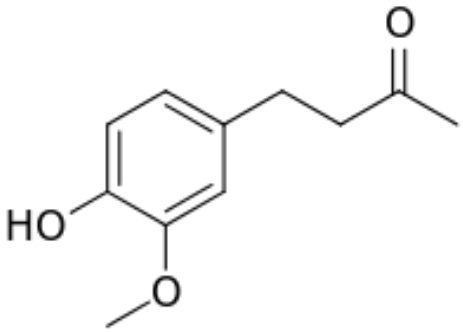
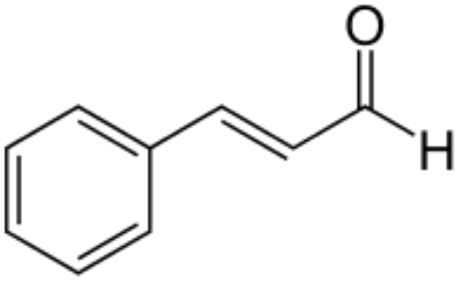
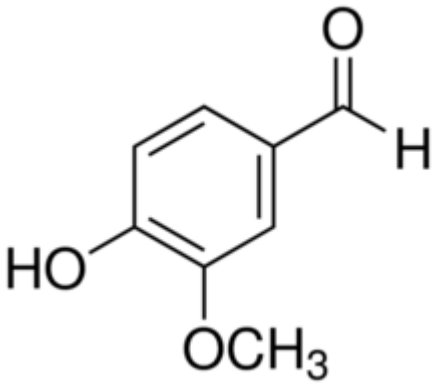


Figure 2. Diketopiperazine backbone. Biosynthesis of amino acids via claisen condensation yields R1 and R2 groups that are amino acid derivatives.

Natural Product Inhibitor	Structure
Red Algae brominated furanone	
Zingerone	
Cinnamaldehyde	

Vanillin	

Diketopiperazines have been known to have anti-quorum sensing properties, which is interesting considering they are often synthesized naturally by condensation of amino acids. These findings serve as one justification for the virtual screen of natural products. Interestingly, they all share an aromatic ring.

The methodology of the current experiment was to sample a diverse array of natural products *in silico* to find molecules predicted to have high affinity for the LasR binding site. It is plausible that there are allosteric inhibitors of the enzyme, however the scope of the current project pertains to discovering competitive inhibitors [15]. Natural product biosynthesis is under natural selection to always favor production of molecules with a purpose, another justification for sampling natural products. Considering that in a virtual screen the variety of chemical structures sampled will always be a tiny fraction of the limitless possibilities, beginning a drug design project with natural products accelerates the sampling of viable structures. This is key since a method such as a virtual screen relies on the stochastic probability that at least a few molecules from a much larger pool will be effective *in vitro* and *in vivo*. A major drawback of natural products tend to be the large number of rotatable bonds, which rapidly becomes difficult for the DOCK algorithm to deal with since sampling is based on anchoring a non-rotatable portion then sampling the rest of the molecule as rotatable. Another drawback is the chemical nature of the compounds; they may be difficult to synthesize efficiently, or difficult to modify chemically to enhance affinity to the target or uptake by human intestinal tissue. Nevertheless, if a molecule is found to be an effective inhibitor, computational studies can be expanded to determine if any chemical modifications will enhance binding affinity. The UCSF ZINC database has multiple sets of natural products based on geographic source as well as those ready to be sold[17].

In order to perform the experiments, the program UCSF DOCK version 6.6 was used to perform binding energy analysis of the protein and ligands. The algorithm that performs the binding energy calculation can be decomposed into van der Waals interactions and electrostatic interactions. A Lennard-Jones potential describes the van der Waal interaction scheme and a Coulombic potential describes the electrostatic energy. There are several limitations to the DOCK protocol applied to the virtual screening. The active form of the LasR transcription factor is a dimer of identical

subunits, yet only one binding site was put through virtual screen. However, since the binding of one ligand causes an increase in the affinity of the other subunits for the same ligand, it is difficult to tell how an inhibitor would behave in this respect. Also, a high affinity compound may function as an agonist for the dimerization of the transcription factor, rendering the compound useless for antibacterial purposes.

Computational Methods:

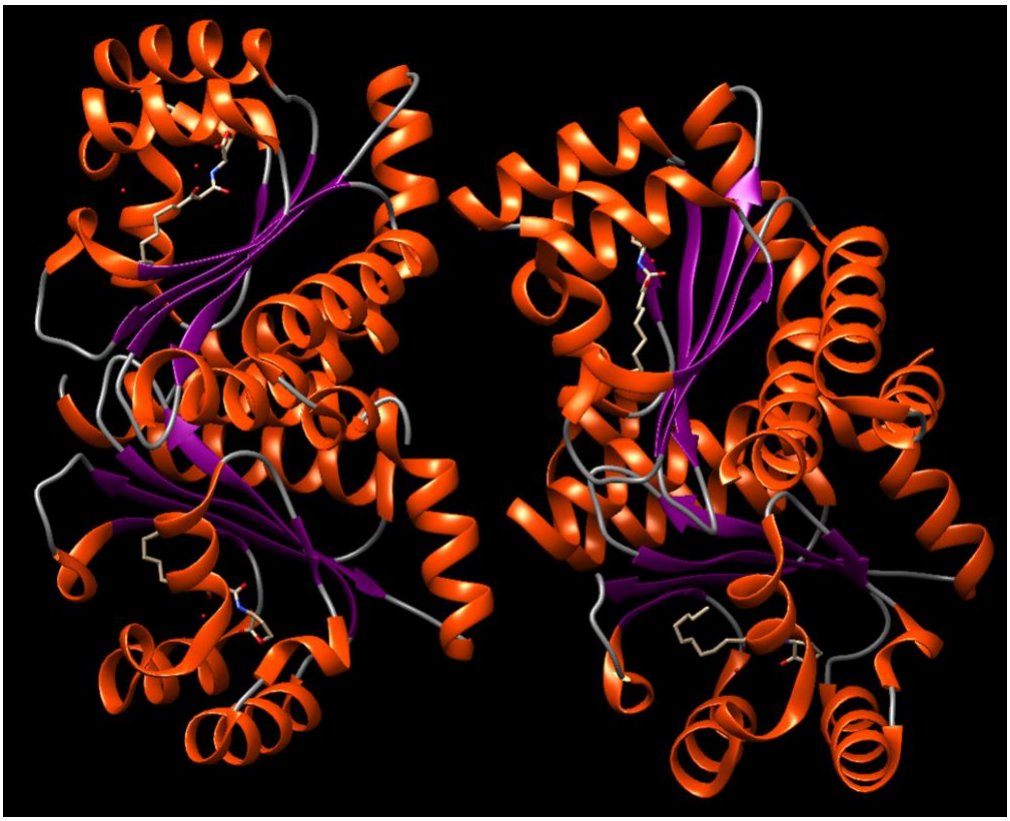
The algorithm that is used to predict binding affinity can be summarized into the equation

$$\Delta E(binding) = \sum(ij)\sum(ji)\left(\left(\frac{X}{r^6} - \frac{Y}{r^{12}}\right) + \frac{332 q_{ji} q_{ij}}{A r^2}\right)$$

The i and j symbols represent the pair-wise interaction of all the receptor atoms with all the ligand atoms. The X and Y terms are related to the van der Waal interaction radii for individual atoms. The A parameter represents the dielectric component of electrostatic interactions. In solvent, electric fields cannot strongly permeate, however in a vacuum or the inside of a protein, individual electric dipoles of atoms can have a significant effect on binding energy. In the current experiment, the dielectric factor was held constant at a value of 4. It should be noted that although the Lennard-Jones 6-12 potential was applied, this parameter can be adjusted based on individual systems or experiments.

Preparation of Receptor for Virtual Screen

In order to prepare the pdb format file of the LasR crystal structure for DOCKing, Chimera was used to modify the structure. The co-crystallized ligand was individually saved for DOCKing later, then deleted. Hydrogen atoms were added to heavy atoms and water molecules were deleted. Only one monomer of the LasR tetramer was saved for DOCKing, the three other subunits were deleted. The full tetramer is shown below



The charges on the remaining monomer were calculated using the FF14SB force field via the program antechamber. The active site of the protein was then used as a reference for the “anchor and grow” DOCK protocol by selecting spheres within 8 angstroms of the active site for ligand sampling. In retrospect this was a somewhat flawed approach since large molecules occupied solvent accessible portions of the protein, however the data was still highly relevant to the potential binding modes of molecules. Once an appropriate box was selected to enclose the binding spheres of interest, the grid potential of the protein was calculated. The grid spacing, or rather “grid resolution”, was set to .4 angstroms, indicating the accuracy with which the protein contribution to binding was calculated. Hypothetically a smaller grid spacing would indicate a more accurate calculation, however the increase in accuracy is not significant enough to warrant the extra calculation time.

Preparing Ligand Structures for DOCKing

The ligand of the native agonist was treated as a control for the expected binding energies of experimental compounds. The structure of the native agonist, an acylhomoserine lactone, was first loaded in chimera then refined by addition of hydrogens and charges via antechamber. Once docked, its predicted binding energy of

50 kcal mol^{-1} was treated as the minimum for any predicted inhibitors to function effectively. The structures of the natural products were all obtained through the ZINC database, whose files are prepared for immediate DOCKing calculations. The databases sampled are listed below.

All DOCKing calculations were run on the LI RED supercomputing cluster.

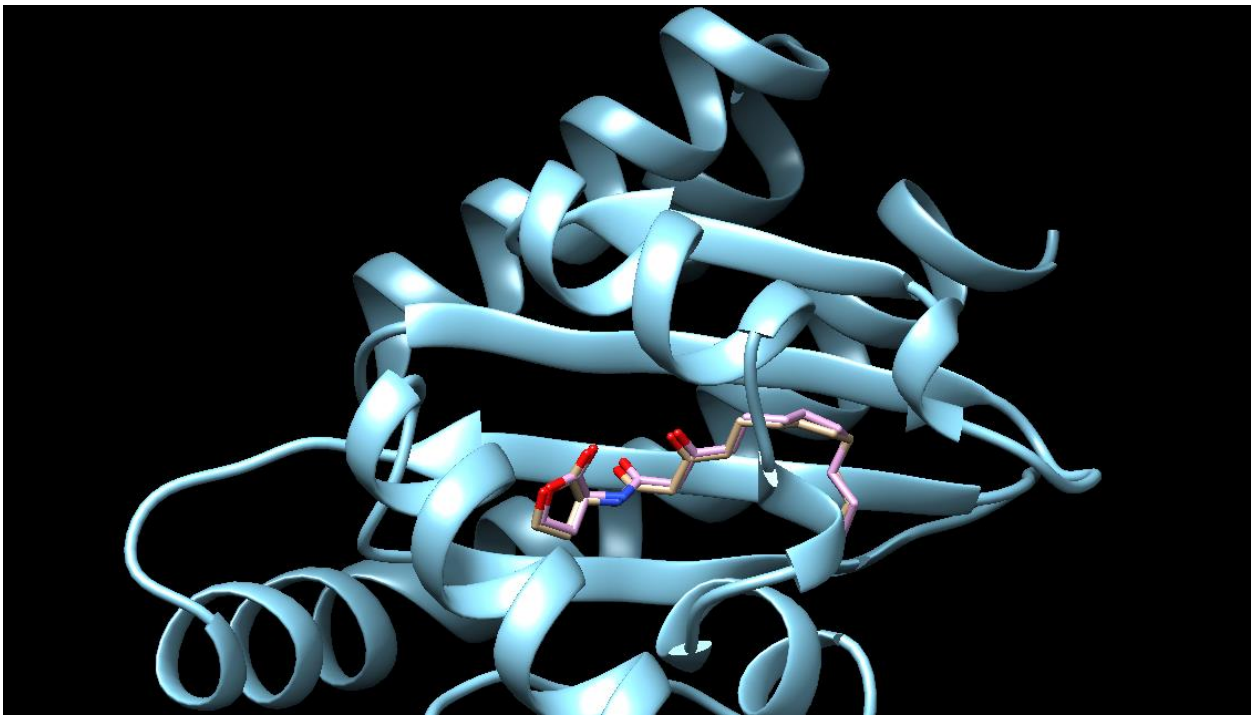
Vendor	Name	Purchasability	ZINC Entries
AfroDb African Natural Products	AfroDb Natural Products	Collaborations Only	885

Vendor	Name	Purchasability	ZINC Entries
ICC INDOFINE	Indofine Natural Products	Items In Stock	144
	Princeton NP	Items In Stock	14084
TCM DATABASE @TAIWAN http://tcm.cmu.edu.tw YC Lab	TCM Database @ Taiwan	Collaborations Only	36043
UEFS.br UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA	UEFS Natural Products	Collaborations Only	473

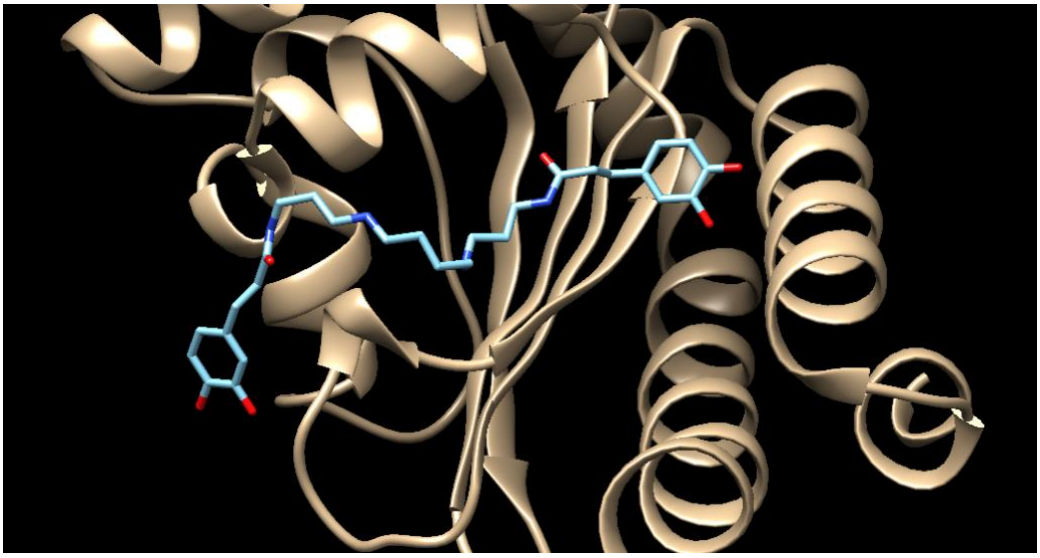
It should be noted that only half of the AfroDB molecule library was sampled.

Results and Discussion:

In order to confirm the robustness of DOCKing on the LasR system, the native ligand was docked back into the active site. The most energetically favorable pose had an RMSD less than 2 angstroms from the X-ray crystal structure under flexible receptor residues. This supports the notion that DOCK is robust enough to sample the active site of LasR effectively. The footprint reference plots demonstrate that hydrogen bonding interactions resemble the empirical results from the crystallization experiment.

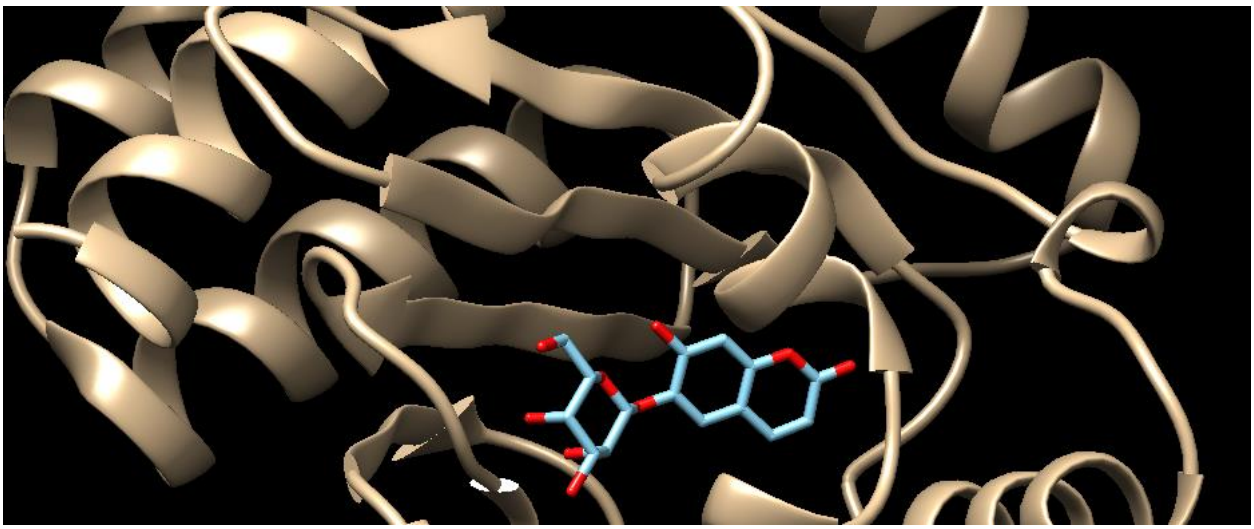


The virtual screen results have elucidated a major caveat to future design of inhibitors. Many compounds, frequently similar to peptides in structure, have highly favorable interactions with the surface of the LasR binding site, as shown below.

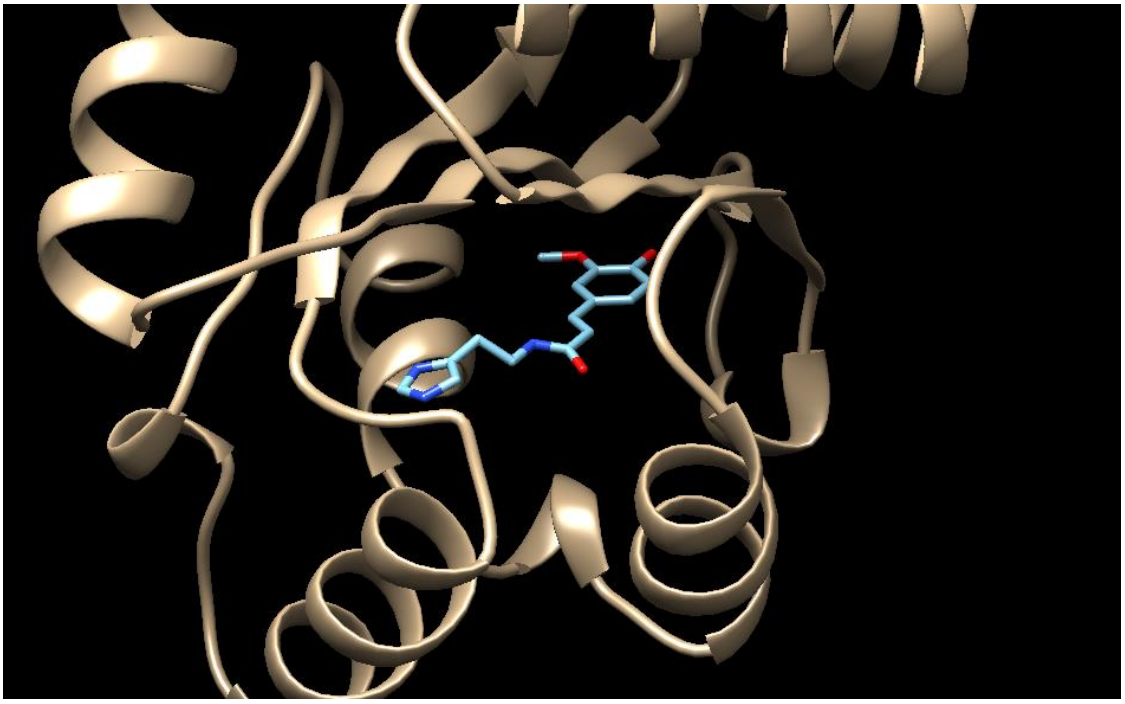


It has been suggested that the DOCK algorithm arbitrarily favors compounds with larger van der Waal surface areas, and the LasR protein model is no exception. However, many modern therapeutic agents are antibodies or peptidomimetics which have large surface areas contacting the protein of interest. This idea may legitimize the current molecule as a drug candidate if some of the rotatable bonds can be restricted by double bonds or some other chemical modification.

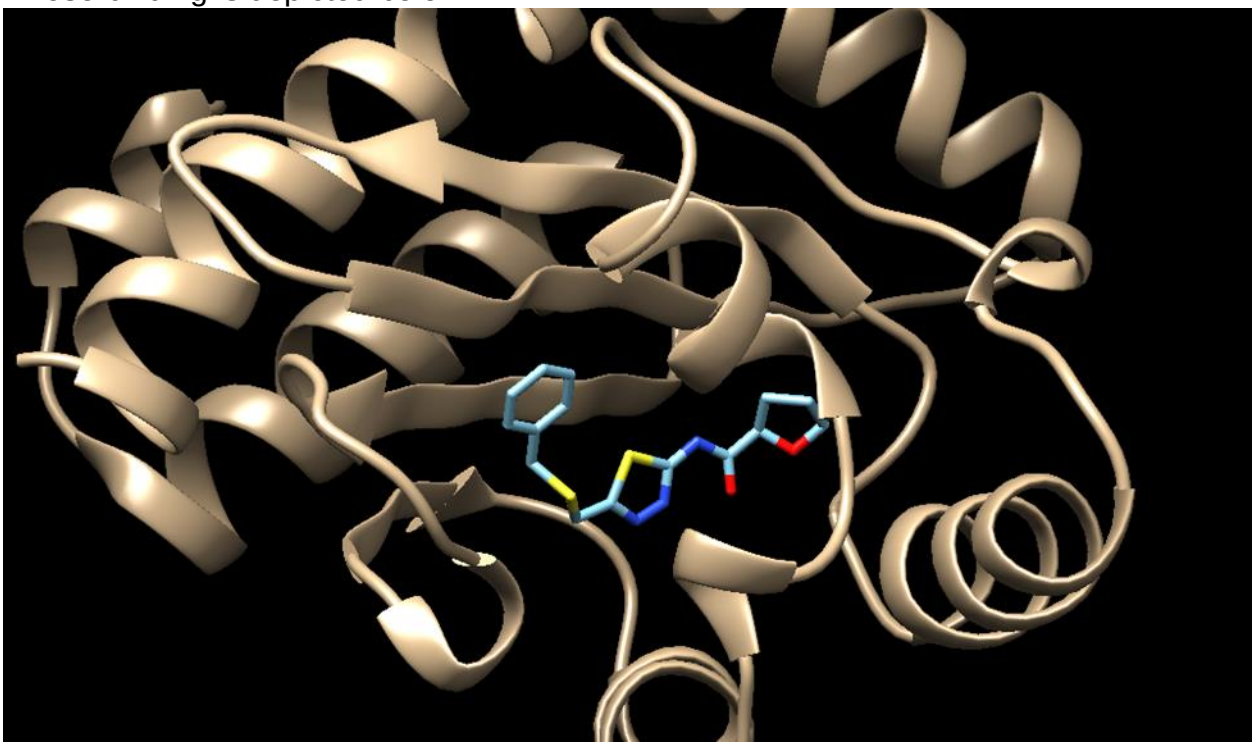
Nevertheless, by simple observation of the native agonist of LasR, it is clear that a competitive inhibitor will have to have a segment that fits well into the hydrophobic pocket adjacent the beta sheet segment of LasR. Interestingly, the molecule depicted above (ZINC ID 13513540) seems to favor formation of a gate which blocks the active site from exposure to solvent. This may hypothetically make it an allosteric or non-competitive inhibitor. Nevertheless, there would almost certainly be issues in absorption or distribution in the body. More conventional examples of inhibitors include the compound below (ZINC ID 03894276)



This compound has elements of both the native agonist as well as diketopiperazine structure. The cyclic ester forms the key hydrogen bond to aspartate 73 in the active site, however it differs in the alcohol groups which protrude off the bicyclic ring. The overall geometry of the molecule renders it a promising lead for an inhibitor. Another promising lead is the compound below (ZINC ID 13397330)



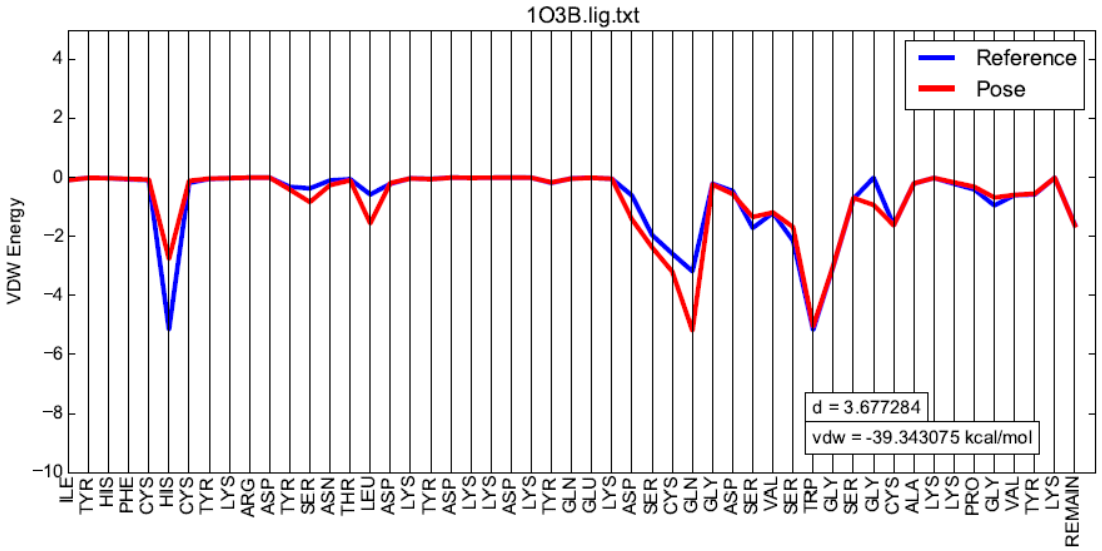
This molecule bears similarity to both the diketopiperazine and cinnamaldehyde structures. The tail portion of the molecule is very different from the hydrocarbon tail of the native agonist. The tail has a hydrophilic amide group flanked by hydrocarbons, which undoubtedly form key hydrogen bonds to stabilize the receptor ligand complex. A molecule which departs significantly from known inhibitor structure is ZINC ID06704819, whose binding is depicted below.

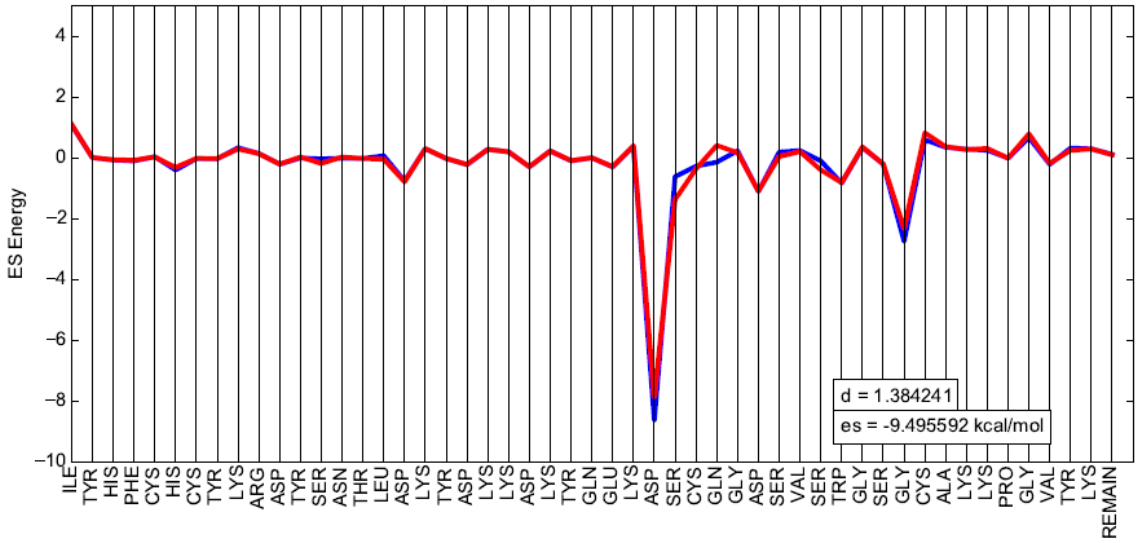


Three ring structures are bonded by thioether and amide groups, and a cyclopropane moiety forms the key nonpolar interactions with the hydrophobic beta strand. The furan group lacks the carbonyl present in the native agonist, however otherwise the furan ring portion is identical to the homoserine lactone. The unique structure of the compounds central ring may be problematic for synthesis, however that remains to be seen.

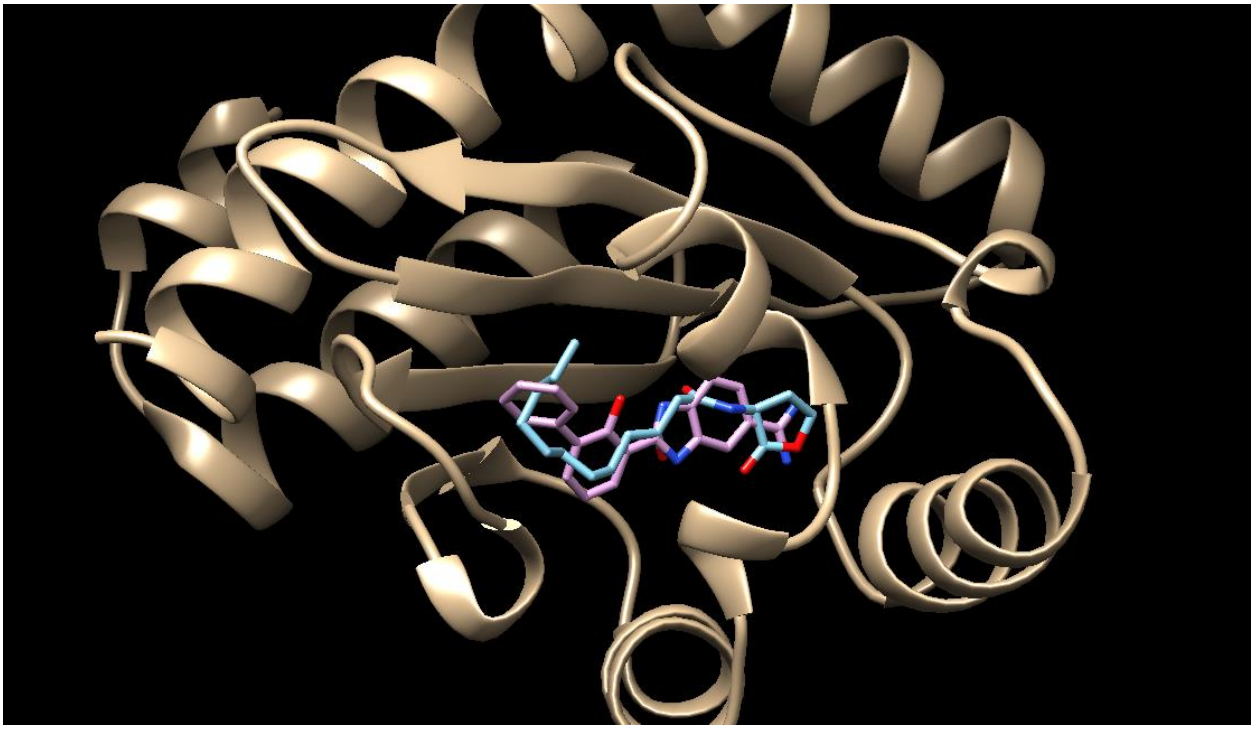
An important question in any drug design project is understanding what interactions or motifs cause agonistic or antagonistic effects on a protein. One computational method of attempting to predict this effect is footprint scoring algorithm of UCSF DOCK 6.6, which decomposes binding energy of a ligand based on which residue the ligand is interacting with. Since the LasR system is inhibited by several analogs of the native agonist such as the brominated furanone, it is reasonable to predict an inhibitor will have a similar footprint to the native agonist. This corresponds to hydrogen bonding with the key aspartate 73 residue and van der Waal packing with valine 76. If this motif turns out to be true, it can help accelerate design of effective quorum sensing inhibitors.

Several compounds screened from ligands in the RCSB have footprints that overlap exceptionally well with the native ligand. One is an inhibitor of beta trypsin, a serine protease. The ligand is from crystal structure 1O3B of the protein data bank with the IUPAC name 3-{5-[Amino (Imino) Methyl]-1H-Benzimidazol- 2-yl}-1,1'-Biphenyl-2-olate. The compound makes a hydrogen bond with a backbone glycine similarly to the native agonist adjacent to the serine 129 residue.



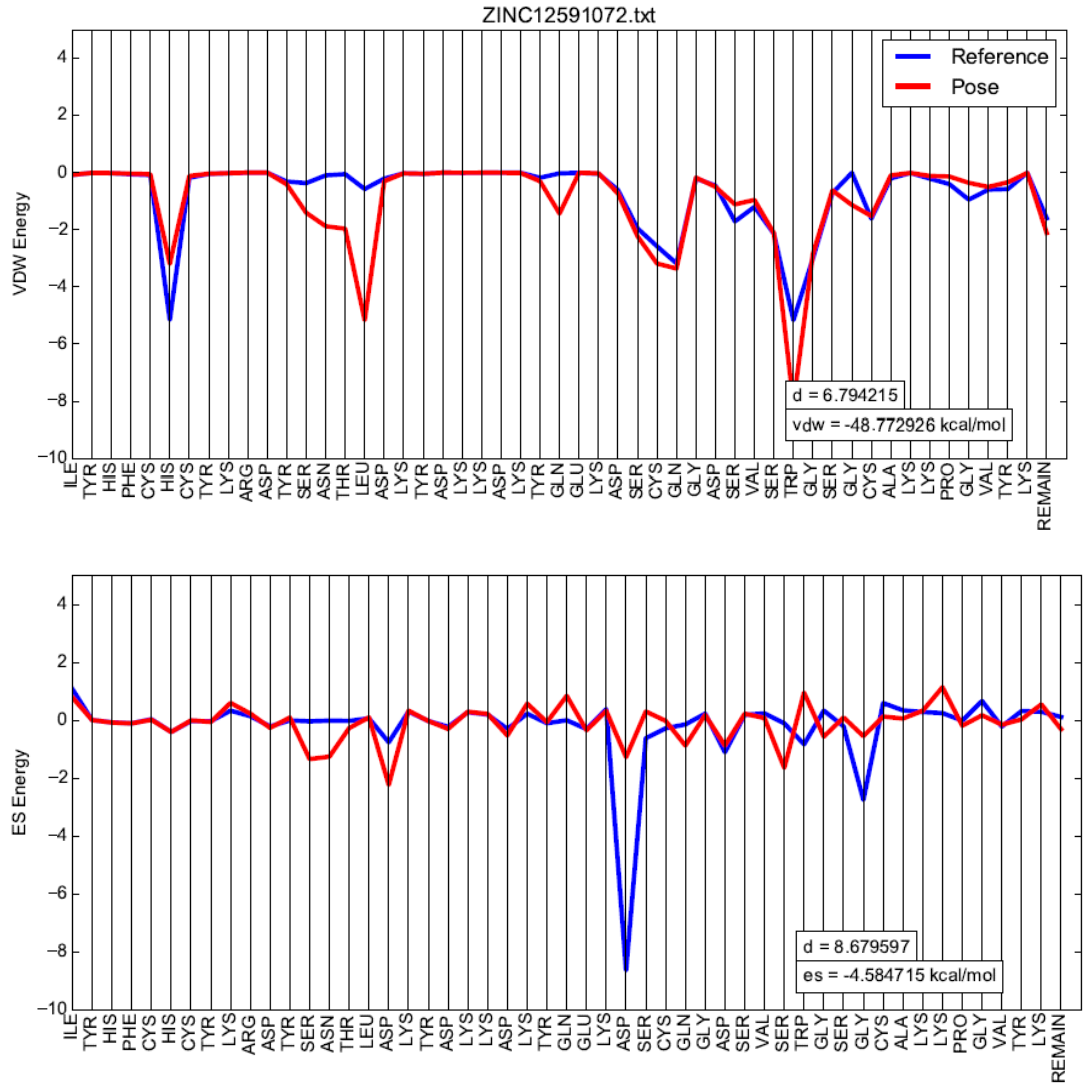


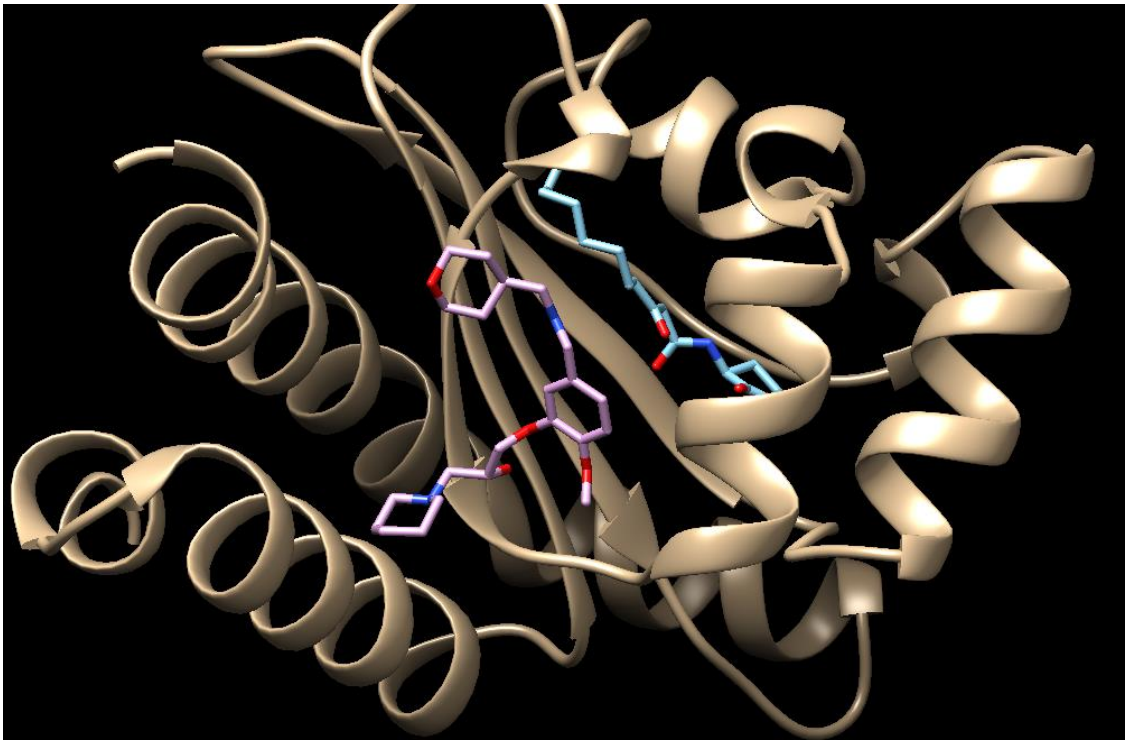
Visual analysis of the binding pose of the native ligand and the protease inhibitor corroborate well with the footprint score result. For this molecule to be a LasR inhibitor, it would need modifications that reduce its specificity to beta trypsin in favor of LasR. The functional groups on the scaffold suggest this is plausible, especially with the reactive imine group.



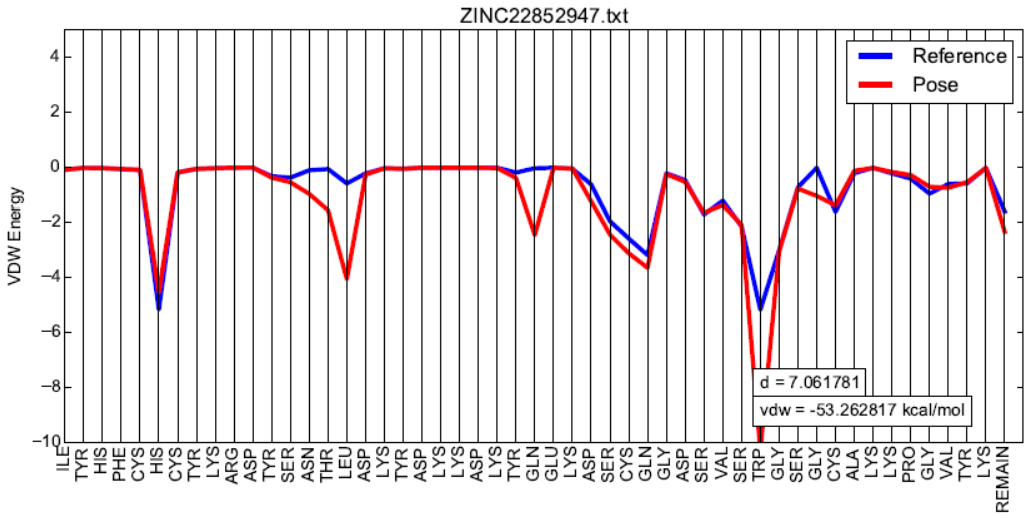
A molecule with novel interactions to the native agonist is ZINC ID 12591072. However, visual analysis of the binding mode revealed the compound forms a gate around the binding site of LasR. This explains why the electrostatic decomposition appears noisy. The fact that this molecule was favored based on footprint scoring is an

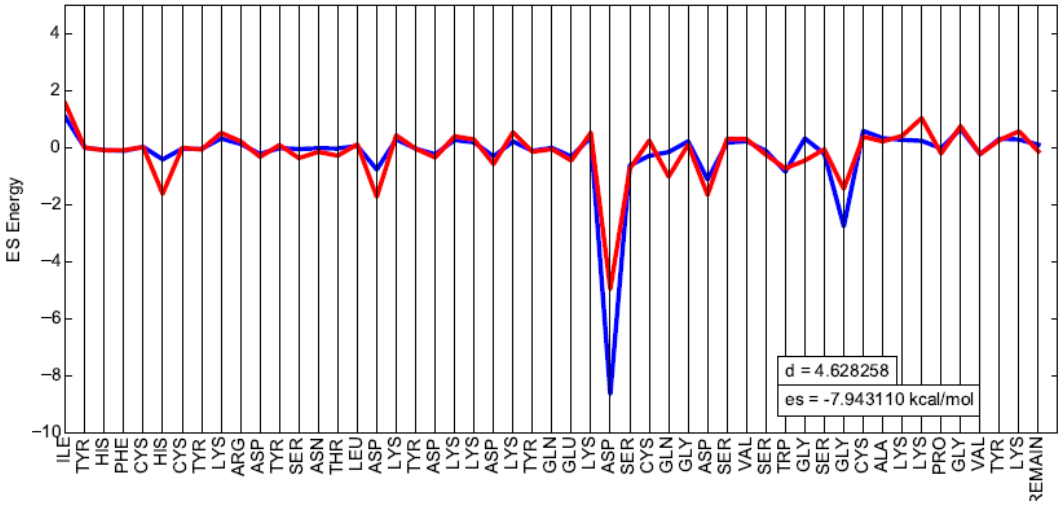
artifact of misguided sampling of the LasR active site. Still, the van der Waal attractions overlap well between the two poses.

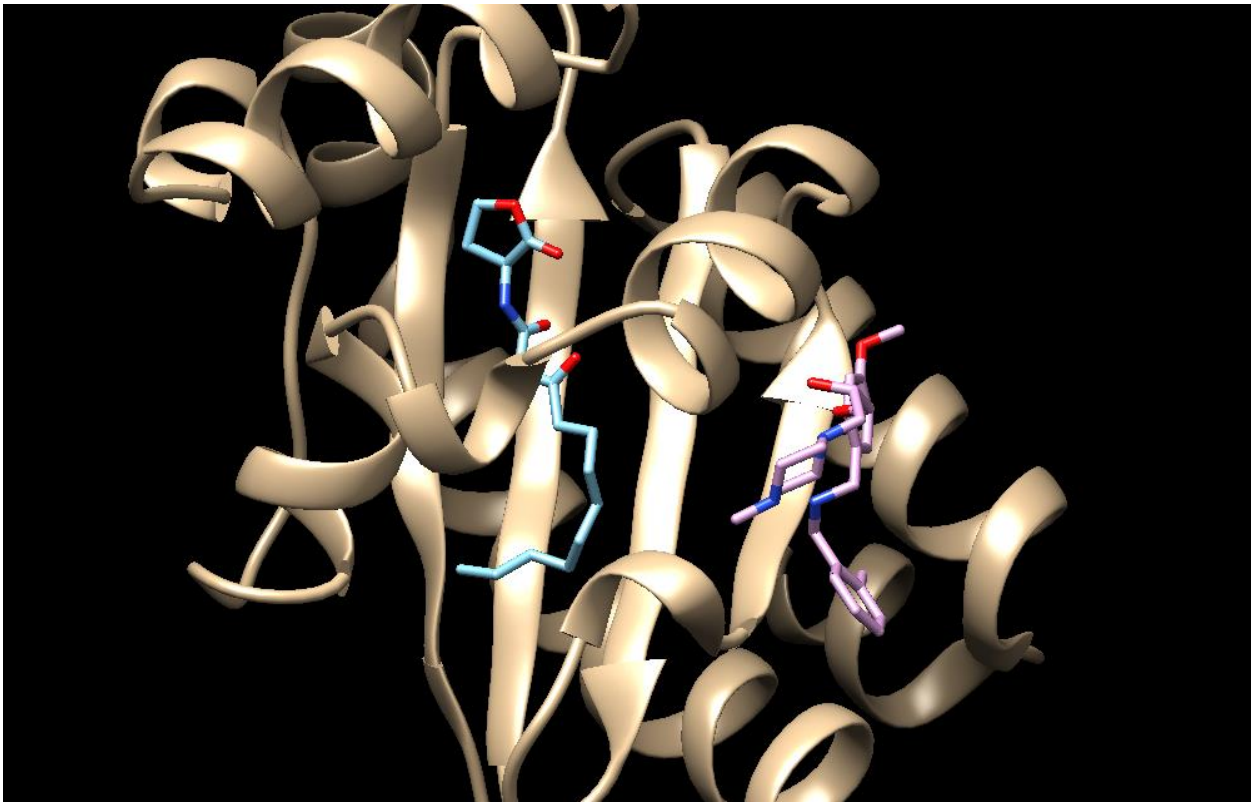




A very similar artifact occurs with ZINC ID 22852947, however with a different segment of the binding pocket. These results indicate that if sampling is not well guided in the initial docking experiment the errors will propagate within footprint scoring. Equivalently, one could argue that if the molecules being sampled are too large, they will dock outside the binding site anyway due to the repulsive energy term of the non-polar interactions.


ZINC22852947.txt
Reference
Pose
VDW Energy
d = 7.061781
vdw = -53.262817 kcal/mol
ILE TYR HIS PHE CYS HIS CYS TYR LYS ARG ASP TYR SER ASN THR LEU ASP LYS TYR ASP LYS LYS ASP LYS TYR GLN GLU LYS ASP SER CYS GLN GLY ASP SER VAL SER TRP GLY SER GLY CYS ALA LYS LYS PRO GLY VAL TYR LYS REMAIN


ES Energy
d = 4.628258
es = -7.943110 kcal/mol
ILE TYR HIS PHE CYS HIS CYS TYR LYS ARG ASP TYR SER ASN THR LEU ASP LYS TYR ASP LYS LYS ASP LYS TYR GLN GLU LYS ASP SER CYS GLN GLY ASP SER VAL SER TRP GLY SER GLY CYS ALA LYS LYS PRO GLY VAL TYR LYS REMAIN



Conclusions:

The results of this experiment have supported existing data on inhibitor design, and yet also provided novel directions for experimentation. Since molecules must have a certain overall geometry and size to fit well in the LasR active site, future libraries for virtual screening should be organized by features such as ligand volume, or even the ratio of length to width. The natural product libraries revealed many solid, rational leads, however the methodology would be exponentially better if the size of the molecules was controlled for.

An interesting pattern of the current experiment is the widespread inefficacy of natural products to bind to the LasR active site. This is most likely a product of natural selection, which would favor a pathogen like *P. Aeruginosa* to have a robust signaling system. Relatively small, drug like molecules with relatively linear geometries are predicted to have the best specificity for the active site. Thus, future work should apply libraries of molecules that have drug like properties, in particular several hydrogen bond acceptors akin to the furanone ring of the native agonist.

Two chemical motifs are consistent across the LasR system; an inhibitor must have distinct amphiphilic and hydrophobic segments. This fact also helps explain the cytotoxicity of the existing synthetic inhibitors; many metabolic intermediates have a similar motif, and off-target binding is a major complication. Thus, it seems the greater hurdle of designing a drug that can effectively inhibit quorum sensing via LasR is specificity for the target, since there is no shortage of molecules with varying degrees of

affinity for LasR. To this end, future work should focus on the ADME of lead compounds.

Acknowledgements:

I would like to thanks Robert Rizzo, Yuchen Zhou, Courtney Singleton, and Jiaye Guo for guidance on the project.

1. Bottomley, M.J., et al., *Molecular insights into quorum sensing in the human pathogen Pseudomonas aeruginosa from the structure of the virulence regulator LasR bound to its autoinducer.* J Biol Chem, 2007. **282**(18): p. 13592-600.
2. Schuster, M., M.L. Urbanowski, and E.P. Greenberg, *Promoter specificity in Pseudomonas aeruginosa quorum sensing revealed by DNA binding of purified LasR.* Proc Natl Acad Sci U S A, 2004. **101**(45): p. 15833-9.
3. Hentzer, M., et al., *Attenuation of Pseudomonas aeruginosa virulence by quorum sensing inhibitors.* EMBO J, 2003. **22**(15): p. 3803-15.
4. Ilangovan, A., et al., *Structural basis for native agonist and synthetic inhibitor recognition by the Pseudomonas aeruginosa quorum sensing regulator PqsR (MvfR).* PLoS Pathog, 2013. **9**(7): p. e1003508.
5. Zou, Y. and S.K. Nair, *Molecular basis for the recognition of structurally distinct autoinducer mimics by the Pseudomonas aeruginosa LasR quorum-sensing signaling receptor.* Chem Biol, 2009. **16**(9): p. 961-70.
6. Chun, C.K., et al., *Inactivation of a Pseudomonas aeruginosa quorum-sensing signal by human airway epithelia.* Proc Natl Acad Sci U S A, 2004. **101**(10): p. 3587-90.
7. Welsh, M.A., et al., *Small molecule disruption of quorum sensing cross-regulation in pseudomonas aeruginosa causes major and unexpected alterations to virulence phenotypes.* J Am Chem Soc, 2015. **137**(4): p. 1510-9.
8. Rybtke, M., et al., *Pseudomonas aeruginosa Biofilm Infections: Community Structure, Antimicrobial Tolerance and Immune Response.* J Mol Biol, 2015. **427**(23): p. 3628-45.
9. Brackman, G., et al., *Quorum sensing inhibitors increase the susceptibility of bacterial biofilms to antibiotics in vitro and in vivo.* Antimicrob Agents Chemother, 2011. **55**(6): p. 2655-61.
10. Anand, R., N. Rai, and M. Thattai, *Interactions among quorum sensing inhibitors.* PLoS One, 2013. **8**(4): p. e62254.
11. Yang, L. and M. Givskov, *Chemical Biology Strategies for Biofilm Control.* Microbiol Spectr, 2015. **3**(4).
12. Jimenez, P.N., et al., *The multiple signaling systems regulating virulence in Pseudomonas aeruginosa.* Microbiol Mol Biol Rev, 2012. **76**(1): p. 46-65.
13. Eibergen, N.R., et al., *Potent and Selective Modulation of the RhlR Quorum Sensing Receptor by Using Non-native Ligands: An Emerging Target for Virulence Control in Pseudomonas aeruginosa.* Chembiochem, 2015. **16**(16): p. 2348-56.
14. Lidor, O., et al., *Mechanistic analysis of a synthetic inhibitor of the Pseudomonas aeruginosa LasI quorum-sensing signal synthase.* Sci Rep, 2015. **5**: p. 16569.
15. Kumar, L., et al., *Zingerone silences quorum sensing and attenuates virulence of Pseudomonas aeruginosa.* Fitoterapia, 2015. **102**: p. 84-95.
16. Moore, J.D., et al., *A Comparative Analysis of Synthetic Quorum Sensing Modulators in Pseudomonas aeruginosa: New Insights into Mechanism, Active Efflux Susceptibility, Phenotypic Response, and Next-Generation Ligand Design.* J Am Chem Soc, 2015. **137**(46): p. 14626-39.
17. Irwin, J.J. and B.K. Shoichet, *ZINC--a free database of commercially available compounds for virtual screening.* J Chem Inf Model, 2005. **45**(1): p. 177-82.