



**INSIGHTS INTO THE BINDING AND CATALYTIC MECHANISMS OF  
*BACILLUS THURINGIENSIS* LACTONASE AND INTERACTIONS  
OF A PTEN-BINDING INHIBITORY PEPTIDE**

---

A Dissertation Presented to  
the Faculty of the Department of Biology and Biochemistry  
University of Houston

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

---

By  
Marc Neil Charendoff  
May 2013

**INSIGHTS INTO THE BINDING AND CATALYTIC MECHANISMS OF  
*BACILLUS THURINGIENSIS* LACTONASE AND INTERACTIONS  
OF A PTEN-BINDING INHIBITORY PEPTIDE**

---

**Marc Neil Charendoff**

APPROVED

---

**Dr. James M. Briggs, Chairman**

---

**Dr. Richard C. Willson**

---

**Dr. Eric R. Bittner**

---

**Dr. William R. Widger**

---

**Dr. Hye-Jeong Yeo**

---

**Dr. Dan E. Wells, Dean**  
**College of Natural Sciences and Mathematics**

## ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. James Briggs, for his guidance and extreme patience during this endeavor. It is of no surprise to me that it is only toward the end of my work that I see the wisdom of the advice he has provided. It is also through him that I have made the slow transition from thinking as an engineer to thinking as a scientist. To that end, though I cannot claim complete success, I have at least learned to ask a much different battery of questions when confronted with a problem.

I would also like to thank Dr. William Widger for his mentorship. Immersing oneself into this course of study is indeed diving into the unknown with only a limited set of tools. Under these constraints one is asked to make sense of at least a small part of his murky surroundings. Only too soon my early exuberance and confidence were rapidly tempered. I more than appreciate Dr. Widger's ability and willingness to reassure me that "all is normal" despite the slow (and certainly not guaranteed) growth of knowledge.

The Briggs research group has been a fantastic collection of folks without whom I would barely have been able to begin. To them I wish only the best in all the corners of the world they have now gone. I also would be remiss without thanking Drs. Lisa and Jerrick Juliette, Dr. Joy Mendoza, Dr. Alan Sopchik, Dr. Nelson Quiros, and Dr. Steve Alexander for the many discussions and support as I have proceeded along this path. To the doctors and researchers at Texas Children's Hospital, thank you for the inspiration and love extended to my family and me.

Lastly, to my wife Teri, and my daughter Chloe, I cannot say enough. They were with me when I made the decision to undertake this effort and have remained my closest supporters.



**INSIGHTS INTO THE BINDING AND CATALYTIC MECHANISMS OF  
*BACILLUS THURINGIENSIS* LACTONASE AND INTERACTIONS  
OF A PTEN-BINDING INHIBITORY PEPTIDE**

---

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Biology and Biochemistry

University of Houston

---

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

By

Marc Neil Charendoff

May 2013

## ABSTRACT

The lactonase enzyme (AiiA) produced by *Bacillus thuringiensis* (*B. thuringiensis*) serves to degrade autoinducer-1 (AI-1) signaling molecules in what is an evolved mechanism by which allows *B. thuringiensis* to better compete with other bacteria. Bioassays have been previously performed to determine whether the AI-1 aliphatic tail lengths have any effect on AiiA's bioactivity; however, data to date are conflicting. To investigate these questions, multiple molecular dynamics simulations were performed across a family of seven acylated homoserine lactones (AHL) along with their associated intermediate and product states. Distance analyses and interaction energy analyses were performed to investigate current bioassay data. Our simulations are consistent with experimental studies showing that AiiA degrades AHLs in a tail length independent manner. Also, a proposed putative oxyanion hole function of Y194 toward the substrate is not observed in any of the reactant or product state simulation trajectories. However, Y194 does seem to show efficacy in stabilizing the intermediate state. Last, we argue through ionization state analyses, that proton-shuttling necessary for catalytic activity is possibly mediated by both water- and substrate-based, intra-molecular proton transfer. Based on this argument, an alternate catalytic mechanism is proposed.

Drug dependence, or addiction ultimately described as the maladaptation of a neurochemical reward pathway that is normally used to reinforce behaviors that promote an organism's survival and/or propagation (e.g., high calorie food seeking and sexual drive). These pathways are mediated via dopaminergic VTA-NAc neuronal networks. The molecular basis for this interaction has been shown to rely on a 5HT-2cR:PTEN interaction via the 3L4F-F1 fragment. The nature of the interaction is unknown. In this

work a series of REMD runs are designed to generate candidate conformations that might exist under physiological conditions. These candidates are then subjected to exhaustive six-dimensional, coarse-grained docking simulations against the PTEN protein to produce a myriad of potential docking sites. From these dockings the top scoring poses are observed to point to a single binding site that in turn lend themselves to spectroscopic validation.

## CONTENTS

1. General Background .....	1
1.1 Bacterial communications.....	1
1.1.1 Quorum signaling and sensing.....	1
1.1.2 Model system <i>Pseudomonas aeruginosa</i> .....	3
1.1.3 Quorum signal-induced phenotype expression (biofilms).....	8
1.1.4 Mitigation of quorum sensing networks .....	10
1.1.5 <i>Bacillus thuringiensis</i> lactonase .....	11
1.1.6 Scope of research .....	21
1.2 Biochemistry of addictive behaviors .....	22
1.2.1 Serotonin, dopamine and addiction.....	22
1.2.2 5-HT <sub>2c</sub> Receptor.....	29
1.2.3 PTEN.....	29
1.2.4 5-HT <sub>2c</sub> R:PTEN interaction .....	37
1.2.5 Scope of research .....	38
1.3 Theoretical background of methods.....	39
1.3.1 pK <sub>a</sub> prediction.....	39
1.3.2 Molecular dynamics.....	43
1.3.3 Replica exchange molecular dynamics .....	48
1.3.4 GRAMM-VINA.....	48
1.3.5 Cluster analysis .....	51
2. Impacts of PKA Perturbations on the Catalytic Mechanism of <i>Pseudomonas</i> <i>aeruginosa</i> Lactonase .....	58
2.1 Computational methods .....	58
2.1.1 Ligand construction .....	58

2.1.2 Molecular dynamics simulations and setup .....	62
2.1.3 Cluster analysis .....	64
2.1.4 pKa predictions and electrostatic mapping .....	64
2.2 Results and discussion .....	65
2.2.1 pKa of ASP108, ASP 109, TYR194, and HIS residues .....	65
2.2.2 Electrostatic mapping.....	65
3. Impact of Acyl-homoserine Lactone Tail Length on Catalytic Competency and Binding/Interaction Energy.....	72
3.1 Computational methods .....	72
3.1.1 Distance mapping.....	72
3.1.2 Binding free energy study .....	76
3.1.3 Interaction energy study – Q.....	77
3.1.4 Active site QM snapshots / optimization of a dizinc 2+ ligand field ....	79
3.2 Results and discussion .....	79
3.2.1 Distance mapping – catalytic competence .....	79
3.2.2 Free energy analysis – reactant state and perturbation .....	89
3.2.3 Interaction energy analysis .....	94
3.2.4 Geometry and energy of dinuclear zinc (2+) ligation field .....	96
3.3 Conclusions.....	100
3.4 Future direction.....	104
4. Conformational Analysis of PTEN:3L4F-F1 Complex .....	110
4.1 Introduction.....	110
4.2 Computational methods .....	110
4.2.1 Molecular dynamics simulations – TIGER2.....	110
4.2.2 Clustering by dPCA .....	111

4.2.3 Coarse, blind docking with GRAMM.....	111
4.3 Results and discussion .....	112
4.3.1 Molecular dynamics simulations .....	112
4.3.2 dPCA analysis.....	121
4.3.3 PTEN docking hot spots and CD validation .....	125
4.4 Conclusion .....	129
4.5 Future direction.....	129
References.....	132

## LIST OF TABLES

Table	Page
1. AHL degradation bioassay comparison .....	20
2. VINA parameters .....	50
3. PROPKA pKa analysis .....	67
4. UHBD pKa analysis.....	68
5. UHBD pKa analysis - no ligand in protein .....	69
6. Reactant state distance analysis .....	82
7. Reactant state distance analysis continued .....	83
8. Intermediate state distance analysis .....	85
9. Intermediate state distance analysis continued .....	86
10. Product state distance analysis.....	87
11. Product state distance analysis continued .....	88
12. Reactant state binding free energies (kcal/mol).....	91
13. LIE $\Delta G$ (kcal/mol) values of signal ligands versus their reduced forms.....	93
14. van der Waals interaction energies (kcal/mol).....	95
15. Replica time at temperature .....	115

## LIST OF FIGURES

Figure	Page
1. Lux system quorum sensing.....	2
2. <i>Pseudomonas aeruginosa</i> quorum signaling molecules .....	5
3. <i>Pseudomonas aeruginosa</i> quorum signaling scheme .....	6
4. Biofilm formation .....	9
5. Lactonase with bound product .....	12
6. Lactonase secondary structure .....	13
7. Lactonase active site with ligating residues.....	14
8. Lactonase active site - alternate view .....	15
9. Oxygen bridging zincs in lactonase active site .....	17
10. Dopamine biosynthetic pathway.....	24
11. Serotonin biosynthetic pathway .....	26
12. Brain map of relevant regions to dopamine and serotonin signaling.....	28
13. PTEN secondary structure and volume.....	31
14. PTEN domains .....	32
15. PTEN with critical catalytic structures .....	33
16. PTEN with exploded view of catalytic site.....	34
17. PTEN and PIP3 signaling process .....	36
18. Molecular dynamics sample code .....	47
19. AHL X, Y, and Z tail positions looking down into catalytic site .....	60
20. AHL Z position threaded into available side pocket .....	61
21. 3DHB molecular dynamics system.....	63
22. Lactonase electrostatic mapping – front .....	70



23. Lactonase electrostatic mapping – back .....	71
24. Measured distances for AHL reactant state .....	73
25. Measured distances for AHL intermediate state .....	74
26. Measured distances for AHL product state .....	75
27. Molecular dynamics system setup for Q .....	78
28. Hexacoordinated zinc 2+ ions at 4.7 angstroms .....	98
29. Pentacoordinated zinc 2+ ions at 3.5 angstroms .....	99
30. Proposed mechanism for lactonase catalyzed AHL ring-opening .....	103
31. Artificial enzyme surface for AHL ring-opening or peptide hydrolysis .....	108
32. Replica exchange vs. time .....	114
33. 600K Replica secondary structure data .....	117
34. 476K Replica secondary structure data .....	118
35. 378K Replica secondary structure data .....	119
36. 300K Replica secondary structure data .....	120
37. 300K replica dPCA analysis .....	122
38. Helical conformations per dPCA analysis .....	123
39. Turn conformations per dPCA analysis .....	124
40. 3L4F-F1 Helix GRAMM docking scores .....	126
41. 3L4F-F1 Turn GRAMM docking score .....	127
42. GRAMM and CD data .....	128
43. PTEN tryptophan residues .....	131

## LIST OF ABBREVIATIONS

3L4F	Third loop fourth fragment
3L4F-F1	Third loop fourth fragment – first sub-fragment
5HT	5-hydroxytryptamine (serotonin)
5HT-2cR	5-hydroxytryptamine receptor
AHL	n-acylhomoserine lactone
AiiA	Autoinducer inactivation enzyme
<i>B. thuringiensis</i>	<i>Bacillus thuringiensis</i>
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CD	Circular dichroism
DFT	Density functional theory
dPCA	Dihedral principal components analysis
GPCR	G protein coupled receptor
GROMACS	Groningen machine for chemical simulations
HSL	Homoserine lactone
MD	Molecular dynamics
NAMD	Not another molecular dynamics program
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
QM	Quantum mechanics
QSS	Quorum signaling/sensing
REMD	Replica exchange molecular dynamics

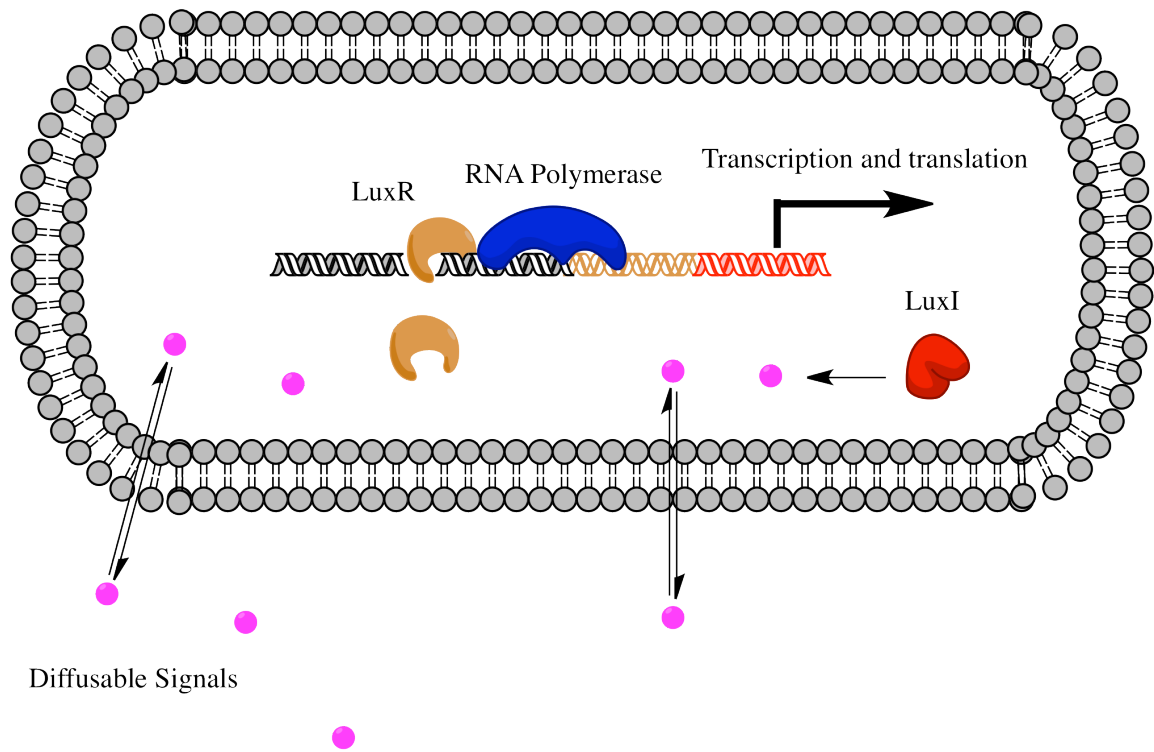
## CHAPTER 1: GENERAL BACKGROUND

### 1.1 Bacterial communications

#### 1.1.1 Quorum signaling and sensing

Quorum signaling/sensing (QSS) is an intercellular communication mechanism where chemical signals allow bacteria to detect their local population and change their gene expression, so as to competitively optimize their behavior in their local environment. These signaling molecules are synthesized within the bacteria and diffuse out to the extracellular environment. As the concentration of signaling molecules in the environment increases beyond a threshold level (the “quorum” concentration), diffusion of the signaling molecules back into the bacteria promotes a change in gene expression mediated by quorum signal-bound transcription factors. Gram-positive and Gram-negative bacteria display this phenomenon through the utilization of different types of diffusible chemical species. Both Gram-positive and Gram-negative bacteria synthesize families of ribose-like molecules that allow for interspecies communication known as autoinducer-2 (*Vibrio fischeri*).<sup>1</sup> Gram-positive bacteria employ signaling molecules that are comprised of small peptides whereas Gram-negative bacteria largely use families of acylated homoserine lactones autoinducer-1 (*Vibrio fischeri*).<sup>1</sup> Sensing networks involving acylated homoserine lactones (AHLs) are well studied and have provided information regarding the importance of QSS to both intra and interspecies communications. An example of such a system is seen in the LuxR/LuxI QSS as shown in Figure 1.

**Figure 1. Lux system quorum sensing**



---

*Vibrio fischeri* QSS archetype utilizing LuxI/R synthase and receptor.

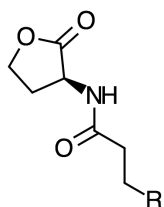
The roles of QSS in interspecies communications are complex, in fact, it has been posited that if one thinks of a collection of organisms as an organ existing in balance with its environment, interspecies communications can be seen as analogous to higher life forms where hormones provide for organ system communication.<sup>2</sup> Bacteria, however, exist with their hosts as symbiotes, commensals, or pathogens, all of which are relationships for which QSS has been implicated. Just a few examples of host biological effects mediated by QSS include cell apoptosis, inhibition of lymphocyte production, vasorelaxation, inhibition of sperm viability and function, stimulation of pro-inflammatory mediators, and regulation of plant root development. It is also notable to understand that quorum signals themselves exert effects on their hosts without any kind of cognate sensing system, thus quorum signaling and QSS are not one and the same especially when considering host-guest relationships.

### **1.1.2 Model system *Pseudomonas aeruginosa***

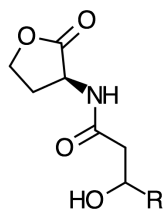
Elucidation of the difference between QSS and host quorum signal sensitivities are provided by numerous studies of the organism *Pseudomonas aeruginosa* (*P. aeruginosa*). *P. aeruginosa* is a Gram-negative bacterium that has a well-studied quorum sensing system whose signaling molecules consist of AHLs. These molecules consist of a homoserine lactone head connected via an amide linkage to a tail with lengths of 4 to 16 carbons. Additionally, they sometimes contain hydroxyl or carbonyl groups at the third carbon atom as shown in Figure 2.<sup>3</sup> There are two AHL-dependent quorum sensing systems in *P. aeruginosa*, *las* and *rhI* that each contain homologs of the LuxR (quorum signal receptor) and LuxI (quorum signal synthase) proteins. The *las* and *rhI* systems direct quorum sensing mediated by N-(3-oxododecanoyl)-L-homoserine lactone (3OC<sub>12</sub>-

HSL) and N-butyryl-L-homoserine lactone (C<sub>4</sub>-HSL), respectively.<sup>4</sup> las and rhl operationally comprise a layered system whereupon activation the LasR-autoinducer complex induces rhIR expression. The LasI autoinducer in turn interferes with binding of autoinducer to RhIR, presumably to ensure sequential activation of the las and rhI cascades.<sup>5</sup> The end result of these cascades is the expression of many different genes including the las-promoted lasA, a protease; lasB, an elastase; toxA, or ExotoxinA; and aprA, an alkaline phosphatase. Additional genes expressed by the rhI system are rpoS, a stationary phase sigma factor; rhlAB, a rhamnosyltransferase involved in synthesis of the biosurfactant/hemolysin rhamnolipid; and lecA which encodes a cytotoxic lectin.<sup>5</sup> The LasI and RhII synthases are, of course, also produced in these autoinduction cycles. In addition to LasR and RhIR, *P. aeruginosa* also possesses a third orphan receptor, QscR that lacks a cognate synthase. This receptor senses 3OC<sub>12</sub>-HSL and serves to control its own as well as other QS genes. A representation of this signaling network is shown in Figure 3.

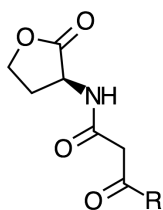
**Figure 2. *Pseudomonas aeruginosa* quorum signaling molecules**



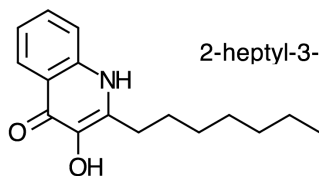
n-acyl homoserine lactone (AHL)



3-hydroxy-AHL



3-oxo-AHL

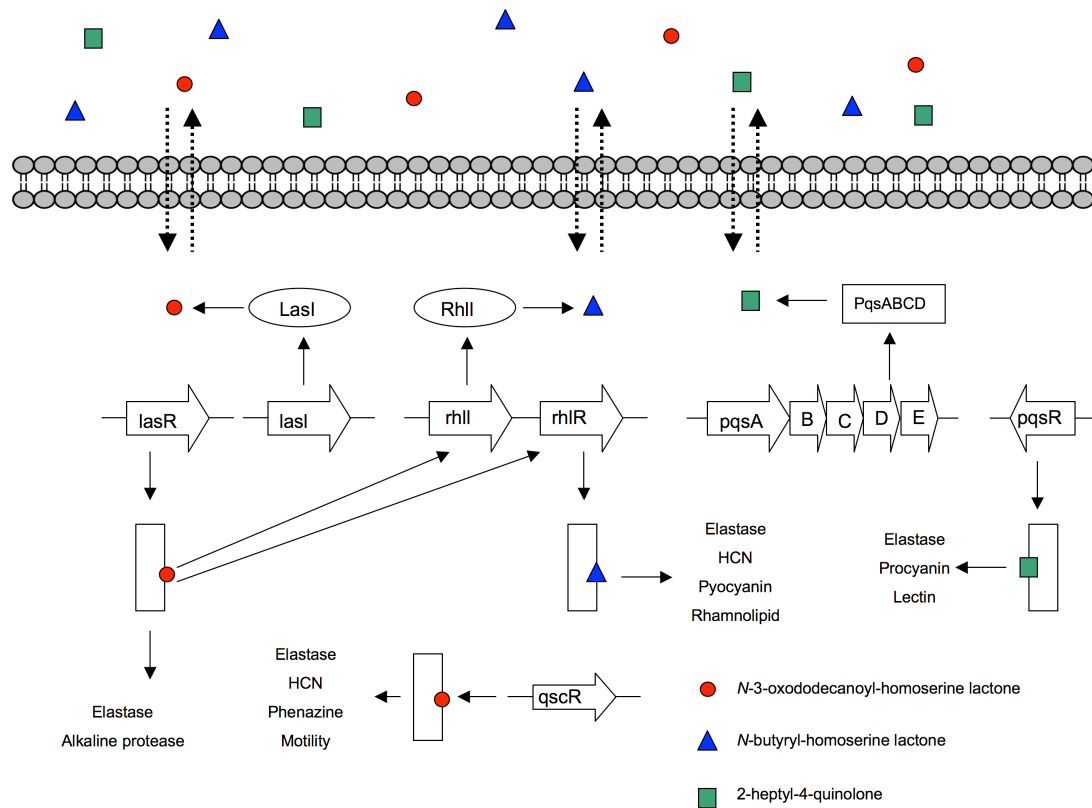


2-heptyl-3-hydroxy-4-quinolone (PQS)

---

These are representative chemical species seen in Gram-negative bacteria. Note presence of alkyl tails across all molecules.

**Figure 3. *Pseudomonas aeruginosa* quorum signaling scheme**



Adapted from Caetano, L. et. al *Microbiology* (2010), 156, 2271-2282.



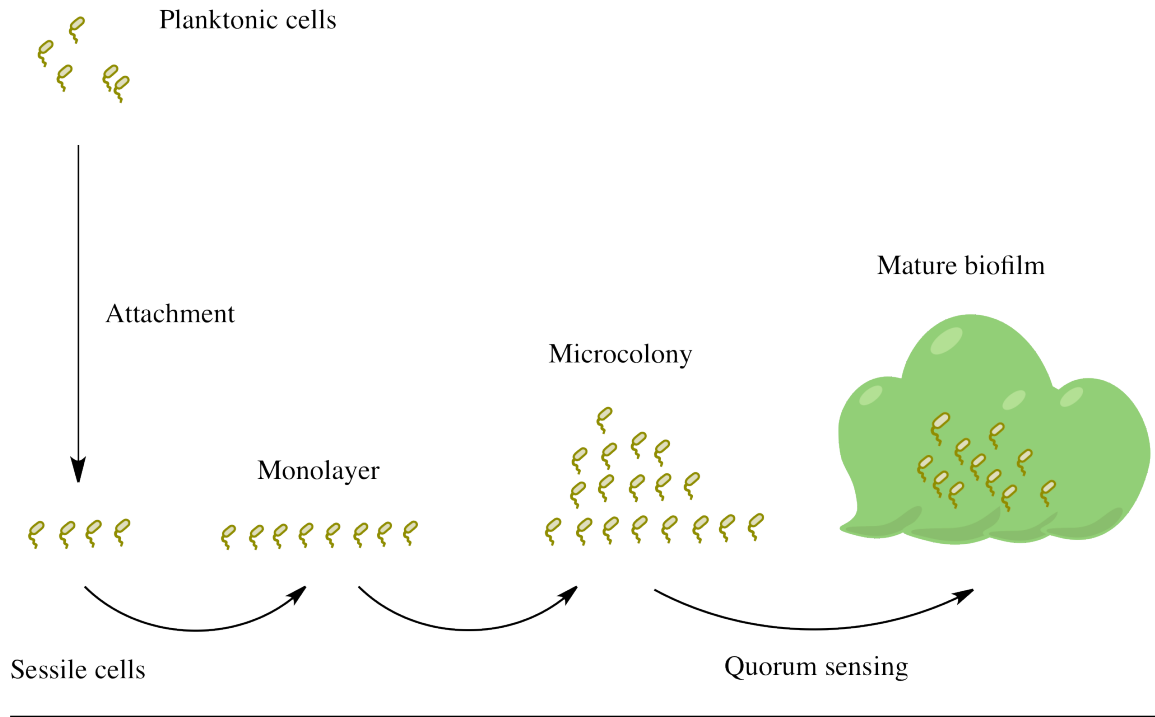
In addition to the use of *P. aeruginosa*'s quorum signals for cognate sensing system function, the signals themselves elicit responses in the host. In mammalian cells, these effects are mediated by 3OC<sub>12</sub>-HSL, but not by C<sub>4</sub>-HSL, or other similar AHL species including the 3OC<sub>12</sub>-HSL enantiomer, indicating structural specificity.<sup>6-8</sup> 3OC<sub>12</sub>-HSL has been found to induce apoptosis in certain types of mammalian cells including macrophages, neutrophils, fibroblasts, lymphocytes, monocytes, mast cells, breast cancer cells and endothelial cells.<sup>6-9</sup> The 3OC<sub>12</sub>-HSL apoptotic pathways proposed are many including the mitochondrial pathway via caspases-9, 3, and 7, an extracellular pathway via caspase 8, or through extracellularly induced, intracellular Ca<sup>++</sup> bursts found to correlate to cell deaths.<sup>6,7,10,11</sup> Along with apoptotic effects, cell structures have also been reported to affect 3OC<sub>12</sub>-HSL exposures. For example, gut epithelial cells exposed to 3OC<sub>12</sub>-HSL had tight-junction proteins that underwent phosphorylation leading to actin cytoskeleton alterations, and lowered transepithelial electrical resistance.<sup>12</sup> These modifications could in turn explain *P. aeruginosa*'s translocation and infection of the intestine in certain disease states.

Given the complexity and severity of both the quorum signal exogenous effects on the host along with the resultant upregulated toxin expression, one might wonder how an organism's defenses might get any better. In the case of *P. aeruginosa*, it turns out that all of these deleterious effects are further exacerbated through the concomitant expression of biofilm producing genes as an integral part of quorum system response. The biofilm produced provides a safe harbor for bacterial colonies while they continue to compete with the host and other bacterial species.

### **1.1.3 Quorum signal-induced phenotype expression (biofilms)**

Generally, biofilms exist as aggregates of bacterial, fungal, and algal communities that sequester themselves in a glue-like matrix typically composed of exopolysaccharide material. Not only is there great diversity across species within a biofilm, but even within a single species there can be many different levels of genetic expression observed depending on location within the film. The development of a bacterial biofilm starts with free floating, or planktonic bacteria finding a low flow area on which it can attach and become sessile. These sessile species slowly increase at the same time that they also are producing quorum signals. At some point in time, when the increasing surface concentration of bacteria are able to achieve the quorum concentration, the quorum sensing receptors inside the bacteria bind their respective promoter regions and begin to up-regulate genes including those responsible for biofilm production. This allows for progression toward a mature biofilm status where the bacteria are engulfed in a protective matting of exopolysaccharide material providing protection to the colony from the environment (Figure 4). Given the ubiquity of AHLs as signaling molecules, expected cross talk promotes biological diversity as more than one species may benefit from each other's signal (e.g., Gram-negative bacteria).

**Figure 4. Biofilm formation**



Biofilm formation revealed as a change in cell status from planktonic state to full fledged biofilm-forming state via QSS.

While the biofilm itself is a significant phenotypic change it should be noted that the change in gene expression dictated by QSS is even more impressive as organisms like *P. aeruginosa* have proteome expression levels that have been observed to change by as much as 50% from its planktonic state to its full-fledged, mature biofilm state.<sup>13</sup> This results in phenotypic changes including production of exopolysaccharides and biofilm formation concomitant with changes in metabolism and pathogenic toxin production as already discussed. Indeed, not only has biofilm production been suggested as a means of facilitating antibiotic resistance, but genes under the influence of QS molecules are also known to code for antibiotic efflux pumps that further protect the bacterium from external challenges.<sup>14</sup> Clearly, these phenotypical expressions are undesired from many different standpoints and comprise challenges in many different fields from medicine to engineering and public health.

#### **1.1.4 Mitigation of quorum sensing networks**

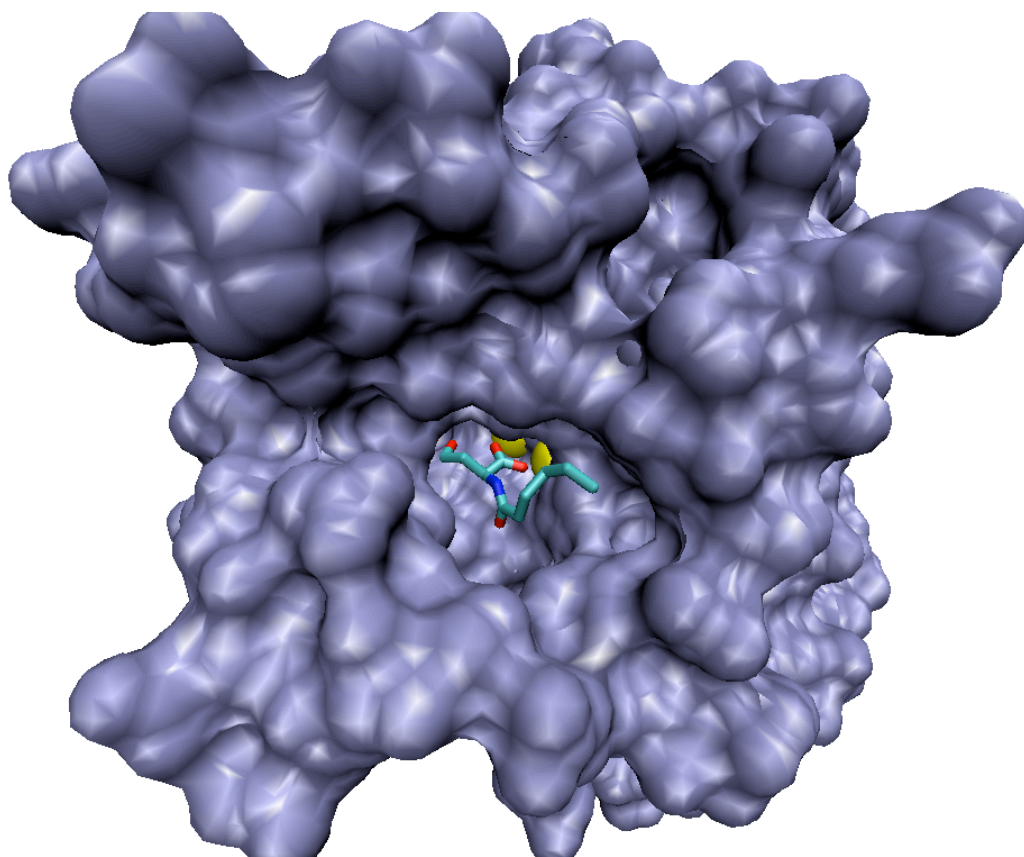
The existence of enzymes whose purpose is to degrade quorum signaling molecules has been known for at least the last decade. These quorum quenching enzymes exist in prokaryotes and eukaryotes to facilitate prokaryote-prokaryote and prokaryote-eukaryote interactions. In some cases, quorum signaling bacteria themselves can produce quorum quenching enzymes that target their own signal molecules. In one study, quorum signal lactonase activity was observed in human airway epithelial cells upon testing transgenic expression of the *AiiA* gene in cell lines.<sup>15-18</sup> The lactonase activity was conferred to the paraoxonase (PON) gene that displays activity toward over 30 different non-AHL type lactones. Later the same PON enzymes were found to also exhibit hydrolytic activity in other capacities, among which were nerve agent detoxification, and

drug metabolism.<sup>19</sup> It has been demonstrated that a Gram-positive soil-dwelling bacterium, *Bacillus thuringiensis* (*B. thuringiensis*), expresses a lactonase enzyme (AiiA) that targets and destroys the AHLs of *P. aeruginosa* through hydrolysis of the lactone ring.<sup>20-22</sup> The enzyme has been subject to study and is of interest as its resultant effect is to disrupt *P. aeruginosa*'s (and that of other Gram-negative bacteria) ability to communicate, thus, preventing quorum-induced gene expression. Some understanding of the enzyme's mechanism of action has been obtained, but other functions have yet to be fully elucidated.

#### **1.1.5 *Bacillus thuringiensis* lactonase**

The *B. thuringiensis* AiiA enzyme is a 250-residue metalloprotease containing two zinc ions in close proximity (Figures 5 and 6).<sup>20-22</sup> Sequence analysis reveals a conserved HXHXDH motif that is shared with the metallo- $\beta$ -lactamase superfamily of proteins. The zinc ions are identified by the coordinating ligands of the protein, namely, H104, H169, H106, and H235, H109, and D108 for Zn1 and Zn2, respectively (Figures 7 and 8).

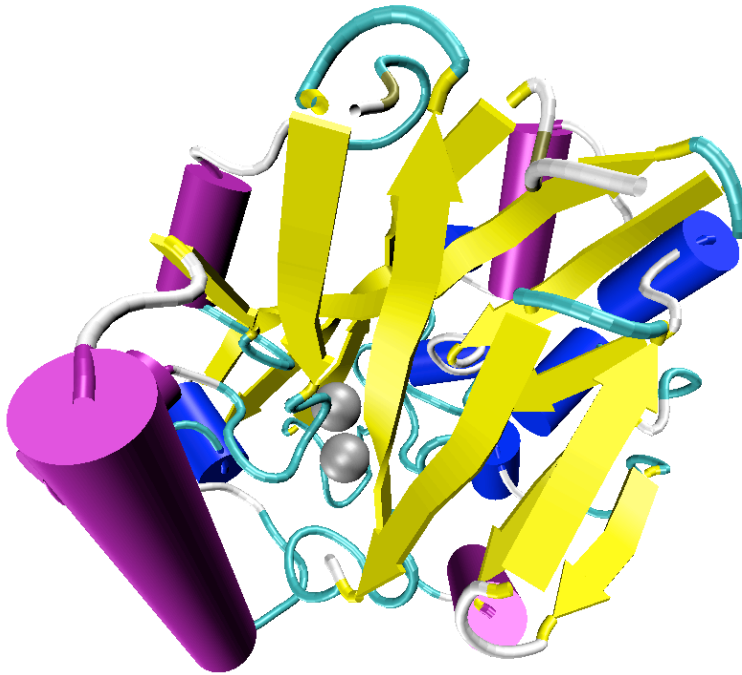
**Figure 5. Lactonase with bound product**



---

Lactonase crystal structure 3DHB.pdb shown with product in complex at the active site. Note zinc 2+ ions shown in yellow.

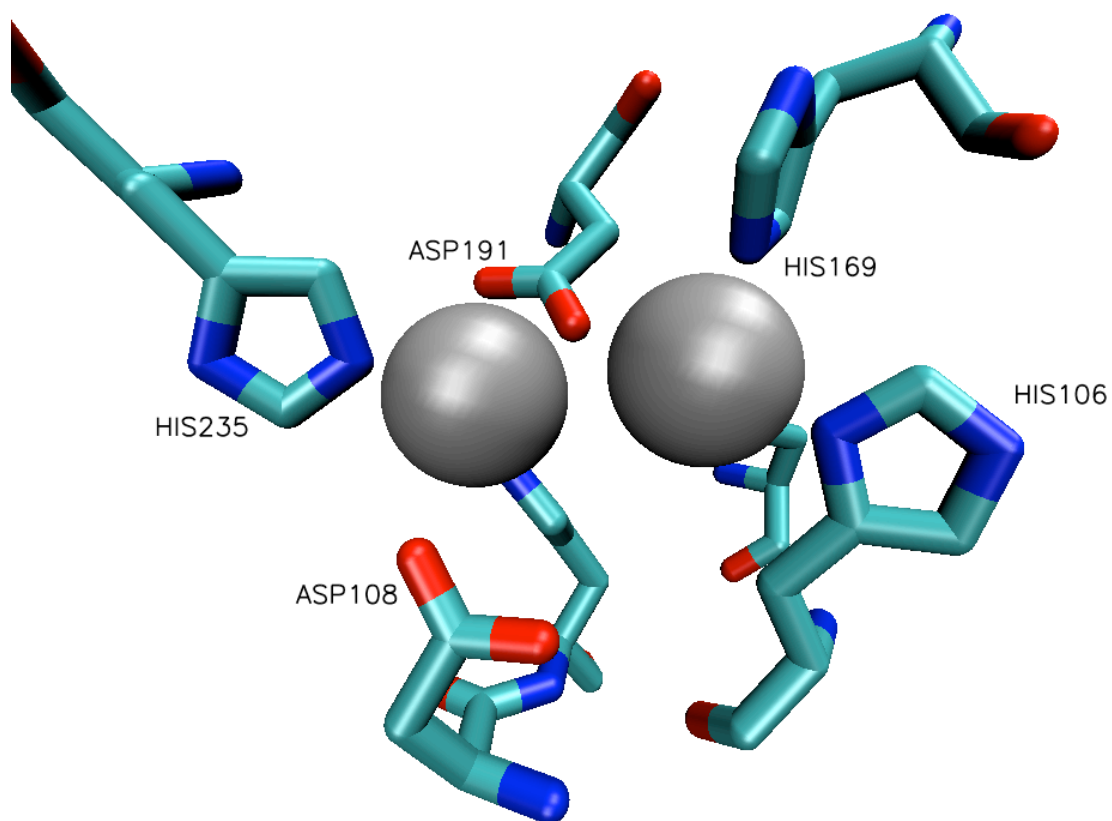
**Figure 6. Lactonase secondary structure**



---

Note metallo- $\beta$ -lactamase-like ( $\alpha\beta/\beta\alpha$ ) fold. The two major  $\beta$ -sheets are in parallel formation at the center of structure surrounded by  $\alpha$ -helices. Coordinated zinc  $2+$  ions are rendered as silver spheres between the sheets.

**Figure 7. Lactonase active site with ligating residues**

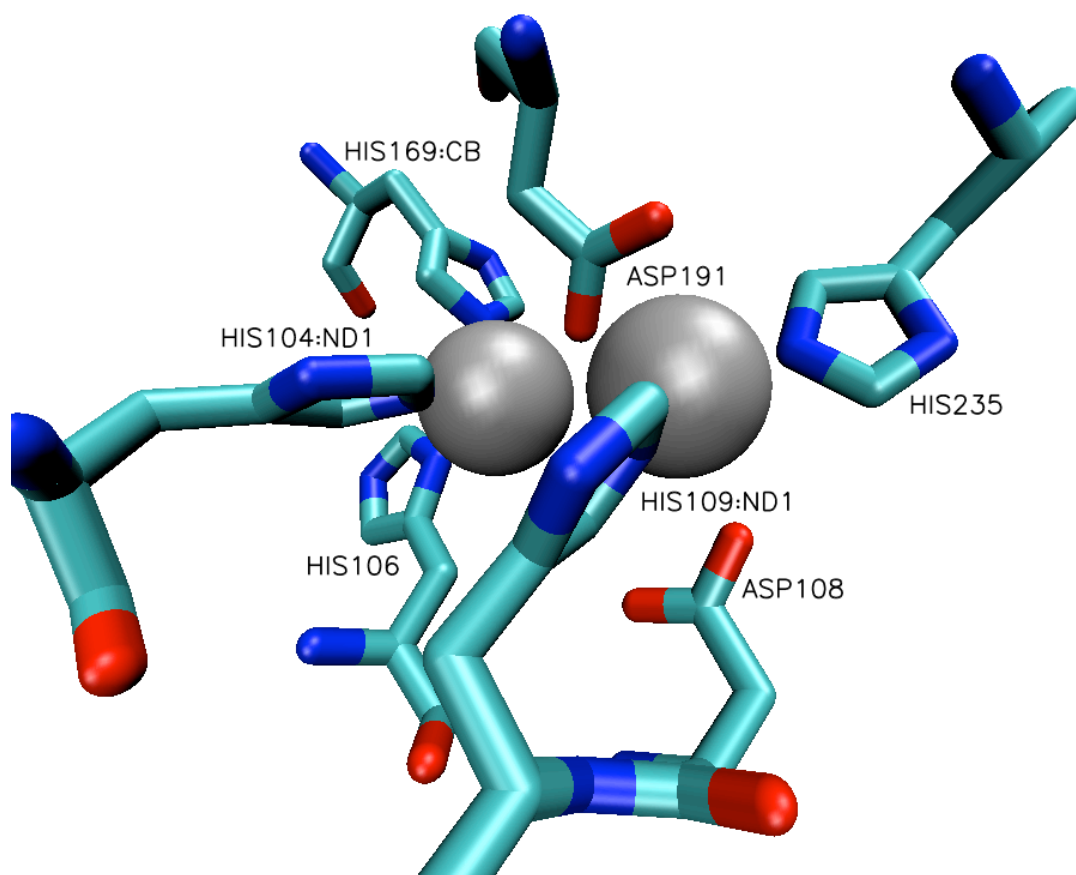


---

This view is to show the catalytic site as presented to ligand facing the reader.



**Figure 8. Lactonase active site - alternate view**

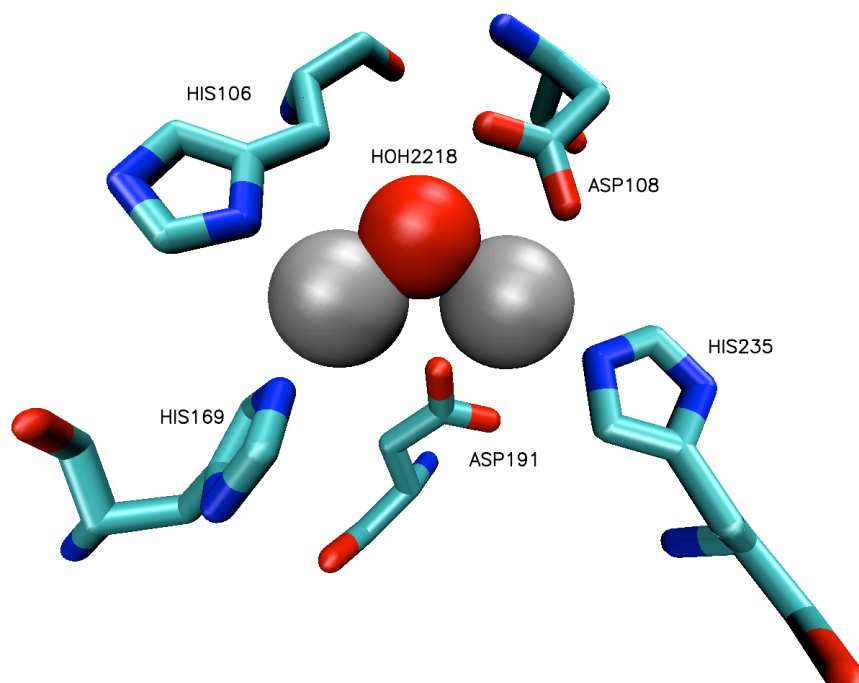


---

This can be seen as the backside of the active site. Note coordination of metal ions.

D191 serves to ligate both Zn1 and Zn2. Crystal structure determinations and quantum mechanical studies have led to the suggestion that the lactone head adopts a bridging position across the zinc ions by orienting its ether and carbonyl oxygens toward Zn1 and Zn2, respectively. Crystal structure studies have elucidated an oxygen density that is consistent with a hydroxide ion or water molecule that also bridges the zinc atoms (Figure 9).<sup>10</sup> Given that metal associated water has a pKa of approximately 7.5, we reasoned that the oxygen density was most likely a hydroxide anion.<sup>23</sup> Kim et al.<sup>7</sup> and Liu et al.<sup>9</sup> have proposed a mechanism whereby the lactone ring is opened by means of base-promoted hydrolysis. The substrate bridging allows for optimal positioning of the hydroxide ion for nucleophilic attack on the lactone's activated carbonyl carbon (through polarization by the zinc ions).<sup>7</sup> While substrate placement and the general mode of catalysis seem to be agreed upon, certain related elements remain unclear.

**Figure 9. Oxygen bridging zincs in lactonase active site**



---

Per pKa analyses and general chemistry principals, the oxygen density seen in the x-ray crystal structure is rationalized as a water molecule or hydroxide ion.

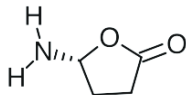
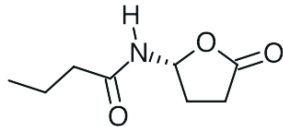
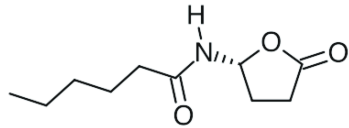
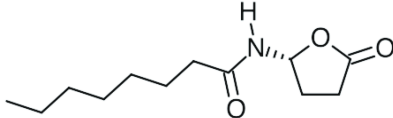
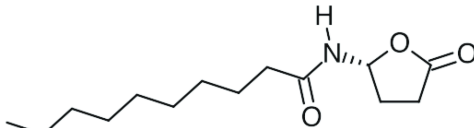
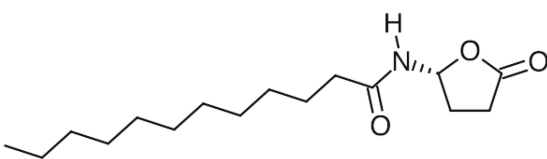
Y194 is a key residue whose function we sought to validate. Y194F mutant studies show a loss of 30% activity that has led to multiple proposed explanations.<sup>20,21,24,25</sup> Liu (2005) et al.<sup>7</sup> initially described the residue as having either an oxyanion hole function to stabilize the substrate tetrahedral intermediate or the residue is a general acid that protonates the leaving group alkoxide. Next, Kim (2005) et al.<sup>7</sup> described Y194 as a general acid as well after measuring the Y194F mutation activity. Liu (2008) et al.,<sup>9</sup> upon further investigation, presented Y194 as an oxyanion hole provider based on substrate lactone carbonyl oxygen to Y194 OH proton distance as observed during a QM/MM molecular dynamics analysis. In addition, the residue is seen to have a stabilizing effect on the lactone ring and was also identified as a key, binding residue. Last, Liao (2009) et al.<sup>8</sup> describe Y194 as an oxyanion hole provider based on quantum mechanics optimization calculations. It is puzzling, however, why an oxyanion hole explanation is favored given the strong, already stabilizing effects of the two  $\text{Zn}^{2+}$  ions that are coordinated to the substrate.

Proton shuttling is a mechanism by which a proton may be transferred to or from a substrate to a residue on its cognate enzyme to promote a given reaction. Upon completion of the reaction, the proton donor or acceptor is subsequently replenished with a proton from the substrate or from the environment via direct or indirect means. While proximity of a residue to its substrate is one way to qualify it as a candidate shuttle, relative pKa's are also critical to determining proton disposition. In the case of lactonase, the source of the proton shuttle remains to be fully addressed. In the base-promoted ring-opening mechanism, at least two proton transfers must occur that may or may not directly involve the enzyme. X-ray crystal structures reveal that a hydroxide ion or water must

bridge across the two zinc ions. Given the low concentration of hydroxide ions at pH 7, it's more likely that water attaches to  $\text{Zn}^{2+}$  first and is subsequently deprotonated given that a metal-bound water molecule's proton is significantly more acidic than a solution bulk water molecule (3-His-Zn(II)-OH<sub>2</sub> pK<sub>a</sub>=7.5 vs. H<sub>2</sub>O pK<sub>a</sub>=15.72).<sup>6</sup> Subsequent to hydroxide ion formation, the substrate enters the binding pocket and orients relative to the active site where the lactone ring carbonyl oxygen coordinates with a  $\text{Zn}^{2+}$  ion. This coordination allows for an electron density pull away from the carbonyl carbon making for a hot electrophilic site. At this point, the hydroxide is well posed for a nucleophilic attack on the substrate's carbonyl carbon. After the nucleophilic attack occurs, the tetrahedral intermediate collapses into an anion that must be re-protonated for ease of release from the cationic active site. Both D108 and Y194 are claimed to be proton shuttles in different studies, however, the presence of the active site zinc ions and ligand exchange effects have made it difficult to delineate which, if either, residue performs this function.

Another distinctive feature of the substrate is the acyl tail that is joined to the lactone head via an amide bond. By virtue of the functional groups in the tail, both the amide bond and the alkyl tail may confer activity of the enzyme toward the substrate. Interestingly, contradictory data exist that show both activity dependence and independence on tail length (Table 1).<sup>3,24</sup>

**Table 1. AHL degradation bioassay comparison**

Substrates	Activities %		
	Kim et al. <sup>10</sup>	Wang et al. <sup>5</sup>	
	HSL	ND	5.40
	C4AHL	61.4	89.1
	C6AHL	100.0	100.0
	C8AHL	116.6	93.9
	C10AHL	217.1	90.3
	C12AHL	31.6	-

Some of the substrate-enzyme interaction has been attributed to a hydrogen bond between the backbone NH group of F107 and the substrate amide carbonyl oxygen mediated by a water molecule, however, this interaction does not support a tail-selective activity. Tail-enzyme interactions to date have been described as not strong enough to have an effect on substrate binding given the amount of disorder seen in the alkyl portion in QM/MD simulations. The atomic B-factors derived from the 3DHB crystal structure for the AHL tail are significantly higher (average 29) than the other atoms (average 21) suggesting more relative disorder of the tail. This is not entirely unexpected; however, it is unclear how this translates to functional capability of the ligand's tail to the enzyme-substrate complex.

#### **1.1.6 Scope of research**

It has been established that in the absence of quorum signaling molecules, biofilms, along with other undesired phenotypes, are not expressed. As stated, the prevention of biofilms is of value in many different fields of endeavor. To that end, the full understanding of how lactonase enzymes function will provide both direct and indirect means of application to the potential design of systems that prevent biofilm formation. Using molecular dynamics, this study is to examine the nature of substrate binding, specifically to glean better understanding as to the function of the AHL alkyl tail. Next, clarification is desired to better understand the function of the Y194 residue when juxtaposed against the dizinc containing active site. Lastly, a description of the lactone ring-opening mechanism consistent with local electrostatic conditions is sought. Moreover, this is the first molecular dynamics based study of AHL-lactonase systems that is, to my knowledge, inclusive of an entire family of QSS molecules.

## **1.2 Biochemistry of addictive behaviors**

### **1.2.1 Serotonin, dopamine and addiction**

The establishment of neurochemical reward pathways can be seen to have evolved so as to reinforce behaviors that promote an organism's survival and/or propagation (e.g., high calorie food seeking and sexual drive). When subject to maladjustment, through whatever means, these same reward pathways can promote deviant behaviors such as those seen in drug addiction. Drug dependence, or addiction can be described as a physical or psychological state whereby a drug-consuming species is compelled to continue taking the substance on a regular basis to avoid the effects of withdrawal and/or to continue experiencing its effects. The drugs in question span an entire spectrum of pharmacological classes including depressants, opioids, stimulants, psychedelics, and inhalants. These substances also include prescription, over-the-counter, and illicitly obtained drugs thus providing very broad public access. The global cost of drug addiction is quite large and dependent on different countries' legal and cultural constructs, indeed targeted studies reveal the magnitude of some facets of the problem. For example, in the U.S. as drug criminalization laws have increased, total costs including crime, health, and productivity induced factors totaled more than \$193 billion in 2007.<sup>26</sup> The U.S. now possesses the highest incarceration rate in the world where as a nation that contains only 5% of the world's population it also contains 23% of the world's incarcerated people.<sup>27</sup> These facts do not even begin to touch on a myriad of other drug-related costs to society.

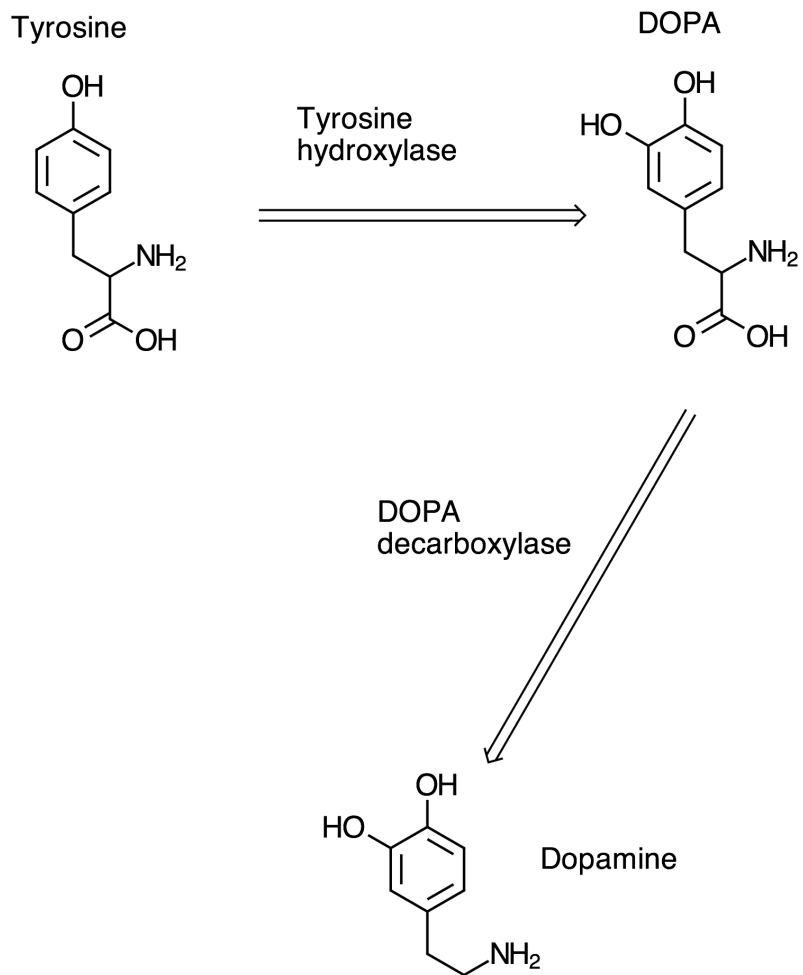
Drug addiction has both physical and cognitive causes that have been studied over decades in an attempt to find potential cures to the diseased state. Over time, the



molecular bases for behavior is defined as the study of numerous brain neurotransmitters and their functions with a vast, complex network of interaction. Both serotonin and dopamine are neurotransmitters that play a critical role in the reward system, each with their own cognate neural receptors and neuronal channels. The interplay of the serotonin and dopamine within the brain is still not fully understood, yet some level of mapping of the neurochemical centers and pathways are starting to be established.

Dopamine is a neurotransmitter synthesized naturally in humans starting from tyrosine and proceeding through reactions first with tyrosine hydroxylase and then with l-amino acid decarboxylase in the presynaptic terminal (Figure 10).<sup>28</sup> The neurotransmitter is released across the synapse for uptake at the postsynaptic cell at multiple receptor sites. The effects of dopamine are many and tend to vary according to the specific pathway taken. One of the effects of dopamine is the control of motor function. Antagonism of this motor control function is readily seen in humans with Parkinsonism. Emesis and lactation are other effects regulated by dopaminergic pathways. Lastly, behavior initiation (e.g., hallucinations, psychoses, and reward behaviors) is also controlled by dopamine, specifically via mesolimbic and mesocortical pathways initiated in the ventral tegmental area (VTA).<sup>28</sup>

**Figure 10. Dopamine biosynthetic pathway**

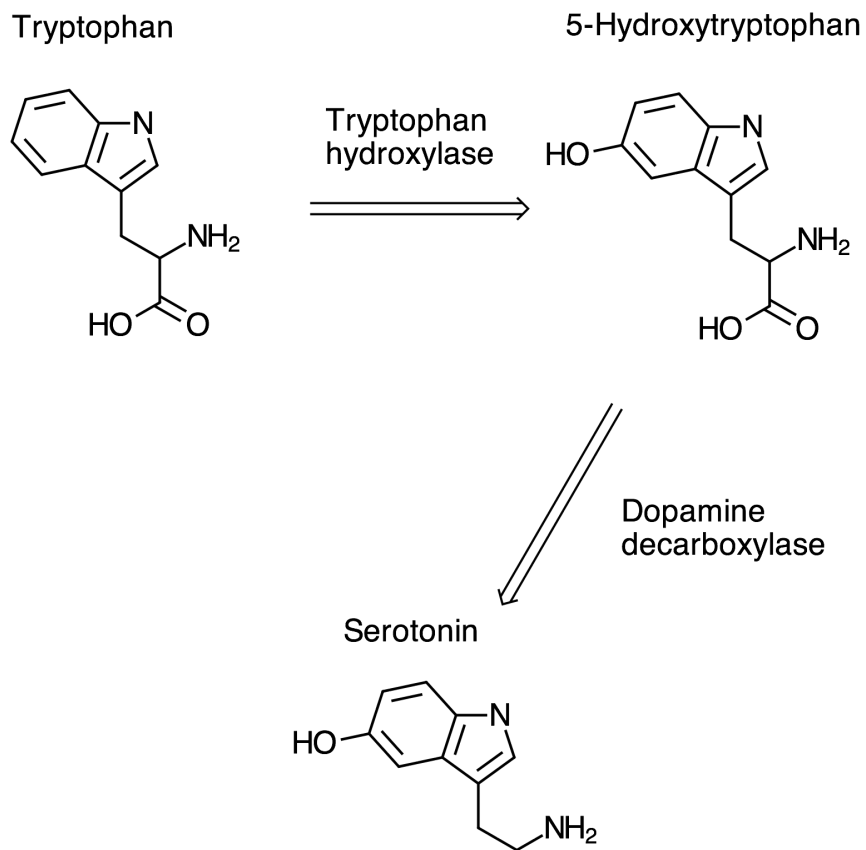


---

This is an abbreviated representation of dopamine biosynthesis starting from tyrosine.

Serotonin or 5-hydroxytryptamine (5-HT) is another neurotransmitter that is synthesized in an almost identical manner to dopamine starting with the amino acid tryptophan and progressing through reactions with its cognate hydroxylase followed by l-amino acid decarboxylase in the presynaptic terminal (Figure 11).<sup>28</sup> The responses effected by 5-HT receptors are many including anxiety, food intake, hallucinations, mood, emesis, obsessive behavior, and sexual function to name just a few.<sup>28</sup> The 5-HT system presents a high level of complexity as the pathways and receptor families are numerous.

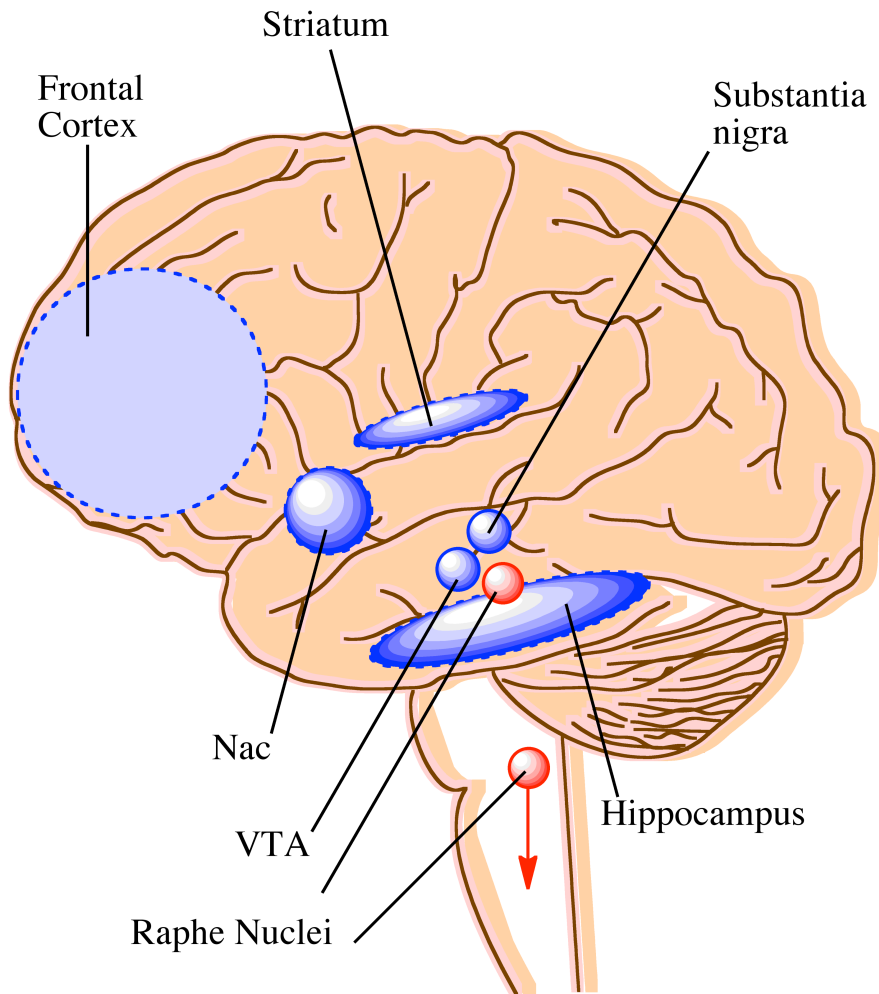
**Figure 11. Serotonin biosynthetic pathway**



This is an abbreviated representation of serotonin biosynthesis starting from tryptophan. Note similarity to the dopamine pathway.

Clusters of 5-HT neurons called Raphe nuclei exist that subdivide between superior and inferior positions in the brain. The inferior nuclei localize near and innervate brain stem structures to serve as regulators for motor activity and autonomic functions. The superior nuclei include both the dorsal (DRN) and median Raphe nuclei (MRN) that are generally held to contribute to reward pathways. The efferent nerves from the MRN project into the hypothalamus, amygdala, hippocampus, and frontal, cingulate, septal, occipital, and parietal cortices. The DRN innervates the locus coeruleus, periaqueductal grey, substantia nigra, striatum, nucleus accumbens (NAc), thalamus, hypothalamus, amygdala, hippocampus, and entorhinal and frontal cortices (Figure 12). It is the action of the 5-HT neurons on dopaminergic pathways and regions in the brain that are thought to comprise the serotonin-dopamine relationship in addictive behaviors.<sup>29-36</sup>

**Figure 12. Brain map of relevant regions to dopamine and serotonin signaling**



---

This figure is to provide an overview of serotonin and dopamine centers in the brain. Connectivity of these centers is still under intense study. The Raphe nuclei are known to contain clusters of 5-HT neurons.

### **1.2.2 5-HT<sub>2c</sub> Receptor**

5-HT receptors are many and present a significant challenge when trying to determine how they influence behaviors. There are 14 recognized 5-HT receptors grouped into seven different families. The 5-HT<sub>1</sub> and 5-HT<sub>2</sub> families have five and three subtypes, respectively. The majority of the receptors are G protein-coupled receptors (GPCR) the exception being the 5-HT<sub>3</sub> receptor, a ligand-gated ion channel. Of particular interest is the 5-HT<sub>2c</sub> receptor (5-HT<sub>2c</sub>R) implicated as having a significant role in dopamine neurotransmission in the midbrain region.<sup>37,38</sup>

Given that 5-HT<sub>2c</sub>R has been identified as a GPCR, the overall structure is taken to adopt the same as those seen in the crystallized GPCRs such as rhodopsin and  $\beta$ 2-adrenergic receptor based on sequence homologies and secondary structure analysis. The protein consists of seven transmembrane (7-TM)  $\alpha$ -helices interconnected by three extracellular (3-EL) and three intracellular peptide loops (3-IL). The N-terminus extends into the extracellular environment while the C-terminus extends into the cytosol. Interaction of both agonist and antagonist ligands within the cell surface accessible region of the 7-TM region induces a conformational change in the GPCR that allows for activation of its associated G-protein allowing for further downstream events.<sup>39</sup>

### **1.2.3 PTEN**

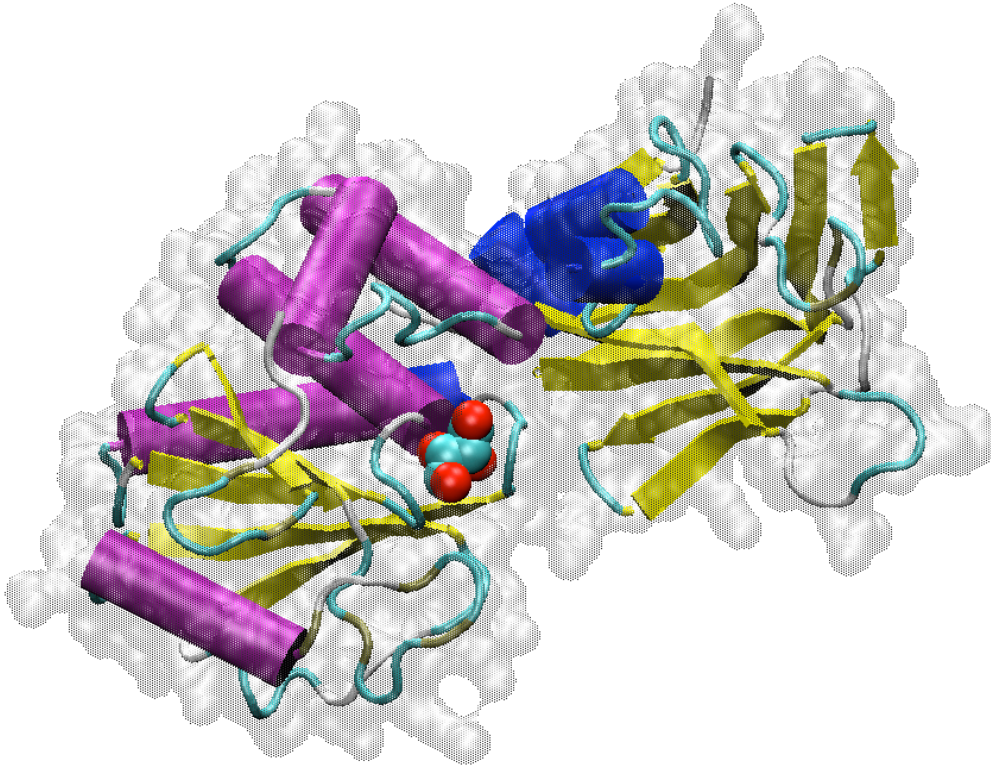
Phosphatase and tensin gene homolog is located on chromosome 10 (PTEN) is a gene that codes for a protein that is widely known for its tumor suppressive activity. PTEN is located on the long arm of chromosome 10 at position 23.3.<sup>40,41</sup> Mutation of this gene is common in many human cancers and is accountable for 50% of glioblastomas, endometrial carcinomas, prostate carcinomas, and melanomas. In addition, PTEN

mutations are also implicated in many benign neoplasms such as those induced by Cowden's and Bannayna-Zonana syndromes.<sup>42-48</sup> Along with general tumor suppression, PTEN has also been identified as having neuron generation and repair influence where its deletion in mice has been observed to facilitate nerve regeneration after injury.<sup>49</sup>

The PTEN protein consists of 403-amino acids containing a HCXXGXXR motif observed in both tyrosine phosphatases and dual specificity protein phosphatases. The N-terminal 190-residue region displays homology to tensin and auxilin while the C-terminal region contains about 220 amino acids with no clear homology. The phosphatase domain consists of a five-stranded  $\beta$  sheet with two  $\alpha$  helices on one side and four on the other (Figures 13 and 14). Mutations to either domain have been shown to reduce or abrogate PTEN's ability to check tumorigenesis.<sup>50</sup> The HCXXGXXR signature motif forms what is termed a "P-loop" that is similarly found in both protein tyrosine (PTP) and dual specificity phosphatases (DSP). This structure is critical to both substrate binding and to catalysis via the CYS-124 and ARG-130 residues. The other part of the binding pocket is formed by the "TI-loop", a 4-residue insertion to the p $\alpha$ 5- $\alpha$ 6 relative to the reference dual specificity phosphatase enzyme VHR. The final part of the pocket is made up of the WPD loop, or the p $\beta$ 4- $\alpha$ 3 section consisting of residues 91-94. The structural difference in the pocket relative to their related PTP and DSP structures, is the size, to accommodate the PI(3,4,5)P<sub>3</sub> substrate (Figures 15 and 16).



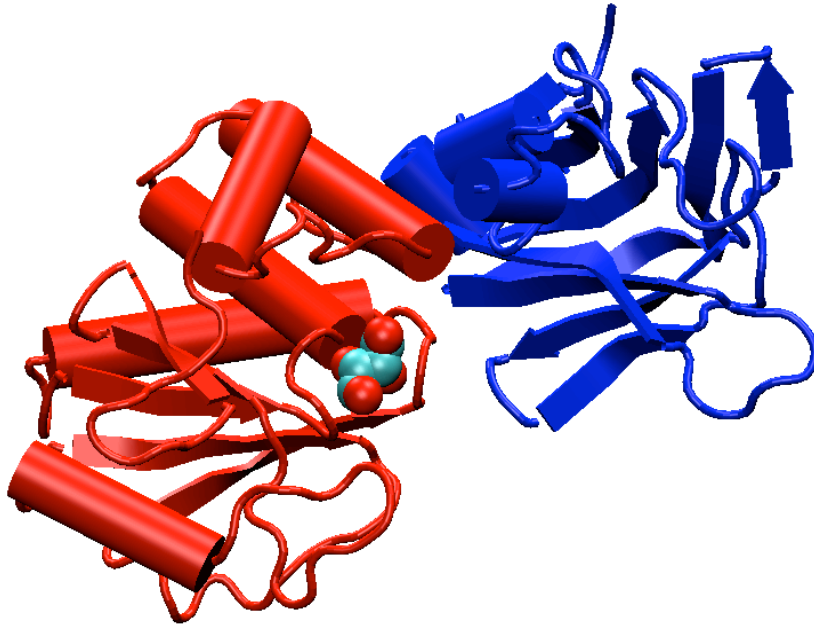
**Figure 13. PTEN secondary structure and volume**



---

Translucent rendering of human PTEN with underlying secondary structure revealed. Note the five stranded beta sheet (yellow) in the phosphatase domain and alpha helices (magenta).

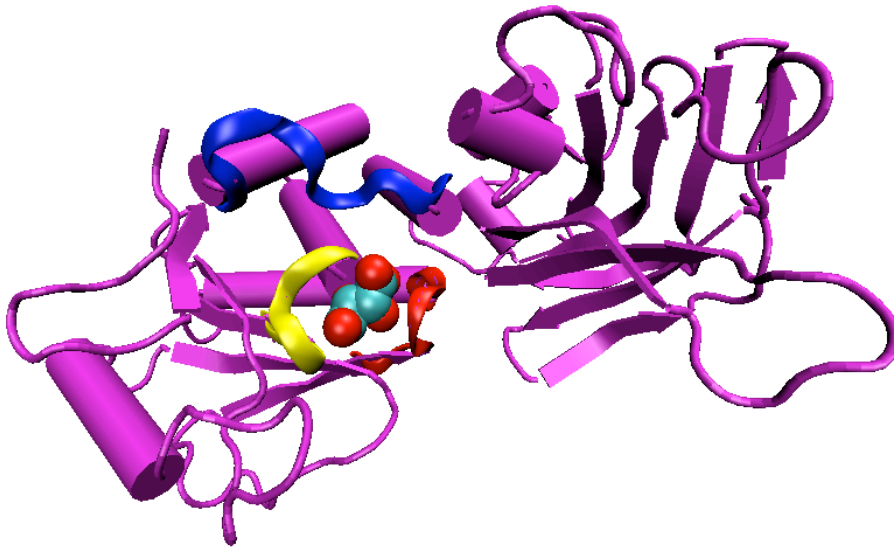
**Figure 14. PTEN domains**



---

van der Waals rendering of the substrate in the catalytic pocket of the phosphatase domain shown in red. The blue region is the membrane-associating domain.

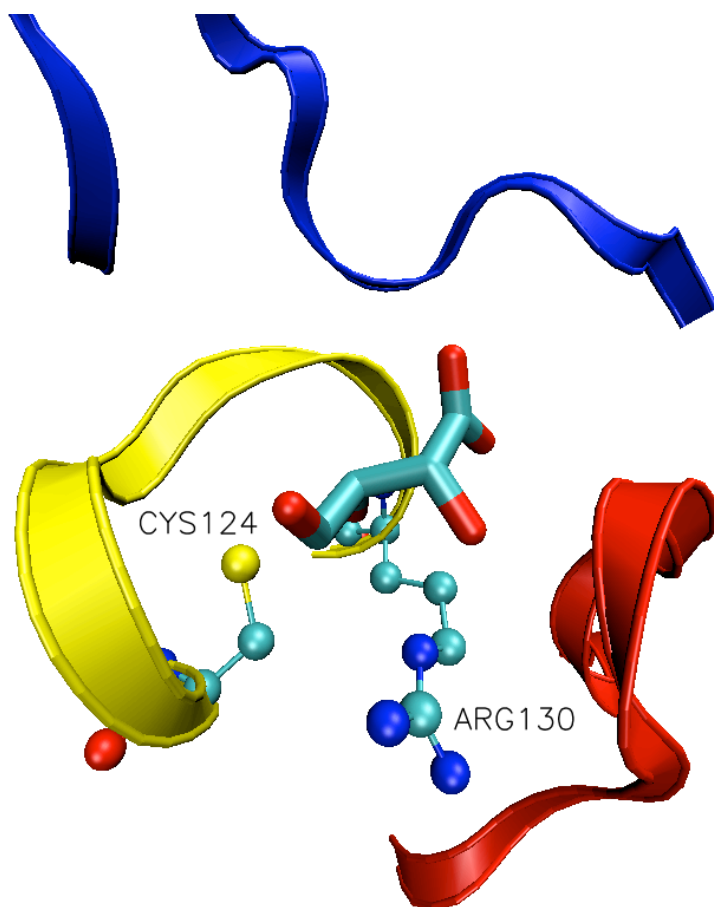
**Figure 15. PTEN with critical catalytic structures**



---

WPD loop rendered in red, P loop rendered in yellow, and TI loop rendered colored in blue. Substrate rendered as a van der Waals surface while the P loop is rendered as a yellow ribbon. TI loop is rendered as a blue ribbon. WPD loop is rendered as a red ribbon. These loops constitute the active site pocket for PIP3.

**Figure 16. PTEN with exploded view of catalytic site**

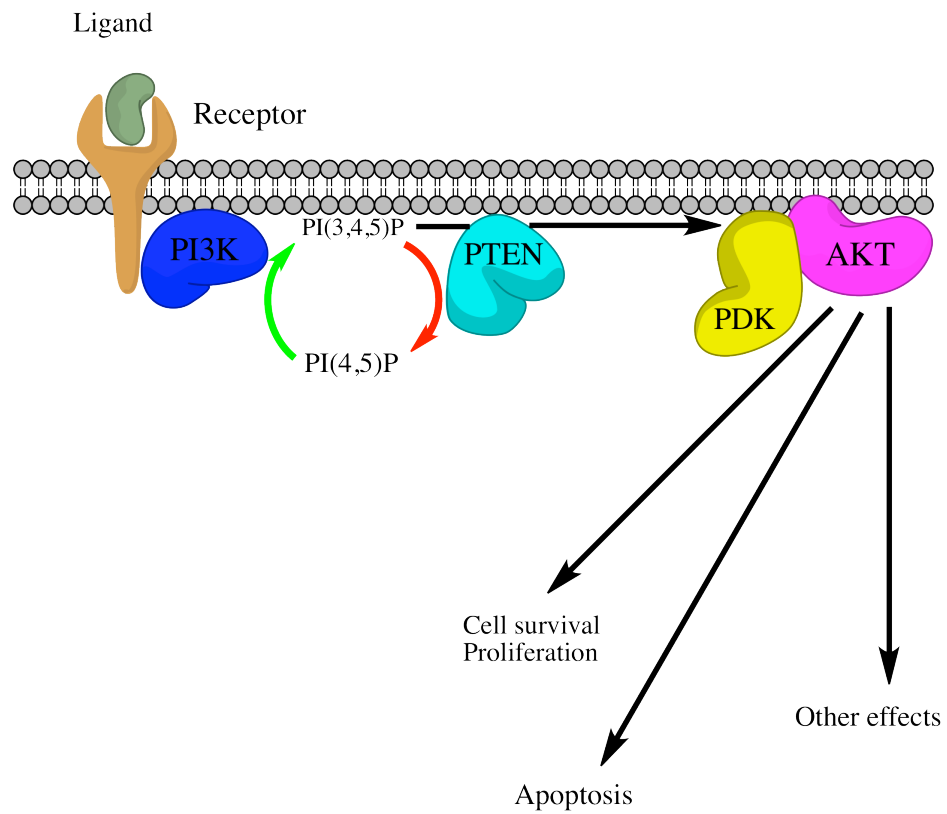


---

P loop critical residues CYS124 and ARG130 shown juxtaposed against the bound substrate analog, tartarate.

PTEN has been shown to dephosphorylate threonine, serine, and tyrosine-phosphorylated peptides. Dephosphorylation of phosphatidylinositol (3,4,5)-trisphosphate (PIP3) has also been observed with specificity toward the D3 position on the inositol ring. As a second messenger, PIP3 is produced via phosphoinositide 3-kinase that in turn activates the Akt/PKB kinases that play a multifaceted role in cell death, senescence, survival, metabolism, and angiogenesis – all of which play roles in tumorigenesis and regression. The general thought is that PTEN dephosphorylates PIP3 at some rate closely coupled to its own production rate thus providing for a PIP3 buffer that damps its tumor generating capability (Figure 17).<sup>51-54</sup>

**Figure 17. PTEN and PIP3 signaling process**



Adapted from Planchon et. al., *Journal of Cell Science* (2008), 121, 249-253.

### 1.2.4 5-HT2cR:PTEN interaction

A potential interaction of 5-HT2cR with PTEN was hypothesized upon taking into account several factors.<sup>55,56</sup> First, as an already identified tumor suppressor, PTEN was found expressed in many if not most body tissues including brain tissues. Next, 5-HT2cR in addition to being found in the gastrointestinal tract, has also been shown to be distributed in central nervous system tissues. Lastly, 5-HT2cR phosphorylation status was observed to be necessary for effective signaling. Thus, given the phosphatase activity of PTEN combined with the required phosphorylation states required for 5-HT2cR activity led researchers to consider a 5-HT2cR:PTEN interaction as a mediator of serotonin-dopamine dependent pathways (e.g., reward behaviors).

In 2005, a study was conducted to investigate the potential for an efficacious 5-HT2cR:PTEN interaction.<sup>55</sup> Immunoprecipitation experiments were first conducted whereby PTEN specific antibody coimmunoprecipitated 5HT-2cR and 5-HT2cR specific antibody coimmunoprecipitated PTEN thus suggesting the hypothesized interaction. The next set of pull-down experiments were conducted to better localize the interaction and revealed the 5-HT2cR third intracellular loop as the point of PTEN association. This loop was further subdivided into five fragments that were again subject to pull-down assays that more pointedly demonstrated the fourth fragment of the third intracellular loop 3L4F as the key region of 5-HT2cR:PTEN association. Band-shift phosphorylation assays were able to reveal PTEN's ability to dephosphorylate 5-HT2cR agonist induced phosphorylation.

Subsequent *in vivo* administration of Tat-3L4F into rat VTA demonstrated preclusion of PTEN:5-HT2cR coimmunoprecipitation suggesting the interaction might

have physiological effects. From this point, colocalization studies were able to determine the location of dopaminergic neurons that expressed both PTEN and 5HT-2cR in rat VTA. Next, neuronal retrograde staining was then able to show how these PTEN and 5HT-2cR containing neurons in the VTA innervate the NAc, a region well known as one part of the brain's "pleasure center". Following that experiment, Electrophysiological testing was able to show how both systemic and intra-VTA injections of Tat-3L4F are able to suppress the firing rate of VTA dopamine neurons. Finally, as a means to demonstrate potential efficacy as a therapeutic, Tat-3L4F was shown to inhibit conditioned place preference induced by tetrahydrocannabinol, and nicotine in rats. Taken together, the experimental evidence was able to show PTEN's role in mediating the rewarding effects of drugs via the VTA-NAc dopamine pathway.

### **1.2.5 Scope of research**

While it has been strongly suggested that disruption of the 5HT-2cR:PTEN complex can suppress addictive behaviors, the exact biophysical nature of this interaction is as yet unknown. More precise knowledge of this association is necessary for further refinement of the 3L4F peptide to a pharmacophore model and/or reduction to an efficacious peptidomimetic or other small molecule. A significant challenge to this knowledge gap is the lack of x-ray crystallographic or NMR derived data that are generally preferred to better define the interaction bases. Thus, within the confines of what is a blind docking problem, the scope of this research is to first define the possible binding modes of the 5HT-2cR:PTEN complex by examining both PTEN and 3L4F biophysical and biochemical properties. Next, reduce the candidate peptide ligands of a competent 5HT-2cR:PTEN complex to a select few that might be chemically transformed



into a peptidomimetic. Last, test and validate both the candidate peptide and/or peptidomimetic ligand against favorable PTEN docking sites to deduce which ligand and PTEN residues are critical and how they inform further ligand modification. While the original 3L4F fragment amino acid sequence, PNQDQNARRRKKKERR, was briefly looked at, the main focus this study will be on the reduced sequence PNQDQNAR (3L4F-F1) whose effective binding efficacy was demonstrated early on in this work and took us yet closer to the goal of a realistic pharmacophore candidate.

### **1.3 Theoretical background of methods**

#### **1.3.1 pK<sub>a</sub> prediction**

In attempting to understand the structural, catalytic, or binding behavior of a protein, knowledge of the chemical character of its comprising residues is critical. These residues in turn, consist of amino acids with non-polar, polar, and ionizable properties. As the titratable nature of the ionizable residues implies a change in protonation state, it is critical that the proper state is used when performing molecular dynamics analyses. The natural place to start when considering ionizable residues is to take into account their acidity (alternately basicity) as defined by their acid dissociation constants or the negative logarithms thereof. pK<sub>a</sub>'s are typically referenced in their aqueous states, however as their proteinaceous states may be significantly different, the problem remains as how to calculate a meaningful perturbation. This tasking is accomplished in the University of Houston Brownian Dynamics (UHBD) program.

The UHBD program methodology uses the finite difference Poisson-Boltzmann (FDPB) algorithm by which to calculate the charge-charge interactions amongst a system

of charges.<sup>57-59</sup> The FDPB algorithm is a numerical solution to the linearized Poisson-Boltzmann equation

$$\nabla \cdot \epsilon(r) \nabla \phi(r) = -4\pi \rho^f(r) + \lambda(r) \bar{\kappa}^2 \phi(r)$$

where  $\epsilon$  is the dielectric constant,  $\phi$  is the electrostatic potential, and  $\rho^f$  is the fixed charge density, all at point  $r$ .  $\lambda$  delineates regions accessible to mobile ions and  $\bar{\kappa}$  is a modified Debye-Hückel parameter. The FDPB algorithm is ultimately used to calculate energies required for the pairwise interaction energy grid used to solve the perturbed  $pK_a$ 's of ionizable residues of a given system.

In considering the protonation of a single residue, the free energy change of that residue to the protonated state from the deprotonated state is

$$\Delta G = 2.303RT(pH - pK_a^{(w)})$$

where  $pK_a^{(w)}$  is the experimental  $pK_a$  of the residue in solution while the pH is the pH of the environment. In UHBD this equation is further extended to allow for electrostatic interaction analysis of several titratable residues within a given protein. In the case where all ionizable residues are deprotonated as a reference state where there are  $M$  ionizable groups, a single Gibbs' free energy state of the entire protein can be described mathematically as

$$\Delta G(x'_1, \dots, x'_M) = 2.303RT \sum_{i=1}^M x'_i (pH - pK_i^{(w)}) + \sum_{i=1}^M x'_i \gamma_i G_{ii} + \sum_{i=1}^{M-1} \sum_{j=i+1}^M [(q_i^\circ + x'_i)(q_i^\circ + x'_j) - q_j^\circ q_j^\circ] \Psi_{ij}$$

where  $x'_i$  is 1 when the residue is protonated and 0 when it is deprotonated,  $q_i^\circ$  is the formal charge of the deprotonated residue, and  $\Psi_{ij}$  is the absolute value of the  $G_{ij}$  components of the interaction matrix.

Given the binary nature of residue protonation, one can easily see that there are  $2^M$  energy states available. To determine which states are most probable, a distribution of protonation states (DOPS) algorithm is utilized. In this algorithm, all ionizable residues start in their protonated state for which the Gibb's free energy is then calculated. Next, another state energy is calculated based on a randomized value for each ionizable residue that is accepted when it is greater than a predetermined acceptance number. This final energy state is compared to the initial state and is accepted or rejected according to the Metropolis algorithm. Iterating through this procedure of calculating and testing the resulting states of a protein based on an ensemble of ionization states, allows one to eventually arrive at a reasonable estimation of a final, most probable energy state ultimately comprised of ionizable residues in their most probable ionization states. The end point of this analysis provides the computational scientist with a means to provide correct protonation states for the titratable residues of his protein as a preprocessing step prior to molecular dynamics modeling.

pK<sub>a</sub> analysis, however, is not limited to just first principals-based methodologies but lends itself readily to more empirically driven algorithms. In PROPKA, for example, the core assumption starts with a residue model pK<sub>a</sub> just like in UHBD.<sup>60,61</sup> However, the change in pK<sub>a</sub> term is derived according to the residue in question. For ASP and GLU the number of hydrogen bonds associated with the carboxyl group correlate to pK<sub>a</sub> shift as the product of the number of hydrogen bonds times a fitted base value. As it is also well recognized that hydrogen bonds are distance and angle dependent, these terms are also incorporated as a distance/angle function to the nearest side-chain hydrogen bond donor/acceptor ( $\Delta pK_{SDC-HB}$ ) along with the same consideration given to backbone

hydrogen bonds ( $\Delta pK_{BKB-HB}$ ). From a physical standpoint, not only are hydrogen bonds important players in  $pK_a$  modification, but (de)solvation effects are also key microenvironment contributors. Subject to hydrogen bond geometry, the relations are:

$$\Delta pK_{SDC-HB} = \begin{cases} C_{HB} & \text{if } D \leq d_1 \\ C_{HB} \cdot \frac{D - d_2}{d_1 - d_2} & \text{if } d_1 < D < d_2 \\ 0 & \text{if } d_2 \leq D \end{cases}$$

$$\Delta pK_{BKB-HB} = \begin{cases} -\cos\theta \cdot C_{HB} & \text{if } D < d_1, \theta > 90^\circ \\ -\cos\theta \cdot C_{HB} \cdot \frac{D - d_2}{d_1 - d_2} & \text{if } d_1 < D < d_2, \theta > 90^\circ \\ 0 & \text{if } d_2 \leq D, \theta \leq 90^\circ \end{cases}$$

As an example, desolvation is known to raise the  $pK_a$  of aspartic and glutamic acid (nominal  $pK_a$  range of 4.5) in hen egg white lysozyme, RNase H, xylanase, and human thioredoxin to values ranging from 6.1 to 8.1. To this effect, parameters that consider the depth of residue burial then become important to predict the degree of  $pK_a$  shift to assign. As a means to capture these parameters, analogously described as solvent accessible surface area, the number of non-hydrogen atoms surrounding the center point of an ionizable group are tallied and multiplied by a fitted base value to generate the resulting  $pK_a$  adjustment. This tally is not exhaustive, but limited to a given, local radius around the residue in question to establish what is deemed to be a local desolvation factor ( $\Delta pK_{LocalDes}$ ). While this initial correlation was found to be adequate for more surface-localized ionizable residues, yet more detail was needed for more deeply buried residues. By “deeply buried” it was determined the number of nonhydrogen atoms within 15.5 Å

of the ionizable group must be equal to or greater than 400 to achieve this designation as a global desolvation factor ( $\Delta pK_{GlobalDes}$ ). These two resulting desolvation factors are indeed additive in effect allowing for simple and fast calculation of the total desolvation  $pK_a$  shift. Lastly, charge-charge interactions ( $\Delta pK_{chgchg}$ ) are accounted for via the same distance correlation as that used for hydrogen bonding effects. Per above, the relationships resemble:

$$\Delta pK_{LocalDes} = N_{Local} \cdot C_{Local}$$

$$N_{15.5\text{\AA}} \geq 400$$

$$\Delta pK_{GlobalDes} = (N_{15.5\text{\AA}} - 400) \cdot C_{Local}$$

$$\Delta pK_{Des} = \Delta pK_{GlobalDes} + \Delta pK_{LocalDes}$$

$$\Delta pK_{chgchg} = \left\{ \begin{array}{ll} C_{chg} & \text{if } D \leq d_1 \\ C_{chg} \cdot \frac{D - d_2}{d_1 - d_2} & \text{if } d_1 < D < d_2 \\ 0 & \text{if } d_2 \leq D \end{array} \right\}$$

### 1.3.2 Molecular dynamics

Within the field of molecular modeling, “molecular dynamics” generally refers to the simulation of a chemical system in a deterministic manner, using molecular mechanics force fields as particle motion drivers. The outputs of these simulations result in trajectories of atomic motions with respect to time that represent samples of the conformational space of a system. Atomistic trajectories and conformational samplings

are key to determining both bulk solution phase properties and to understand both inter and intramolecular behavior of a system. In setting up one of these simulations, certain steps must be taken. First, as already discussed, the states of the ionizable residues must be determined through pK<sub>a</sub> prediction analysis and defined in the appropriate input file. Next, force field parameters must be assigned to the molecular system under study. The system is then solvated in a designated volume, either implicitly or explicitly, to which counterions are added to balance the net charge of the original complex to approximate physiological conditions. Lastly, the physical conditions of the simulation including step size, boundary conditions, execution time, statistical ensemble, and cut-off distances are designated after which all configuration files are submitted for the computational run.

The force fields that are applied to the molecular systems are comprised of both bonded and non-bonded interactions described mathematically as a potential energy below.<sup>62-64</sup>

$$\begin{aligned}
 V(r^N) = & \sum_{bonds} \frac{k_i}{2} (l_i - l_{i,0})^2 \\
 & + \sum_{angles} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 \\
 & + \sum_{torsions} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]
 \end{aligned}$$

$$+ \sum_{i=1}^N \sum_{j=i+1}^N \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}^2}$$

With regards to the bonded force field, the first two terms derive from the mathematical model for the potential energy of a simple harmonic oscillator where the independent variables are bond length and bond angle. The third term is periodic so as to describe the energy derive about a dihedral angle or torsion. The non-bonded elements of the force field are taken as the 12-6 Lennard–Jones potential and Coulomb’s law potential for the fourth and fifth terms of the field respectively. From these potential fields, resultant forces are calculated for each step in a molecular dynamics simulation so that Newton’s laws of motion may then be applied to all atoms of the system under study.

Commonly used force fields such as GROMOS<sup>65</sup> and CHARMM<sup>66</sup> use different methods by which to parameterize their force constants. These methods generally involve the use of molecular quantum mechanics routines to generate potential energy surfaces from which the force constants are derived. Next, molecular dynamics runs of various solvents (e.g., water) are run to check densities, enthalpies, and radial distribution functions versus experimental data. Necessary corrections are made to ensure the force fields can match liquid bulk properties. Further, these force fields may use a united atom approach whereby certain common groups are amalgamated (e.g., –CH<sub>2</sub>–) or explicit atom models may be employed.

As the computational engine for a molecular dynamics simulation rests on integration of Newton’s laws, a velocity equation in addition to the force field is required and goes as follows:

$$v(t) = \frac{r(t + \Delta t) - r(t - \Delta t)}{2\Delta t} + O(\Delta t^2)$$

This velocity equation is based on the Verlet algorithm that is used for kinetic energy determination. An example subroutine is shown in Figure 18.



**Figure 18. Molecular dynamics sample code**

Molecular Dynamics Pseudocode	
subroutine integrate(f,en)	Integrate equations of motion
sumv=0	
sumv2=0	
do i=1, npart	MD loop
xx=2*x(i)-xm(i)+delt**2*f(i)	Verlet algorithm
vi=(xx-xm(i))/(2*delt)	velocity
sumv=sumv+vi	Velocity center of mass
sumv2=sumv2+vi**2	Total kinetic energy
xm(i)=x(i)	Update positions previous time
x(i)=xx	Update positions current time
enddo	
temp=sumv2/(3*npart)	Instantaneous temperature
etot=(en+0.5*sumv2)/npart	Total energy per particle
return	
end	

This pseudocode is only very generally applicable to all molecular dynamics programs. Many enhancements are needed to create physically relevant output.

### **1.3.3 Replica exchange molecular dynamics**

In attempting to effectively sample the potential energy surface of a molecular system, one is challenged by an extremely rugged, multidimensional landscape that owes its shape to the amount of degrees of freedom that are an innate part of that system. The search for most populated conformational states typically goes hand in hand with looking for global minima located within the potential energy surface. The complexity of the surface therefore, makes it possible that a molecular dynamics approach to a conformational search, at a single given temperature, can result in the discovery of a local minimum from which there is little chance of escape (given the energy distribution about the given temperature). Replica exchange molecular dynamics (REMD) provides a methodology providing a means by which to escape local minima thus allowing for more complete searching of the conformational space.<sup>67</sup>

In REMD, multiple molecular dynamic models are run at different temperatures each of which will have different energy distributions. Spacing the temperatures at close enough intervals ensures a probability exists that one trajectory will sample an energy configuration very close, if not identical, to another. Subject to Metropolis algorithm, when this happens, an atom by atom velocity exchange between replicas is implemented so that each replica, in effect, is given the opportunity to sample the others energy space. In classic REMD, the Metropolis criteria is applied to replicas with adjacent temperatures, in other variants such as TIGER2<sup>68,69</sup> the testing is applied globally.

### **1.3.4 GRAMM-VINA**

In molecular docking programs, the general desire is to ascertain the binding competency of a small molecule ligand to a larger, macromolecular receptor. This

competency is in turn described as a family of complex poses that are assigned a numerical score that is in turn based on a series of descriptors with some kind of basis in physical reality (e.g., geometrical fit, intermolecular hydrogen binding, electrostatic effects, and hydrophobic interactions). Often, these scoring functions are combinations of weighted descriptors that are tailored to specific kinds of intermolecular interactions and are associated with known experimental data. Finally, certain assumptions are usually required for these analyses whereby the ligand is given a certain number of rotating bonds, the ligand and/or binding site is in a given charge state, and a search space is defined where the ligand is allowed to adopt certain positions relative to the receptor.

In course-grained methods such as GRAMM,<sup>70-75</sup> both the ligand and receptor are held rigid and are subject to a rough discretization in the form of a step potential that describes their molecular surfaces. The receptor is then held fixed while the ligand is moved in 6-space (x,y,z translations  $\theta, \varphi, \psi$  rotations) around it in an exhaustive manner subject to user-defined translational and rotational gradations. The full complement of 6-space can be explored in a rapid fashion owing to the exploitation of a Fast Fourier Transform (FFT) algorithm. The intermolecular surfaces are checked for overlap and then scored on what is essentially a geometric fit based on a Lennard-Jones step potential. The end product is a series of scored ligand-receptor poses that allows for potential quick hits or informs later, more detailed computational analyses.

In more detailed types of docking like AutoDock VINA,<sup>76</sup> the step potentials are replaced by atomistic descriptions and rigid, ligand-receptor structures are replaced by models allowing for varying degrees of movement.

**Table 2. VINA parameters**

Weight	Term
−0.0356	gauss <sub>1</sub>
−0.00516	gauss <sub>2</sub>
0.840	Repulsion
−0.0351	Hydrophobic
−0.587	Hydrogen bonding
0.0585	N <sub>rot</sub>

The scoring function for AutoDock Vina is derived from X-score and PDBbind that leverage both knowledge-based and empirical scoring. The interaction functions are defined relative to a surface distance:  $d_{ij} = r_{ij} - R_{t_i} - R_{t_j}$ ,

$$f_{t_i t_j}(r_{ij}) \equiv h_{t_i t_j}(d_{ij})$$

where  $R_t$  is the van der Waals radius of atom type  $t$ ,  $r$  is the interatomic distance, and  $h$  is the distance dependent scoring function. Terms included in the overall scoring function include steric, hydrophobic, and hydrogen bonding with weights given in Table 2, the steric terms are given by:

$$\begin{aligned} gauss_1(d) &= e^{-(d/0.5 \text{ \AA})^2} \\ gauss_2(d) &= e^{-((d-3 \text{ \AA})/2 \text{ \AA})^2} \\ repulsion(d) &= \begin{cases} d^2, & \text{if } d < 0 \\ 0, & \text{if } d \geq 0 \end{cases} \end{aligned}$$

The hydrophobic term equals 1 when  $d < 0.5 \text{ \AA}$ ; 0 when  $d > 1.5 \text{ \AA}$ , and linearly interpolated in between. The hydrogen bonding term equals 1 when  $d < -0.7$ ; 0 when  $d > 0$ , and linearly interpolated in between. Metals are considered as hydrogen bond donors and all interaction functions are cut off at  $r_{ij} = 8 \text{ \AA}$ . The optimization of the poses, that is in turn, based on the scores, derives from the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method that takes into account both scores and their gradients.

### 1.3.5 Cluster analysis

Data clustering is an analytical technique by which data set pattern features are meaningfully arranged into groups. These groups, in turn are used to provide insight into a problem at hand or a system under study. In astronomical analyses, for example,

investigators might be tasked with better understanding of astrophysical phenomena that may correlate to the observed galaxy morphologies. These galactic morphological measures lend themselves well to cluster analysis. Another example is chemical species clustering where groups of chemical species might be clustered by drug efficacy and molecular weight as a means to tease out better pharmacophore leads. Yet another example finds cluster analysis widely used in genetic phylogeny determination amongst various biological species. Regardless of the use, the chief advantage of cluster analysis is its use in problems involving *exploratory* data analysis as opposed to alternate methods that are more suited to *confirmatory* data analysis (e.g., linear regression, hypothesis testing, any method utilizing *a priori* knowledge or assumptions).<sup>77,78</sup> Another way to consider cluster analysis is as an unsupervised classification method whereby patterns are sought in unlabeled data. Conversely, supervised classification or learning occurs where a set of data is analyzed relative to a set of training (labeled) data to determine meaning. Algorithms utilizing supervised data methods include those using neural nets or support vector machines.

Generally, cluster analysis can be thought of consisting of three primary steps.<sup>79</sup> The first step is pattern representation. In pattern representation, the size and features of the pattern(s) are taken in to account along with an optional step of further feature derivation or extraction in which more meaningful descriptors are defined from the data at hand. The second step involves some measure of mutual proximity of patterns that are appropriate to the overall data domain. These proximities (or more intuitively, similarities) may consist of any host of measures the simplest, and most familiar being the Euclidean distance. The selection of these measures can be as critical to the overall

cluster analysis as the clustering algorithm itself in attributing meaning to the output. The third and last step is the clustering or grouping algorithm itself. The algorithms employed in cluster analysis are many and may be classified as being either hierarchical or partitioning algorithms.

In hierarchical algorithms, groupings nest in a pyramidal fashion dictated by pattern proximity measures.<sup>77</sup> The output of these analyses form dendrograms or trees familiar to those looking at taxonomical classifications. In partitioning algorithms, the output data tends to form a single partition or group (as opposed to a nested dendrograms) that is mutually exclusive (fuzziness factors may be included, but are neglected here for descriptive purposes). Partitioning algorithms tend to be used for larger data sets where a nested dendrogram structure is less compact and harder to interpret and/or communicate.

With regards to all clustering techniques, certain issues apply that should also be considered.<sup>77</sup> Starting with a single cluster then subdividing or, alternatively, starting with multiple clusters and merging (both subject to some proximity measure and predetermined stopping point) relate to what are called divisive and agglomerative approaches, respectively. Whether the data are subject to sequential or simultaneous grouping are referred to as monothetic or polythetic clustering methods. With regards to cluster edges either hard or fuzzy boundaries may be employed where an additional layer of information is added for fuzzy members that possess degrees of membership in more than one group. Last cluster methods may be stochastic or deterministic where initial partitioning is driven and/or propagated via initial cluster guesses or optimization of clusters to a function (e.g., squared error function). Any or all of these items may appear

in either hierarchical or partitioning algorithms. Care should be taken when selecting a clustering method to ensure its relevance to the system under study and to understand the method's inherent influence on the interpretation of the resultant data.

For the purpose of conformational analysis of proteins, many degrees of freedom may be of interest to the investigator. The gross fluctuations of the quaternary and tertiary structure may be captured using backbone root mean square deviation analysis (RMSD).<sup>80</sup> In this method, the “backbone” is comprised of protein polypeptide bond and alternating alpha carbon atom spatial coordinates while purposefully neglecting side chain atoms. The RMSD is calculated for one structure versus another structure via the following equation:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}$$

where  $N$  is the number of pairs of equivalent atoms, and  $\delta$  is the distance between them.

Extension to three space between, say, two sets of  $\eta$  points  $v$  and  $w$  becomes:

$$RMSD(v, w) = \sqrt{\frac{1}{\eta} \sum_{i=1}^{\eta} (v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2}$$

The significance of RMSD analysis with regards to clustering algorithms is the opportunity provided to explore conformational space. In this situation, a protein structure is subject to a series of molecular dynamics simulations that act as an engine to generate trajectories that contain a multitude of physically realistic conformations. These conformations are then run through an algorithm that produces arrays of RMSDs that act



as features by which to cluster conformations. More specifically, and germane to this study, the Jarvis-Patrick algorithm can be used to perform this operation.<sup>81</sup> In this routine, similarity is governed by a distance metric that is the RMSD. Also, two different parameters are selected, **Neighbors to Examine** and **Neighbors in Common**. The **Neighbors to Examine** parameter determines the number of item's neighbors to take into account when assessing the number of mutual neighbors shared with another item. The value must be a minimum of two. The **Neighbors in Common** parameter determines the minimum amount of mutual nearest neighbors two items must have for them to be in the same cluster. The value must be at least one and be larger the value of the **Neighbors to Examine** parameter. The algorithm partitions the input into non-hierarchical clusters and results in an output with non-overlapping clusters.

Another clustering technique that is frequently used in data analysis is principal component analysis (PCA).<sup>79,82-84</sup> PCA is both exploratory and unsupervised where many variables are correlated against each other by performing eigen-decompositions of covariance matrices. The goals of PCA are to simplify the understanding of how the variables correlate, reduce and or compact the original data set, and discriminate between variables of more/lesser importance. Of relevance to this work, conformational space of smaller peptides is a prime candidate for analysis by this technique. For example, given a peptide of  $N$ -mers a covariance matrix of  $2N \times 2N$  dihedral angles may be constructed. This matrix becomes the basis by which more important correlations are derived amongst dihedrals which then allows for clustering of secondary peptide conformations.<sup>85</sup> Assuming a data set:

$$\tilde{x}_i, i = 1, \dots, m.$$

Compute the mean:

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m \tilde{x}_i$$

Next, compute the zero-mean data matrix:

$$x_i = \tilde{x}_i - \bar{x}$$

$$A = (x_1, x_2, \dots, x_m)$$

Now create the covariance matrix:

$$C = AA^T$$

Of note, the covariance matrix is symmetric and positive definitive, therefore the eigenvalues of the matrix will be real and non-negative.

From an eigenvalue problem:

$$Cv_i = \lambda_i v_i$$

Where in matrix form:

$$CV = V\Lambda$$

and

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$$

The eigenvalues are placed in decreasing order from  $\lambda_1$  to  $\lambda_N$ . The eigenvectors corresponding to the computed eigenvalues are:

$$V = (v_1, v_2, \dots, v_N)$$

So that the eigen-decomposition is computed as:

$$C = V\Lambda V^{-1} = V\Lambda V^T$$

At this stage, the “principal components” are those corresponding the higher value eigenvalues. The investigator at this point may examine further to see where truncation of less important factors might be reasonably applied.

## CHAPTER 2: IMPACTS OF PKA PERTURBATIONS ON THE CATALYTIC MECHANISM OF *PSEUDOMONAS AERUGINOSA* LACTONASE

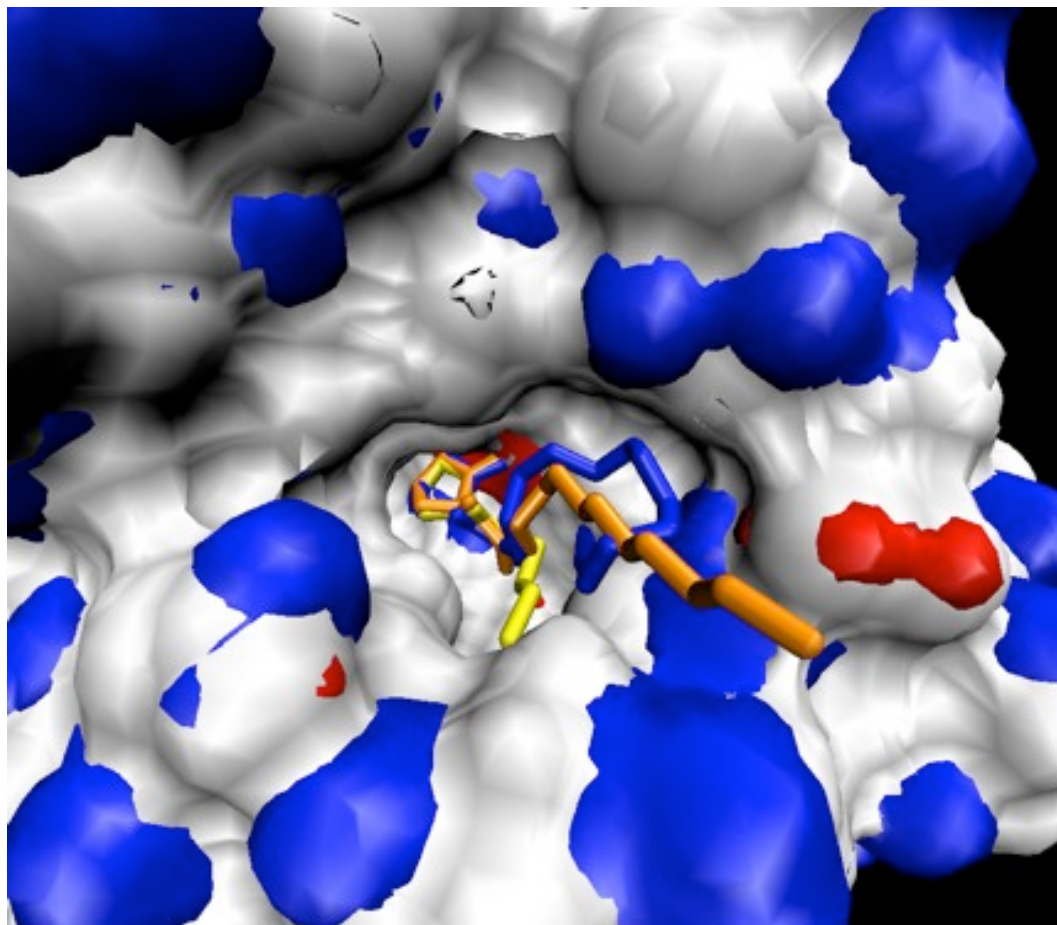
### 2.1 Computational methods

#### 2.1.1 Ligand construction

Protein coordinates were obtained from the crystal structure 3DHB.<sup>20</sup> The 3DHB structure was specifically used as it possesses a substrate well associated with the active site, thus indicative of an early product state. The substrate coordinates were extracted into SPARTAN<sup>86</sup> where, for the reactant state, the lactone ring was reformed and a constrained minimization was performed to retain the tail coordinates and lactone ring orientation relative to the zinc ions. A similar approach was used for the intermediate and product states. Also for the reactant states a hydroxyl ion was assigned to occupy the bridging position between the two zinc ions. For C6 and shorter AHLs, the original coordinates from the crystal structure were maintained while for C8 and longer AHLs, three candidate conformations (termed X, Y, and Z) were constructed so as to occupy likely poses in the binding cavity as shown in Figures 19 and 20. The coordinates from the 3DHB PDB file were imported into SPARTAN and methyl(ene) groups were added to achieve the desired chain length after which a molecular mechanics minimization was performed on only the methyl(ene) groups to ensure proper bond lengths and angles. These initial minimizations were performed with the AHL ligands in complex with the enzyme and with only the newly added methyl(ene) groups allowed to adjust to their environment. Last, the acyl tail was evaluated to ensure proper C-C bond lengths and angles (Figs 1 and 2). In order to study the role of protein backbone hydrogen bonding to the amide carbonyl group in the AHLs, the amide bond carbonyl groups (C=O) were

replaced with methylene (-CH<sub>2</sub>-) groups, while the amide nitrogens were left in a neutral state. Substrate topologies were generated by the PRODRG2.5 server,<sup>87</sup> while substrate partial charges were calculated through ESP fitting with SPARTAN using DFT with the 6-31G\* basis set.

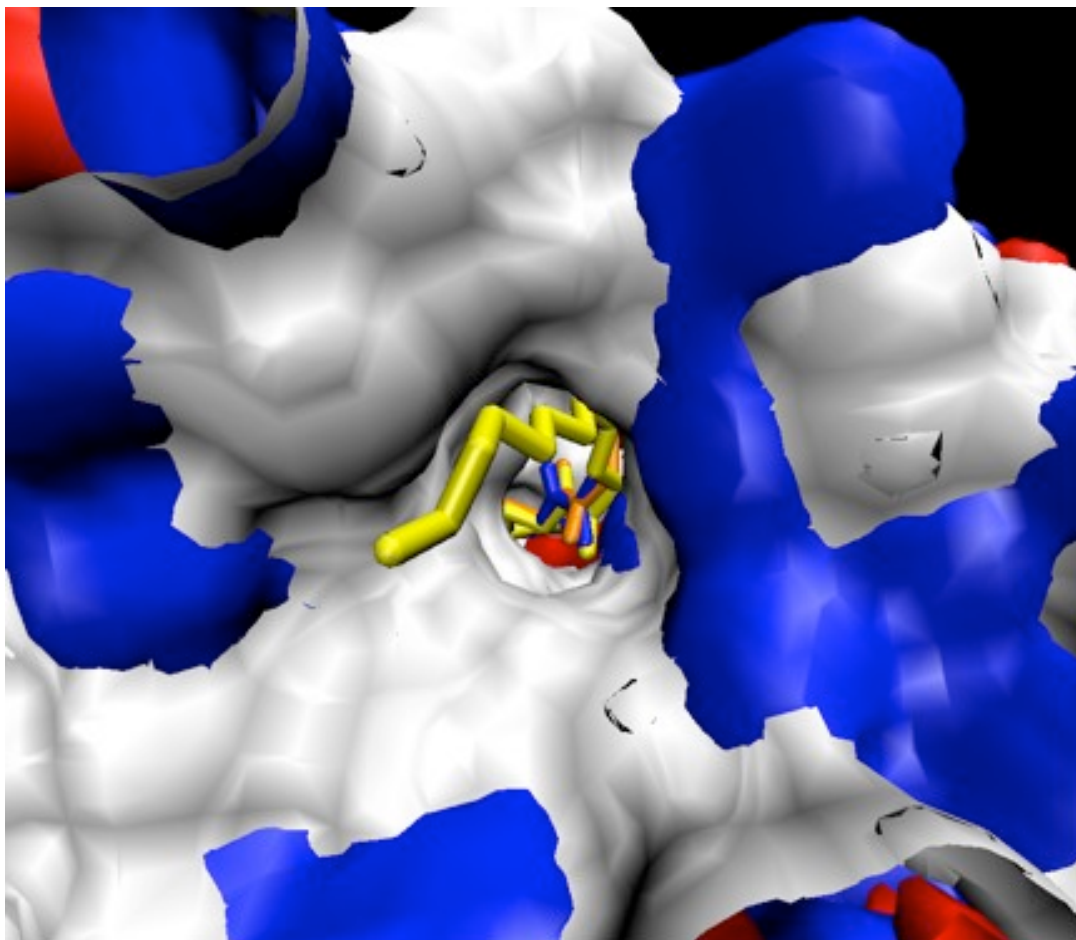
**Figure 19. AHL X, Y, and Z tail positions looking down into catalytic site**



---

The generation of different tail paths in this system is to provide for a sampling of conformational space that the alkyl tail might visit.

**Figure 20. AHL Z position threaded into available side pocket**



---

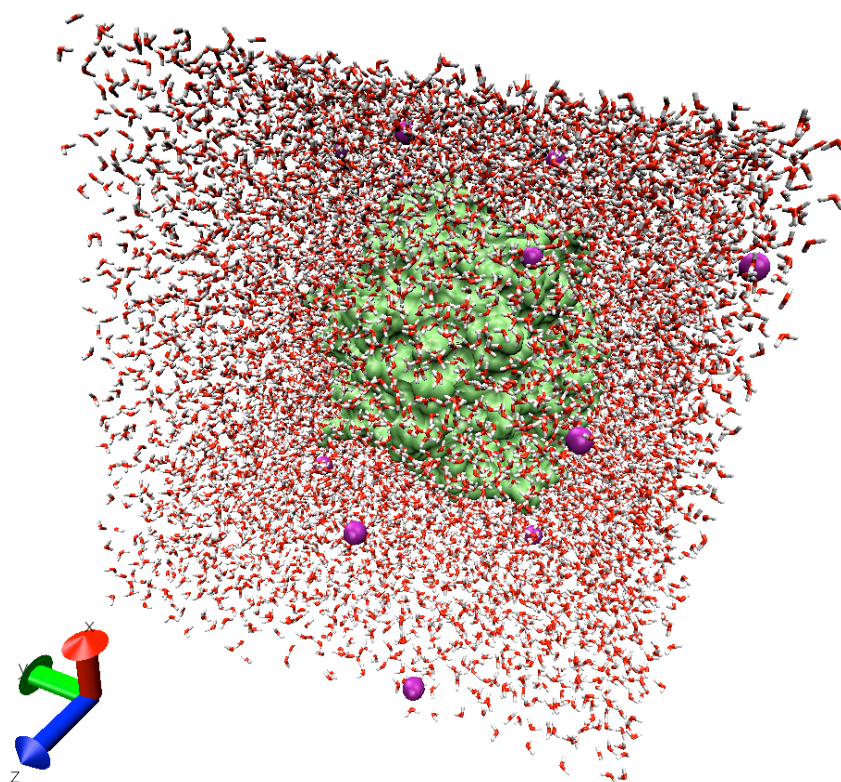
Alternate view of the Z conformation where the alkyl tail is threaded down a nearby pocket.

### 2.1.2 Molecular dynamics simulations and setup

All molecular dynamics simulations were run using GROMACS 3.3<sup>80</sup> with a GROMOS-96<sup>65</sup> force field for the protein. The protein was taken from PDB entry 3DHB from which all crystal waters were retained. The 3DHB structure was specifically chosen as it more clearly represents an early enzyme-product state. All protein substrate complexes were centered in a water box with simple point charge (SPC)<sup>88</sup> waters such that none were within 3.5 Å of any existing atom and water molecules extended 10 Å beyond the furthest protein atom in each Cartesian direction (Figure 21). The total system charge was neutralized using the appropriate number of Na<sup>+</sup> or Cl<sup>-</sup> counter-ions. An energy minimization was first performed to remove interatomic clashes using 500 steps of the steepest descent algorithm. Next, position restrained molecular dynamics was run for 100 ps allowing both the substrate and waters to relax about the restrained protein. A set of 1.0 ns molecular dynamics simulations were carried out for each of the ligand-protein complexes. An additional 4.0 ns production run was performed for both the ligand and reduced ligand-protein complexes. All simulations were run at a temperature of 300K in the NPT ensemble with a 2 fs time step employing periodic boundary conditions and the LINCS algorithm<sup>89</sup> to constrain bonds to hydrogen atoms. Both Berendsen pressure and temperature baths<sup>90</sup> were employed with relaxation coefficients of 0.1 and 1.0 ps, respectively. Lastly, long-range electrostatic interactions were managed using particle-mesh Ewald methodology beyond a cutoff of 9.0 angstroms.



**Figure 21. 3DHB molecular dynamics system**



---

Cubic explicit solvent (water) box as utilized for the lactonase MD analyses. The lactonase enzyme is denoted as a surface in green, the chloride ions are denoted as magenta spheres. Waters are rendered in red and white.

### 2.1.3 Cluster analysis

Molecular dynamics trajectories of the 3DHB protein both with and without *n*-hexanoyl homoserine lactone were subject to cluster analysis. This was done in an attempt to capture random motion fluctuation effects on critical residue pKas. Ligand influence was taken into account as well. The Jarvis-Patrick algorithm was employed through the use of GROMACS's `g_cluster` program whereby the RMSD average structure was generated from each cluster and subject to pKa analysis. The number of nearest neighbors was set to one while the number of identical neighbors needed to form a cluster was set to seven.

### 2.1.4 pKa predictions and electrostatic mapping

The protonation states for all ionizable residues in the protein systems were predicted at a pH of 7.0 using the single-site method encoded in the UHBD program.<sup>57-59</sup> The calculations were conducted utilizing a dielectric constant of 80 for water and 20 for the protein. An ionic strength of 150 mM and an ionic radius (for the Stern layer) of 2.0 Å were used. The CHARMM22 polar hydrogen only force field was employed.<sup>66</sup> Hydrogen atoms were added to the molecular systems using the CHARMM program<sup>91</sup> to generate the neutral reference state for the pKa prediction procedure. Both ligand-bound and ligand-free states were analyzed. pKa predictions were also performed using the PROPKA program as a means to cross validate UHBD predictions. Electrostatic mapping was performed using APBS.

## **2.2 Results and discussion**

### **2.2.1 pKa of ASP108, ASP 109, TYR194, and HIS residues**

pKa predictions performed on the enzyme-substrate complexes of both C6AHL and its corresponding ring-opened product did not reveal any significant perturbations to the pKa's of the Y194 or D108 residues. In the cases where the C6AHL ligand was included, the molecule was treated as a non-ionizable residue. As the Y194 and D108 residues are considered candidate proton shuttles for the zinc-bridging water, it is noteworthy that the pKa study revealed no shift of their pKa's toward that necessary to deprotonate a metal-bound water (that being around 7.5). Y194 pKa's averaged  $11.6 \pm 0.1$  across ligated and unligated protein states while D108 and D191 titrated below their reference pKa of 4.4. That D108 and D191 remain in their anionic state is not surprising given their proximity to the zinc di-cations, thus validating their function as metal ligands. H104, H169, H104, H235, and H109 were also found to be in their neutral state, consistent with their role in metal ion coordination. No significant discrepancies were noted between UHBD and PROPKA results (Tables 3-5).

### **2.2.2 Electrostatic mapping**

Electrostatic mapping of the lactonase enzyme was performed on the lactonase enzyme without the complex substrate as it is improbable that it would have any effect on the binding site. The mappings are most easily viewed by projecting the electrostatic field onto the protein's solvent accessible surface. These constructs are often used to examine the surface for relevant geometric and electrostatic features. In this case, the lactonase is rendered with the active site both facing the reader ("front" – Figure 22) and rotated 180 degrees away from the reader on a vertical axis ("back" – Figure 23). The most

prominent feature noted is the highly positively charged region of the catalytic site containing the two  $\text{Zn}^{2+}$  ions.

**Table 3. PROPKA pKa analysis**

Residue	Model pKa	3DHB Crystal	Cluster				
			1	2	3	4	5
D108	3.8	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
D191	3.8	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H104	6.5	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H106	6.5	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H109	6.5	<b>8.47</b>	<b>5.29</b>	<b>5.18</b>	<b>5.08</b>	<b>4.68</b>	<b>4.9</b>
H169	6.5	<b>2.83</b>	<b>2.06</b>	<b>1.62</b>	<b>1.86</b>	<b>1.93</b>	<b>1.27</b>
H235	6.5	<b>6.2</b>	<b>2.98</b>	<b>3.92</b>	<b>1.62</b>	<b>5.66</b>	<b>3.08</b>
Y194	10	<b>11.69</b>	<b>11.56</b>	<b>11.84</b>	<b>12.02</b>	<b>11.93</b>	<b>11.71</b>

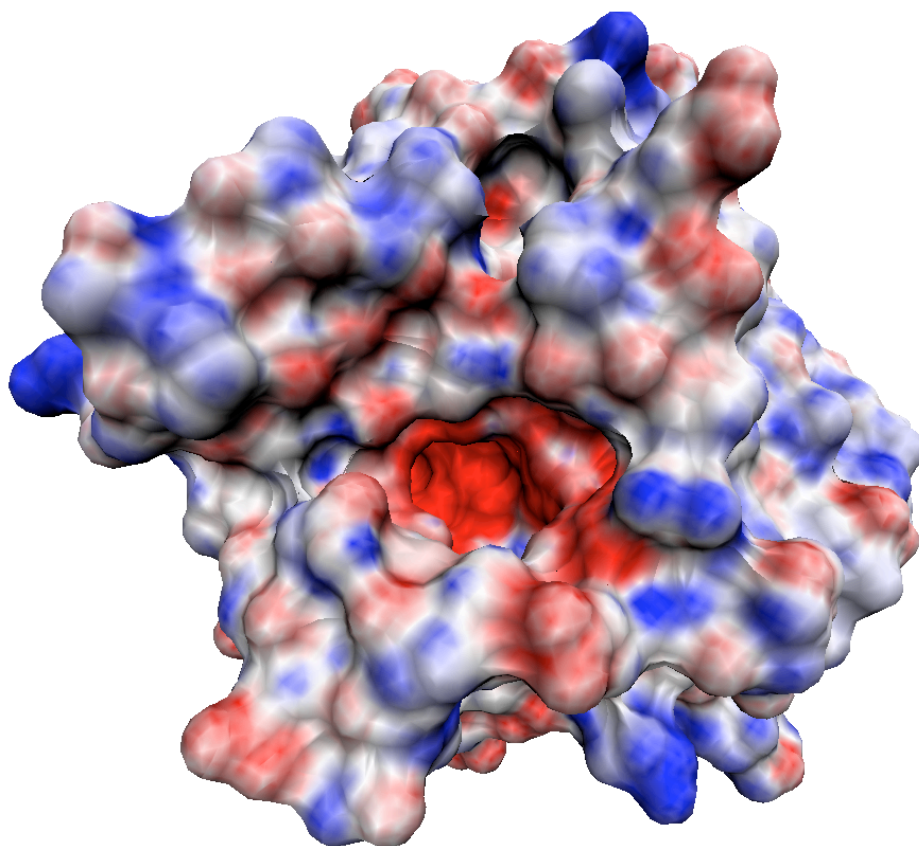
**Table 4. UHBD pKa analysis**

Residue	Model pKa	3DHB_xtal	Cluster				
			1	2	3	4	5
D108	4	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
D191	4	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H104	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H106	6.3	<b>3.127</b>	<b>1.717</b>	<b>2.257</b>	<b>2.082</b>	<b>2.004</b>	<b>2.443</b>
H109	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H169	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H235	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
Y194	9.6	<b>9.971</b>	<b>10.659</b>	<b>12.717</b>	<b>11.603</b>	<b>10.639</b>	<b>11.709</b>

**Table 5. UHBD pKa analysis - no ligand in protein**

Residue	Model pKa	Cluster				
		1	2	3	4	5
D108	4	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
D191	4	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H104	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H106	6.3	<<pKa	<b>1.349</b>	<b>2.857</b>	<b>2.857</b>	<b>2.163</b>
H109	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H169	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
H235	6.3	<<pKa	<<pKa	<<pKa	<<pKa	<<pKa
Y194	9.6	<b>11.963</b>	<b>11.121</b>	<b>11.409</b>	<b>11.409</b>	<b>11.942</b>

**Figure 22. Lactonase electrostatic mapping – front**

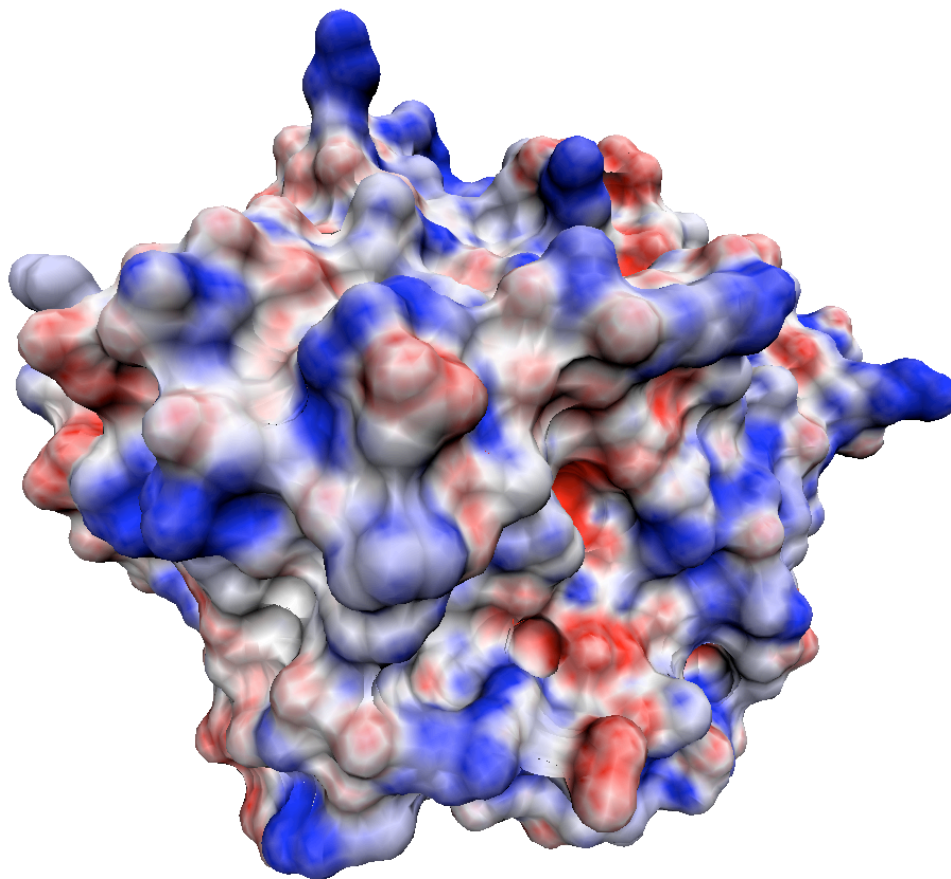


---

Of note is the catalytic site (red pocket) that displays a significant electrostatic charge.



**Figure 23. Lactonase electrostatic mapping – back**



---

Backside of the lactonase protein reveals no significant regions of electrostatic charge.

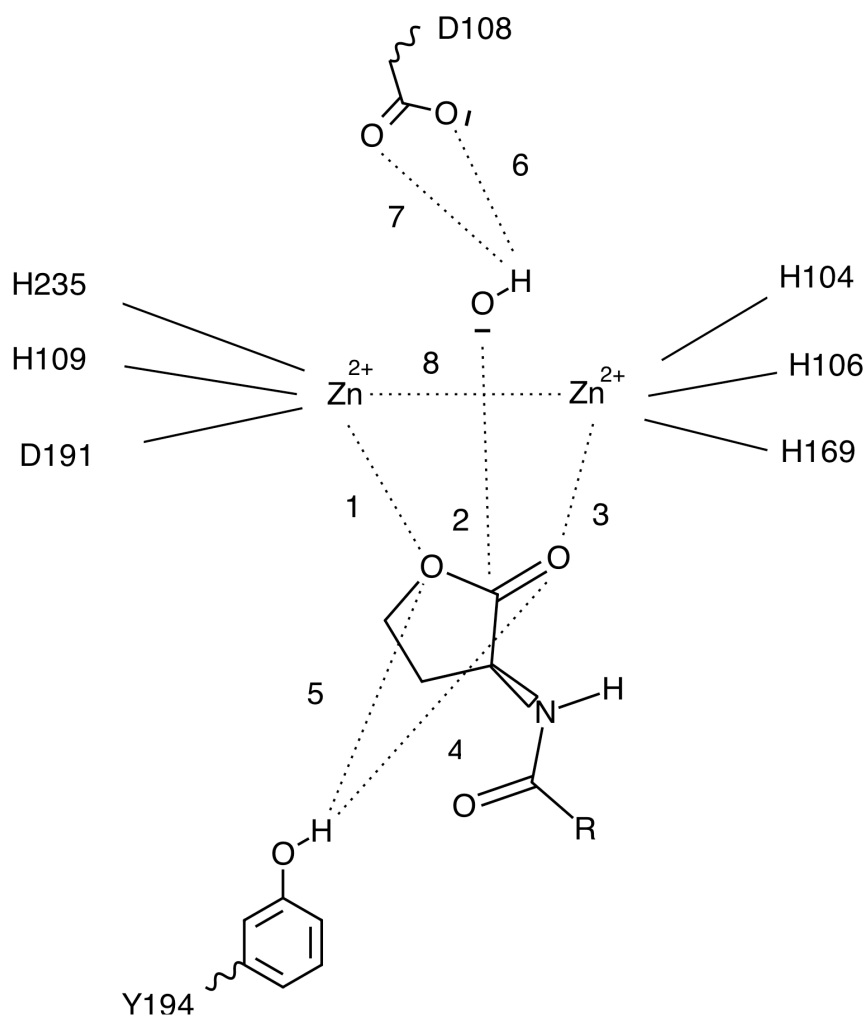
## **CHAPTER 3: IMPACT OF ACYL-HOMOSERINE LACTONE TAIL LENGTH ON CATALYTIC COMPETENCY AND BINDING/INTERACTION ENERGY**

### **3.1 Computational methods**

#### **3.1.1 Distance mapping**

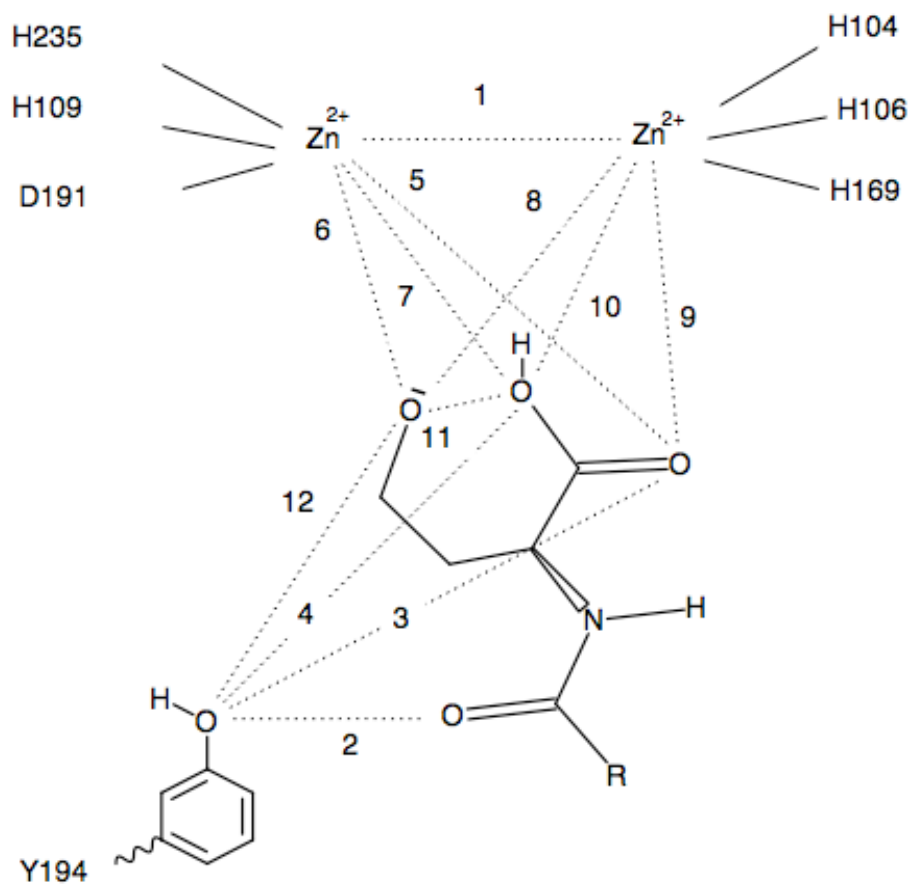
Molecular dynamics runs were set up and run as already described for pKa analyses. Distances across MD trajectories were calculated to ascertain catalytic competence and other residue contributions using GROMACS' g\_dist program.<sup>80</sup> Reactant, intermediate, and product state distances were all analyzed (Figs 24, 25, and 26, respectively). Specifically, distances 1, 2, and 3 of the reactant state were observed to monitor for the lactone ring bridging the zinc atoms. Distances 4 and 5 of the reactant state were observed to investigate the contribution of Y194 to ligand stabilization and/or early tetrahedral intermediate stabilization. Distances 6 and 7 of the reactant state were investigated to study how D108 might play a role in hydroxide ion stabilization. Intermediate and product state residue contributions were also investigated in a similar manner and mapped. Likewise, product and intermediate state oxygen to zinc ion distances were examined to determine how thermal variation effects substrate binding and release. Last, reactant, intermediate, and product state distances were examined for pre-transition state cues that either support or debunk current ideas in the overall mechanism.

**Figure 24. Measured distances for AHL reactant state**



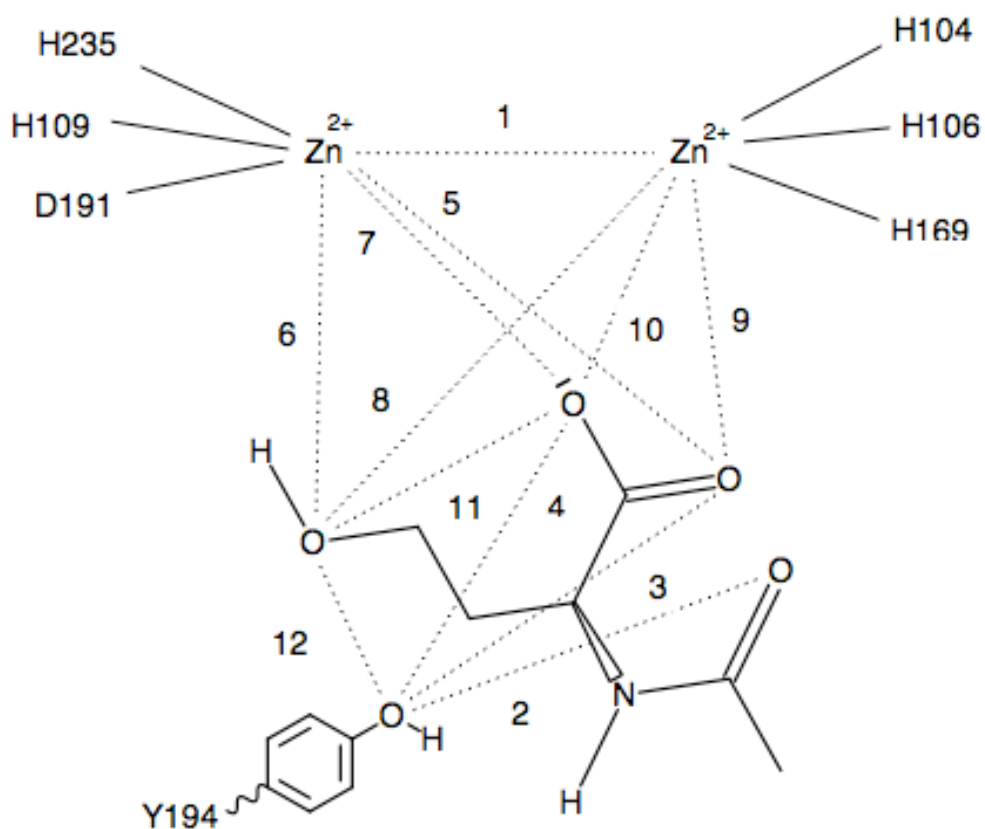
In the reactant state, distances of interest include the nucleophile to electrophile distance, the lactone oxygen distances to the zincs, and amino acid contributions to stabilization of the reactants.

**Figure 25. Measured distances for AHL intermediate state**



Of particular interest is distance 11, whose distance is consistent with intramolecular proton transfer.

**Figure 26. Measured distances for AHL product state**



In this figure the distances studied revealed a release of the alkoxide oxygen along with a promiscuity of the carboxylate oxygens between both zinc ions.

### 3.1.2 Binding free energy study

The binding energy calculations of the protein-substrate complexes were performed using the linear interaction energy (LIE) method.<sup>92</sup> In this computation, MD trajectories are generated for both the substrate in solution and for the protein in complex with the same associated substrate. These two distinct substrate environments define two separate thermodynamic states whose difference in turn approximates the free energy of binding. Both electrostatic and Lennard-Jones energy interactions were calculated for each state between the substrate and its environment and combined as follows:

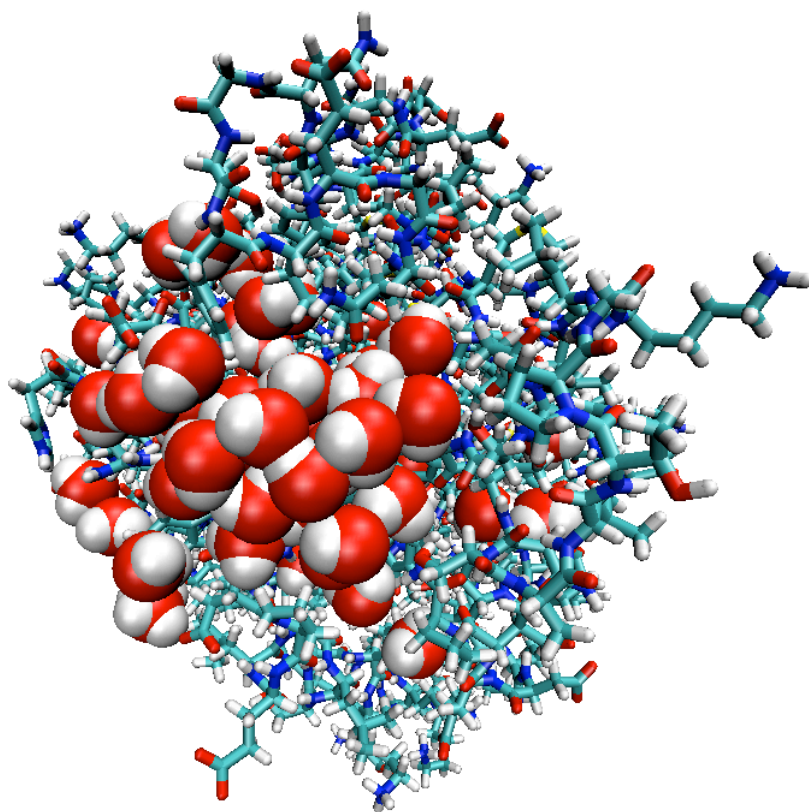
$$\Delta G_{\text{bind}} = \alpha(\langle V_{\text{l-s LJ}} \rangle_{\text{p}} - \langle V_{\text{l-s LJ}} \rangle_{\text{w}}) + \beta(\langle V_{\text{l-s el}} \rangle_{\text{p}} - \langle V_{\text{l-s el}} \rangle_{\text{w}})$$

Where  $\langle V_{\text{l-s LJ}} \rangle_{\text{p}}$  and  $\langle V_{\text{l-s LJ}} \rangle_{\text{w}}$  are the average Lennard-Jones energies of the protein-complexed and solvated substrates, respectively.  $\langle V_{\text{l-s el}} \rangle_{\text{p}}$  and  $\langle V_{\text{l-s el}} \rangle_{\text{w}}$  in turn represent the average electrostatic energy interactions of the protein-complexed and solvated substrates.  $\alpha$  and  $\beta$  are scaling constants for the Lennard-Jones and electrostatic energies. As the electrostatic  $\beta$  term is a derived value (equal to 0.5), the  $\alpha$  coefficient is the only value that is usually fit to experimental data for different families of substrates. In our case, an  $\alpha$  coefficient of 0.18 was used that was based on a broad family of ligands including ones containing the same elements as the AHL molecule under study. In addition, a sensitivity analysis using other  $\alpha$  coefficients confirmed no obscuring of the overall trend we desired to explore, that of the potential existence of a proportional response of binding energy to tail length.

### 3.1.3 Interaction energy study – Q

For van der Waals interaction energy analysis of protein-ligand interactions in the enzyme complex, the Q program<sup>93</sup> was utilized. For these cases, end-state conformations were obtained from the GROMACS production run trajectories as starting states for the Q trajectories. For these calculations the CHARMM22 force field as implemented by Q was used. A 15Å water-solvated boundary sphere, replacing crystal waters, were implemented and centered on zinc 252 to include the ligand and active site, using pdb structure 3DHB as a template (Figure 27). All residues outside the boundary sphere were retained with ionizable residues neutralized. A heat-up procedure was run for 160 ps from 50K to 300K with restrained ligand and protein followed by a 1.0 ns production run.

**Figure 27. Molecular dynamics system setup for Q**



---

Lactonase enzyme as set up in Q. Note how waters “cap” over the active site and are rendered as van der Waals surfaces. The substrate and active site in this figure are not seen as they are under the water cap.



### **3.1.4 Active site QM snapshots / optimization of a dizinc 2+ ligand field**

Single-point energy calculations on truncated substrate and active site residues were conducted with GAUSSIAN 03<sup>94</sup> using DFT with a B3LYP functional. The 6-31G(d,p) basis set was employed. The single point energy analyses were conducted on substrate poses that were observed to be well represented over all the reactant, intermediate, and product state trajectories. To observe the effect of zinc to zinc distance on the ligand field, two water, hexacoordinated zinc +2 ions were gradually moved in stepwise from 4.7 to 3.5 angstroms using B3LYP/6-31G\* level of theory.

## **3.2 Results and discussion**

### **3.2.1 Distance mapping – catalytic competence**

To achieve the proper bridging position for the reactant state, the substrate ring oxygens should orient across each zinc ion in the active site at a consistently close distance. The resulting geometry would in turn provide for a near-attack positioning of the hydroxide ion oxygen to the lactone ring carbonyl carbon thus promoting an optimal nucleophilic attack trajectory. Across the HSL, C4AHL, C6AHL, and the X and Z conformations of the C8-C12AHL species, distance 3 is seen as the shortest and most stable (Table 6). The stability and proximity derives from the fact that the carbonyl oxygen carries a more negative partial charge and least steric hindrance of the two lactone ring oxygen atoms. The Y conformation of the C8-C12AHL species shows a doubling of distance 3. Close inspection of the equilibration MD trajectories revealed that, in each case, a water molecule comes into close enough contact with the zinc atom so as to displace the carbonyl oxygen to the alternate zinc ion. This observation is not surprising as the SPC water molecule oxygen is more negatively charged than the

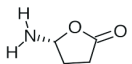
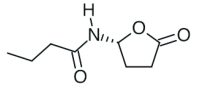
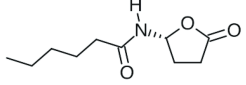
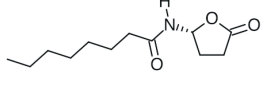
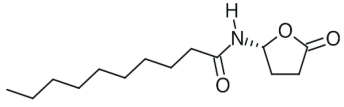
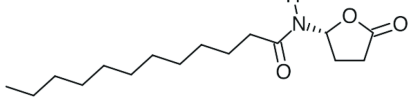
carbonyl oxygen. Distance 1 shows the lactone ring ether oxygen atom residing at 1.5 to 2 times the length of distance 3. This can be explained by the fact that the lactone ring ether oxygen carries less partial negative charge and is more sterically hindered than the carbonyl oxygen. The overall span of distance 1 across all ligands is consistent. The C8-C12AHL Y conformations for distance 1 exhibit less difference across each ligand as the ether oxygen and displaced carbonyl oxygens both ligate the same zinc ion. Distance 2 exhibits consistent and stable values spanning 3.1-3.7 angstroms across HSL, C4AHL, C6AHL, and the X and Z conformations of the C8-C12AHL ligands. These distances appear close enough to promote nucleophilic attack of the hydroxide ion on the carbonyl carbon of the lactone ring. The Y conformations of C8-C12AHL ligands reveal a notably longer distance 2. This longer distance, in addition to the loss of the bridging position of the lactone ring oxygens across the two zinc ions, suggests the loss of catalytic competence for the C8-C12AHL Y conformations. The zinc-to-zinc distance 8 remains constant across all ligands.

Distances 4 and 5 were analyzed through the MD trajectories to see how Y194 might contribute to stabilizing the lactone rings of the AHL ligands through hydrogen bonding to the carbonyl oxygens or ether oxygens, respectively. Distance 4 exhibited high variability across ligands with proton to oxygen distances well outside those of hydrogen bonding range. Distance 5 also exhibited high variability across ligands with proton to oxygen distances well outside those of hydrogen bonding range. Interestingly, a similar analysis performed on the Y194 oxygen to amide carbonyl oxygens revealed a trajectory with a tight, persistent, distance of  $2.4 \pm 0.1$  Å. The corresponding O-H-O angle distribution was consistent with that of a hydrogen bond throughout the trajectory.

This potential hydrogen bond was not seen in the initiating structure or in the energy minimization or soaking steps, but was formed early in the production simulation.

Distances 6 and 7 were examined to determine how D108 might influence the stability of the hydroxide anion. Distance 6 exhibited a fairly wide variance and longer distance when compared to distance 7. Distance 7 shows the hydroxide proton to aspartate oxygen distances that come to within a hydrogen bonding distance across many of the ligands and conformations. The distances appear to become more consistent through the C6-C12AHL species. Furthermore, these distances combined with the hydroxide bridging position across the zinc atoms makes for a stable network by which to ensure an incoming substrate is presented in the best position for nucleophilic attack, provided that the substrate can achieve its own bridging orientation and appropriate binding affinity.

**Table 6. Reactant state distance analysis**

Ligand	Distance (Å)			
	1	2	3	4
	4.1±0.9	3.7±0.2	2.1±0.1	4.1±0.5
	4.1±0.3	3.6±0.3	2.1±0.1	6.3±0.7
	4.2±0.3	3.1±0.1	2.1±0.1	3.2±0.3
	4.0±0.3	3.2±0.1	2.1±0.1	3.8±0.3
	3.8±0.3	3.9±0.1	3.8±0.2	3.8±0.3
	3.9±0.3	3.1±0.1	2.1±0.1	3.2±0.5
	3.7±0.4	3.6±0.3	2.1±0.1	4.8±0.8
	4.0±0.2	3.8±0.2	4.0±0.2	3.8±0.6
	3.8±0.3	3.1±0.2	2.1±0.1	5.9±1.2
	4.2±0.6	3.7±0.3	2.0±0.1	4.7±0.5
	4.1±0.1	4.1±0.1	4.1±0.2	3.5±0.4
	3.6±0.3	3.1±0.1	2.1±0.1	5.4±0.3

**Table 7. Reactant state distance analysis continued**

Ligand	Distance (Å)			
	5	6	7	8
HSL	3.3±0.9	3.7±0.1	3.8±0.2	3.5±0.1
C4-AHL	5.7±0.7	3.6±0.1	3.8±0.1	3.6±0.1
C6-AHL	2.4±0.4	3.0±0.3	2.2±0.3	3.5±0.1
C8-AHL	4.3±0.4	3.4±0.8	2.7±0.5	3.6±0.1
	4.3±0.4	3.4±0.8	2.7±0.5	3.5±0.1
	3.0±0.4	2.7±0.4	2.8±0.4	3.5±0.1
C10-AHL	3.7±1.1	3.6±0.1	3.8±0.1	3.5±0.1
	4.9±1.0	4.1±0.2	2.7±0.2	3.5±0.1
	5.3±1.4	2.5±0.5	2.4±0.6	3.6±0.1
C12-AHL	3.9±0.9	3.6±0.1	3.8±0.1	3.6±0.1
	4.1±0.5	3.9±0.2	2.4±0.2	3.5±0.1
	5.6±0.4	2.8±0.5	2.8±0.5	3.4±0.1

In both the intermediate and product states, the zinc-to-zinc distance 1 (Tables 8 - 11) is significantly longer and more variable than seen in the reactant state. One possible reason for this distance increase is that upon ring opening, the resulting alkoxide and carboxyl oxygens that bridge the two zinc ions are now four carbons apart and able to move around much more freely to allow for different modes of metal coordination. The broadening of the zinc-to-zinc distance also provides for more water access to the zinc ions. Whereas in the reactant state, the C8-C12AHL Y conformations only show a water molecule ligating to a zinc, in both the intermediate and product states. About 50% of all complexes show one or more waters accessing the carboxyl(ate) bound zinc ion suggesting at least part of a means for product release.

Another interesting observation is the alkoxide oxygen - carboxyl oxygen length shown in distance 11 of both the intermediate and product states. The intermediate state distances are significantly shorter than those seen in the product state which approach twice that of the later. This is significant because it suggests that the functional group ends are close enough to allow for an intramolecular proton transfer. Further investigation into this phenomenon reveals that the Y194 oxygen to intermediate state alkoxide oxygen distance 12 is close enough to create a stabilizing effect. The same stabilizing effect is not seen in the product state. When put together, these observations suggest that Y194 stabilizes the second transition state in addition to or instead of the first transition state.

**Table 8. Intermediate state distance analysis**

IS- Ligands	Distance (Å)					
	1	2	3	4	5	6
C4	4.8±0.8	6.9±0.6	5.5±1.0	5.7±0.8	6.7±0.6	1.9±0.04
C6	4.2±0.2	5.9±0.3	3.1±0.2	4.4±0.2	4.3±0.2	1.9±0.04
C8X	4.5±0.5	6.6±0.4	3.1±0.3	4.6±0.4	4.8±0.3	1.9±0.1
C8Y	4.8±0.2	6.6±0.5	3.2±0.4	4.6±0.5	4.8±0.2	1.9±0.04
C8Z	3.6±0.1	6.0±0.4	2.7±0.2	4.4±0.6	4.6±0.3	2.0±0.1
C10X	5.4±0.6	6.0±0.7	2.9±0.3	4.0±0.9	5.2±0.6	1.9±0.04
C10Y	4.6±0.3	6.5±0.9	3.3±0.3	4.8±0.5	4.6±0.3	1.9±0.04
C10Z	5.3±0.1	5.7±0.5	4.3±0.4	4.1±0.4	2.1±0.1	1.9±0.04
C12X	4.8±0.2	6.3±0.3	3.1±0.2	4.2±0.2	4.9±0.2	1.9±0.04
C12Y	4.8±0.2	5.7±1.0	3.2±0.3	4.5±0.4	4.8±0.3	1.9±0.04
C12Z	4.3±0.4	5.6±0.7	5.2±0.3	6.9±0.6	3.7±0.4	1.9±0.04

**Table 9. Intermediate state distance analysis continued**

IS- Ligands	Distance (Å)					
	7	8	9	10	11	12
C4	7.1±0.3	4.7±0.8	6.5±1.8	6.6±0.9	5.5±0.2	3.5±1.0
C6	3.3±0.2	4.1±0.1	2.0±0.1	2.2±0.1	3.1±0.1	2.8±0.2
C8X	3.9±0.6	3.7±1.0	2.0±0.1	2.6±0.7	3.2±0.4	3.4±0.6
C8Y	3.7±0.2	4.3±0.2	2.0±0.1	2.2±0.1	3.1±0.2	3.6±0.7
C8Z	5.8±0.7	2.0±0.3	2.1±0.1	4.0±0.3	4.3±0.4	3.0±0.3
C10X	3.8±0.3	4.5±0.3	2.0±0.1	2.6±0.6	2.9±0.4	4.2±0.7
C10Y	3.5±0.2	4.4±0.3	2.0±0.1	2.2±0.2	3.1±0.2	4.1±1.0
C10Z	4.2±0.1	5.9±0.2	3.7±0.1	2.2±0.1	4.4±0.2	4.5±0.4
C12X	3.2±0.2	4.3±0.2	2.0±0.1	2.2±0.1	3.1±0.2	2.9±0.3
C12Y	3.7±0.2	4.3±0.2	2.0±0.1	2.2±0.2	3.2±0.2	3.5±0.8
C12Z	3.8±0.5	3.8±0.5	2.1±0.3	4.1±0.4	4.2±0.2	4.7±0.4



**Table 10. Product state distance analysis**

PS- Ligands	Distance (Å)					
	7	8	9	10	11	12
C4	4.3±0.3	5.2±0.5	2.8±0.4	2.1±0.1	4.2±0.6	5.0±0.7
C6	5.1±0.2	4.4±0.2	2.2±0.1	2.1±0.1	5.7±0.6	3.6±1.1
C8X	3.9±0.1	4.6±0.1	2.4±0.1	2.1±0.1	6.9±0.4	5.3±0.7
C8Y	2.1±0.1	6.7±0.5	2.2±0.1	3.9±0.3	7.0±0.5	6.7±1.0
C8Z	2.1±0.1	7.0±0.4	2.1±0.1	3.7±0.1	7.3±0.3	7.0±1.1
C10X	3.8±0.1	4.6±0.1	2.3±0.1	2.1±0.1	7.1±0.4	5.4±0.7
C10Y	2.1±0.1	7.3±0.3	2.1±0.1	3.7±0.1	7.0±0.5	6.3±0.5
C10Z	2.1±0.1	6.0±0.7	2.2±0.1	4.1±0.2	7.1±0.4	5.8±1.1
C12X	4.0±0.2	4.6±0.1	2.3±0.1	2.1±0.1	7.1±0.4	4.6±0.9
C12Y	2.1±0.1	7.8±0.7	2.5±0.6	3.9±0.6	7.0±0.4	8.6±1.7
C12Z	2.1±0.1	6.8±0.3	2.2±0.1	3.6±0.1	7.0±0.4	7.3±0.8

**Table 11. Product state distance analysis continued**

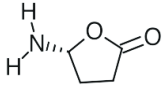
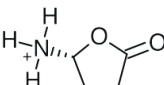
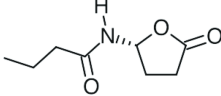
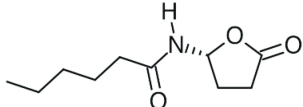
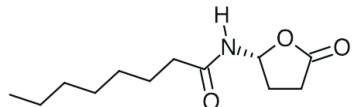
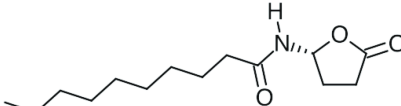
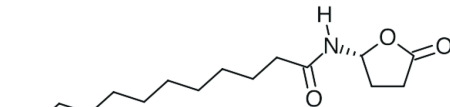
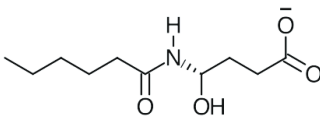
PS- Ligands	Distance (Å)					
	1	2	3	4	5	6
C4	4.7±0.4	4.5±0.6	3.8±0.6	2.9±0.2	2.3±0.4	2.1±0.1
C6	4.9±0.2	6.7±0.8	4.8±0.8	3.6±0.5	3.7±0.2	2.0±0.1
C8X	3.8±0.1	5.9±0.7	5.5±0.5	3.8±0.4	2.3±0.1	2.0±0.1
C8Y	4.9±0.3	5.0±1.0	4.6±0.5	4.6±1.1	3.4±0.3	4.5±0.5
C8Z	5.1±0.1	6.6±0.6	5.0±0.5	5.2±0.6	3.9±0.1	4.5±0.3
C10X	3.8±0.1	6.5±0.7	5.9±0.6	4.1±0.5	2.3±0.1	2.0±0.1
C10Y	5.2±0.1	5.6±0.3	4.4±0.3	4.4±0.4	3.9±0.1	4.5±0.3
C10Z	5.0±0.2	6.6±0.9	4.8±0.5	5.4±1.0	3.5±0.2	3.2±1.5
C12X	3.8±0.2	7.1±0.7	5.1±0.7	3.7±0.5	2.3±0.2	2.0±0.1
C12Y	4.3±0.3	7.1±0.8	5.3±1.2	5.4±1.6	3.0±0.4	5.9±0.8
C12Z	4.9±0.2	7.6±0.7	5.8±0.6	5.6±0.6	3.7±0.2	4.8±0.5

### 3.2.2 Free energy analysis – reactant state and perturbation

A binding free energy study was performed across all ligands including HSL and the ring-opened product to gain a fuller understanding of how they might better inform us of the discrepancy seen between the two sets of bioassay data (Table 12). For those ligands where there may exist, multiple viable conformations (e.g., X, Y, and Z), the  $\Delta G$  values were averaged to provide a better picture of what might be happening across conformational space. The C4-C12AHL ligands exhibit only a very moderate increase in binding free energy throughout the series. The increase in binding free energy is due to the increasing Lennard-Jones interaction free energy components afforded by the incrementally longer alkyl tails of the AHL molecules. This trend, in turn, suggests a favoring of the reference 2 bioassay data in Table 4. To better understand the HSL bioassay data, both protonated (HSL(+)) and neutral ligands were examined as it is probable that the HSL amine group is protonated at pH 7. The initial MD trajectory of the HSL(+) ligand led to an unbound state, therefore, additional MD runs were carried out for both the HSL and HSL(+) complexes to ensure better sampling. The trajectories of the HSL complex exhibited a well-centered spread of binding free energies whose average matches well with the AHL species. This observation is inconsistent with the lack of activity seen for HSL in the bioassays. However, for the three trajectories of the HSL(+) complex, one trajectory is significantly weaker in binding free energy than the AHLs while another trajectory is clearly unfavorable. That one of the three HSL(+) trajectories is energetically favorable is not surprising as during this trajectory its protonated nitrogen retained a shell of water molecules. It therefore seems possible these waters stand a

chance to diffuse past the +1 charged HSL(+) and more favorably associate with either of the +2 zinc ions, thus, displacing the ligand.

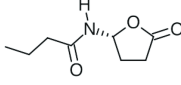
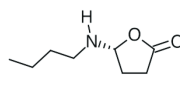
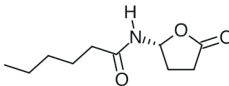
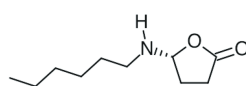
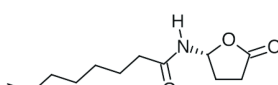
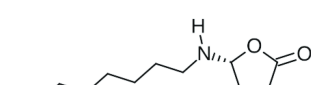
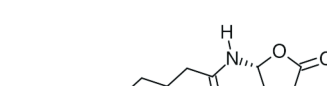
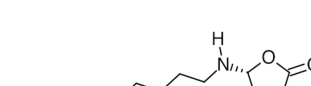


**Table 12. Reactant state binding free energies (kcal/mol)**

Substrates	Activities %			LIE $\Delta G$
	Kim et al. <sup>10</sup>	Wang et al. <sup>5</sup>		
	HSL	ND <sup>a</sup>	5.40	-5.88±0.38 -5.28±0.28 -6.44±0.54
	HSL(+)	—	—	Note 1 -7.66±0.41 -3.83±0.81
	C4AHL	61.4	89.1	-5.90±0.26
	C6AHL	100.0	100.0	-6.74±0.36
	C8AHL	116.6	93.9	-6.50±0.61
	C10AHL	217.1	90.3	-6.49±1.37
	C12AHL	31.6	—	-7.15±0.54
	C6-Ring-Opened	—	—	-28.46±1.59

<sup>a</sup> First trajectory led to an unbound state.

In an attempt to capture the relevance of the F107 residue to the binding free energy, an approach was taken whereby the AHL amide carbonyl group was reduced to a methylene group. It was thought that this might disrupt energetically favorable hydrogen bonding that occurs between the carbonyl oxygen and the F107 amide hydrogen or through an intervening water molecule. In this study, the nitrogen of the amide bond was left with its amide partial charge and neutral as the effect we were searching for was the loss of the carbonyl oxygen, not the gain of a protonated nitrogen. Both the resulting nitrogen and methylene carbon were then changed from sp<sup>2</sup> to sp<sup>3</sup>-hybridized atoms to reflect the loss of the amide bond. The differences in free energy between the natural and reduced species did not show a decrease, but rather an increase in free energy of binding across all ligands (Table 13). While this does not disprove the contribution of the backbone of F107 to the binding free energy, it may show how the reduced alkyl amine is less soluble in water thus making protein binding more favorable. In addition, the sp<sup>3</sup> hybridized carbon and nitrogen atoms of the reduced AHL molecules allow them to fit more intimately within their respective binding pockets than the natural sp<sup>2</sup> hybridized ligands.

**Table 13. LIE  $\Delta G$  (kcal/mol) values of signal ligands versus their reduced forms.**

Ligands	$\Delta G(\text{Ligand})$	Reduced Ligands	$\Delta G(\text{Reduced Ligand})^a$
	$-5.90 \pm 0.26$		$-8.03 \pm 0.27$
	$-6.74 \pm 0.36$		$-7.65 \pm 0.19$
	$-6.50 \pm 0.61$		$-8.52 \pm 0.63$
	$-6.49 \pm 1.37$		$-8.87 \pm 0.37$
	$-7.15 \pm 0.54$		$-8.81 \pm 0.68$

<sup>a</sup>  $\text{C=O} \Rightarrow \text{-CH}_2\text{-}$  and the resulting amine was kept in its neutral state.

### 3.2.3 Interaction energy analysis

An interaction energy study was performed across all reactant, intermediate, and product states to gain a fuller understanding of how they might better inform us of the discrepancy seen between the two sets of bioassay data. As our primary interest was the potential for tail-mediated bioactivity, a series of analyses was conducted on the van der Waals interaction energies across all ligand-enzyme complexes using the Q program (Table 5). As can be seen from the data, the overall van der Waals contributions for the ligands are small, especially relative to electrostatic contributions of the ligand-enzyme complex. These electrostatic interactions can be expected to be governed by the  $\text{Zn}^{2+}$  ions with interaction energies of at least an order of magnitude higher than the more stable van der Waals interactions of the reactant states. Moreover, the changes in van der Waals interaction energy of the reactant states per mole per carbon addition were calculated to be 0.43 kcal/mol, agreeing well with 0.38 kcal per mole per carbon addition seen in *n*-alkyl alcohols in water/*n*-octanol partitioning experiments. Lastly, the per mole per carbon energy addition agrees well with previous bioassay approximations across the same species where the energy delta was calculated to be 0.24 kcal per mole per carbon increment.<sup>22</sup> The van der Waals interaction energies of the intermediate and product states, while noisier due to their anionic states, offer no information to refute trending observations made on the reactant states (Table 14).



**Table 13. van der Waals interaction energies (kcal/mol)**

AHL	RS	IS	PS
C4	−13.6	−4.5	3.9
C6	−12.9	−5.1	−0.6
C8X	−13.2	−0.2	−3.7
C8Y	−12.9	2.3	−2.7
C8Z	−16.3	0.4	−4.5
C10X	−14.3	−4.6	−0.3
C10Y	−14.5	−3.5	−5.7
C10Z	−16.8	−0.7	−4.2
C12X	−21.6	−4.4	−4.2
C12Y	−12.5	−0.4	−2.8
C12Z	−17.0	2.4	−3.3

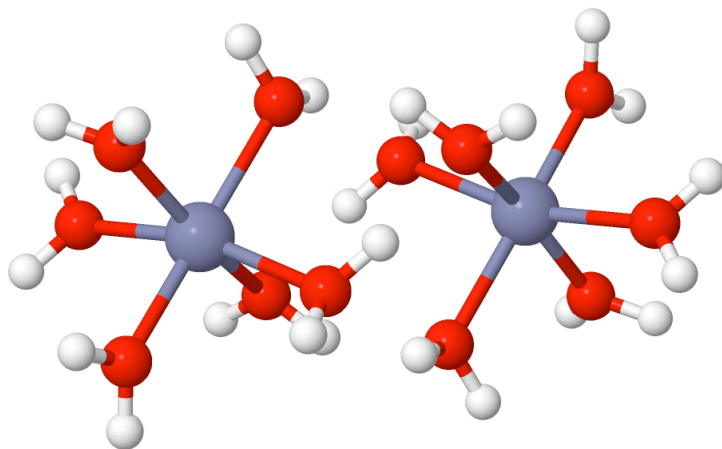
### 3.2.4 Geometry and energy of dinuclear zinc (2+) ligation field

Another point of interest was to consider what other interactions might be occurring between the substrate head (lactone or n-hexanoyl homoserine) and the active site. To that effect, all MD trajectories across all states we observed with a focus on substrate head orientation relative to the active site. Snapshots of poses that were observed to be distinct in space and sustained in time (greater than 1.0 ns) were extracted. The 20 poses across (5-7 poses per state) all states were subject to B3LYP/6-31G\* single point energy calculations to provide for a rough energy sampling and trending across all states. The most discernable trend was a significant clustering of all reactant states as being higher in energy than that of any of the intermediate or product states. An investigation of interatomic distances revealed that the zinc to zinc ion distances were on average  $3.5 \pm 0.1$  Å for the reactant states while the same distance for the product and intermediate states was  $4.7 \pm 0.7$  Å. That the intermediate and product state poses were found more stable under these zinc to zinc distances is reasonable given the additional degrees of freedom of the substrate relative to the reactant states allowing for better ligation of the zinc ions. These observations led to general questioning about the ligation field of the two zinc 2+ ions.

Previous computational studies have served to discuss the energies and optimized geometries of various zinc 2+ ion hydration states. In addition, it has been shown experimentally that in water, zinc 2+ tends to be hexahydrated. To understand what happens when two hexahydrated zinc 2+ ions are brought into close proximity (relevant to our active site observations), optimizations were carried out at the B3LYP/6-31G\* level of theory. These optimizations were conducted with the zinc ions constrained at

both 4.7 and 3.5 Å relative to each other. The difference in energy between the two distance constraint optimizations was  $-0.38$  kcal/mol relative to the 3.5 Å state. More importantly, while the coordination state of each of the zinc ions was hexacoordinated in the 4.7 Å model (Figure 28), the 3.5 Å model adopted a pentacoordinated state with each zinc ion cluster excluding a water to the secondary solvation shell (Figure 29). While it is recognized that the zinc ions in the lactonase active site see a different ligation field, it is also recognized that the ligands are in fact stronger than water overall and thus may indeed display stronger reorganization effects. This reorganization potential coupled with observations made that all the poses all show hexacoordinated zinc ions suggest that reorganization energy and ligand exchange plays an important role in this enzyme's mechanism of action. Furthermore, this ligand exchange was found amongst the pose sampling data where H104, 106, and 235 were found to stray well away from first solvation shell, and Y194 was found to move well within first solvation shell distance all while preserving each zinc ion's hexacoordination in the active site. These ligand field observations are new and have not been considered in alternate computational studies where either constrained QM calculations preclude observing a ligand exchange effect or QM/MM simulations are not considered across a family of ligands and chemical states.

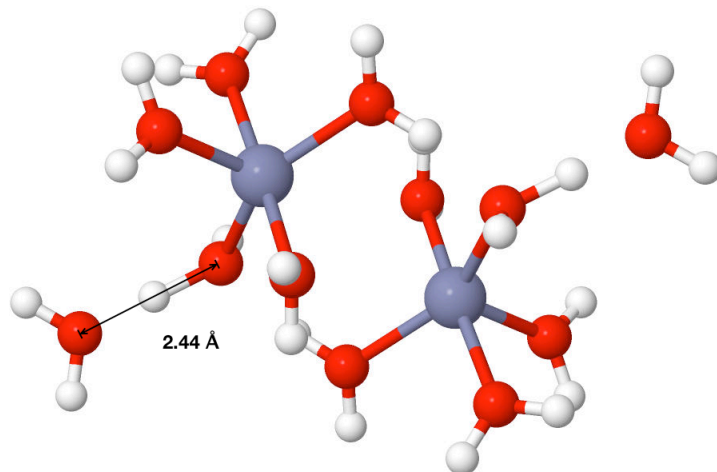
**Figure 28. Hexacoordinated zinc 2+ ions at 4.7 angstroms**



---

Note retention of hexacoordinated state at a zinc-to-zinc ion distance of 4.7 angstroms.

**Figure 29. Pentacoordinated zinc 2+ ions at 3.5 angstroms**



---

Shift to pentacoordinated state displaces two water molecules to secondary solvation shell thus creating an acidic water species in the first solvation shell.

### 3.3 Conclusions

The AiiA lactonase enzyme produced by *B. thuringiensis* exhibits bioactivity toward the quorum signaling molecules of Gram-negative bacteria. This activity appears to be an evolved mechanism where one organism competes with another. The quorum signal degradation occurs through action of the AiiA enzyme in what is seen as an activity that may be tail-length dependent. The bioassay data however, are inconsistent and the key interactions are not well understood. To gain more insight into the enzyme-substrate interactions, we performed multiple MD simulations across a family of Gram-negative quorum signaling molecules, acyl homoserine lactones. These simulations were careful to include the AHLs in their reactant, intermediate, and product complexed states.

We observed consistent catalytic competence with regard to geometric distances across the lactone ligands with the Y-conformation of the C8-C12AHL ligands standing out as one that demonstrates how important water exclusion is to enzyme function. Water exclusion is a critical and common theme in enzymology and so it is of no surprise that some conformations may be favored over others to this effect. The tail appears to serve as a “plug” barring water access to the active site. It was further observed that half of the intermediate and product state retained water at the zinc ions. This was facilitated by the ring-opened ligand allowing the enzyme to “breathe” thus forcing the zincs to move further apart to allow for water ligation. Next, we produced binding free energy data that correlate well with bioassays suggesting that the tail only crudely orients the lactone head in the active site. The markedly reduced activity of AiiA toward HSL was correlated to the protonated HSL(+) rather than the neutral species as demonstrated by a broad range of  $\Delta G$ s, including an unbound state. This protonated HSL(+) is consistent with the

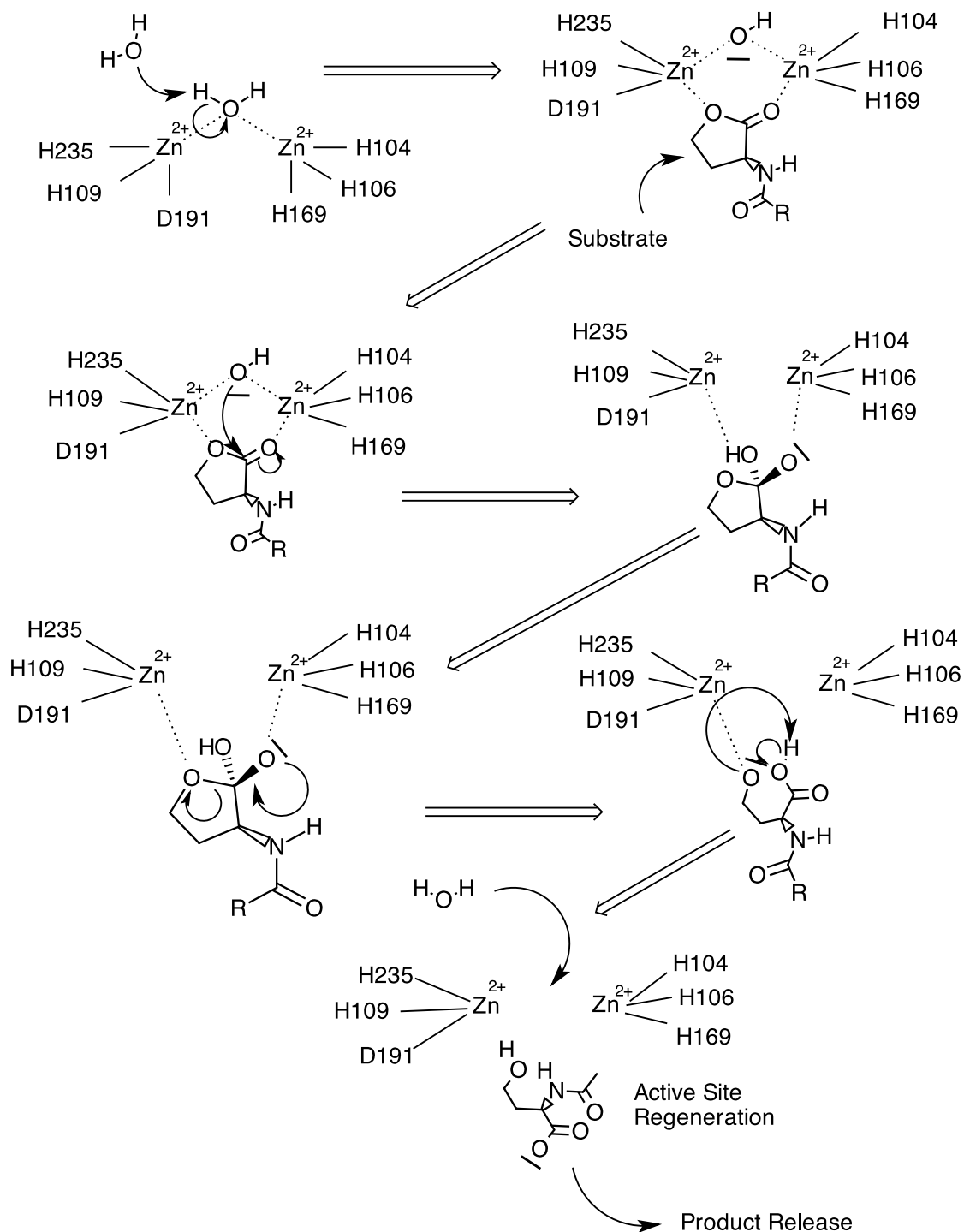
expected protonation state for a pH 7 environment. Thus, the full range of  $\Delta G$ s for lactone ligands is consistent with the bioassay data observed by Kim et al.<sup>10</sup>

Residues F107, D108, and Y194 were investigated for their influence on binding energetics and substrate stability. Also, both D108 and Y194 are referred to as having the potential to serve as proton shuttles for the zinc-bound water and protonation of the ring-opened alkoxide and so were examined for that effect. The binding free energy study on the putative amide hydrogen bonding did not confirm the contribution of F107 to the binding free energy. Instead the reduced ligands lower hydrophilicity actually strengthened the association interaction. The distance analysis of Y194 showed no early potential in forming an oxyanion hole. This outcome may be due to the simulations with reactants that have less developed negative charge on the lactone carbonyl oxygen ( $q = -0.46$ ) than that of the anionic transition state (closer to  $-1.0$ ). That the reaction is not totally quenched by the Y194F mutation suggests that Y194 serves as a stabilizing influence rather than as a true oxyanion hole. This is strengthened by observations of the influence of Y194 on the intermediate state carbonyl oxygen but not on the reactant or product state. That Y194 influences the intermediate state suggests late involvement in the first transition state or conversely, early involvement in the second. D108 is observed to be close enough to the hydroxide ion to perhaps provide some stabilizing influence through hydrogen bonding. Lastly, pKa analysis showed no evidence of either D108 or Y194 having altered pKa's in the range necessary to serve as effective proton shuttles for the zinc-bound water and ring-opened alkoxide. However, this does not preclude water itself from serving as the (de)protonating medium. The pKa of bulk water is 15.7 which is more than basic enough to deprotonate a zinc-bound water (pKa 7.5), whereas the

resulting hydronium ion is definitely acidic enough ( $\text{pK}_a -1.7$ )<sup>95</sup> to protonate adjacent waters. Moreover, studies of similar dimetalloproteins and dimetallic enzyme analogs suggest that the  $\text{pK}_a$  of a water molecule-bridging zinc ions is lowered to no more than about 5.8 - still two orders of magnitude higher than D108's acid dissociation constant.<sup>96,97</sup> Inspection of the binding pocket reveals a structure that is stable and well exposed to its environment thus providing an avenue for proton transport through a water molecule network that exists prior to substrate binding. Upon substrate binding, the zinc-bridging hydroxide anion is effectively protected from re-protonation by the plug-like function of the acyl tails whereupon the ring-opening reaction can proceed. Subsequent protonation of the product's alkoxide group can then be accomplished by the intramolecular proton transfer from the intermediate state's carboxylic acid group where the resulting carboxylate group better accommodates the still-neighboring zinc ion. An alternate mechanism supporting these aforementioned items is rendered in Figure 30.



**Figure 30. Proposed mechanism for lactonase catalyzed AHL ring-opening**



Key points in this alternate mechanism are intramolecular proton transfer in the intermediate state, and water as a deprotonating base to the zinc-bound water.

### 3.4 Future direction

The prime motivator for this work was to see how biofilm formation might be prevented. Specifically, these biofilms are seen as sources of biological contamination in spacecraft environments that can promote disease or even equipment failure. The closed environment is problematic in that the astronauts serve as both sources and potential victims of what they originally introduced into a very constrained space. It is also problematic in that elimination of the biofilms and/or their inoculants becomes much harder to do with traditional biocides as the enclosed environment provides an easy means for biocide buildup that can be harmful to the occupants. Last and equally important is the issue of payload whereby both biocides and or alternate technologies to date require some significant amount of mass that must be considered during mission planning. The ultimate thrust of the research undertaken was to better learn requirements for quorum signaling molecule degradation so as to translate the knowledge toward a more passive means of biofilm prevention. This in turn might help in reducing payload requirements.

At this point in time, our research has revealed several aspects of catalyzed lactone hydrolysis that might be leveraged toward practical design of artificial catalysts displaying either lactonase or peptidase activities. For instance the lactonase's evolved use of metals serves multiple purposes. First, is the ability of the metals to orient the substrates for best pre-attack geometry. This is a classic enzymatic function that serves to lower the activation energy and was shown in our modeling via the juxtaposition of both lactone and hydroxide bridging. Next is the ability of the electrophilic metal to coordinate with water to create a more reactive species. This is critical to the hydrolysis reaction and

shows an evolved ability of the protein to recruit discrete members of its surrounding environment for bond making and breaking interactions with its substrate. Last is the metals' function to both knit the protein together while still allowing for "breathing," where breathing can be thought of as the alternating between pentacoordinated and hexacoordinated states. The alternation between coordination states allows for free energy driving of water functionalization, substrate docking, and substrate release.

Another aspect seen by our research is the lactonase-substrate complex's ability to exclude water. This water exclusion prevents bound substrate and critical residue displacement from their respective metal ions thus promoting geometric stability of both near-attack and transition states. Conversely, this water exclusion is followed by the ability of the complex to admit water to the site after the reaction is complete to aid in both displacing the ligand and regenerating the water substrate for the next reaction cycle. As in substrate orientation, the creation and control of the active site's hydrophobic pocket is a common theme in enzyme mechanism. However, the interesting observation (and challenge to the design problem) is how the enzyme evolved (in the case of AHLs) to use the substrate's own alkyl tail to serve as a water barring plug. This is significant as the protein need not develop more complex domains or features to facilitate the creation of a water-free environment. It essentially remains a simple funnel to a binding and active site.

In attempting to leverage our observations of the lactonase-AHL system one can immediately see how various chelant moieties might be utilized to capture the benefits of metal ions. For future work, studies of which chelant-metal pairings in which environment might be both electronically and geometrically closest to those same

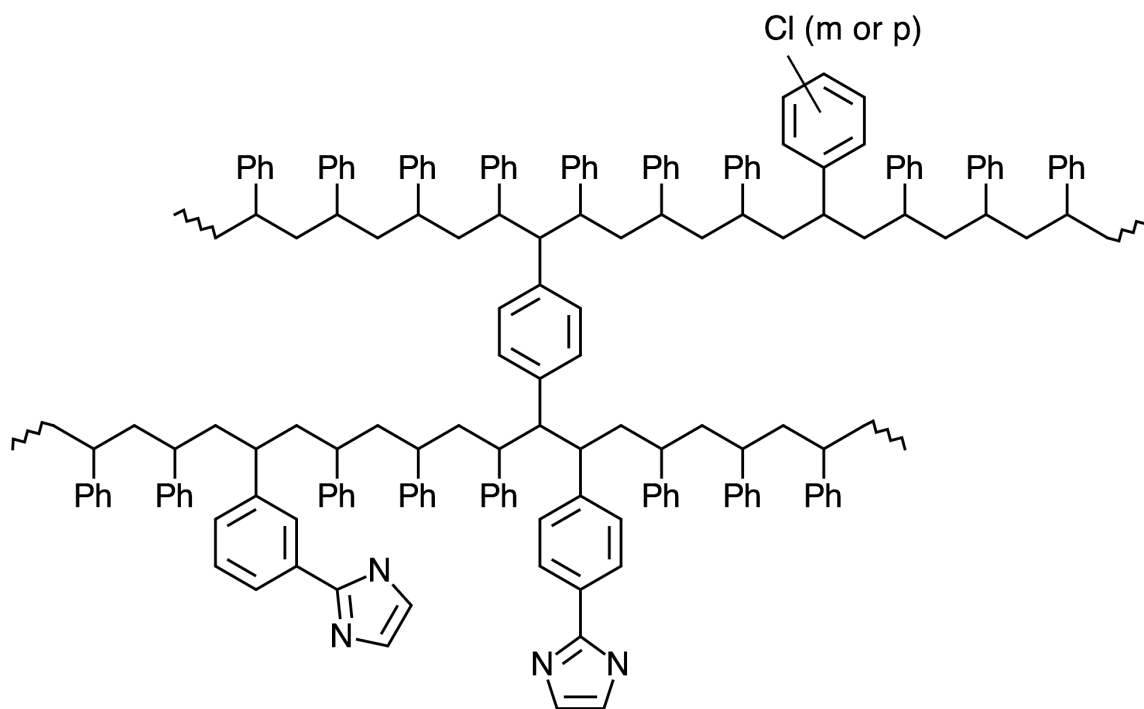
measurements we drew from our observations. Indeed these chelant-metal pairs need not exist as pairs themselves (e.g., dimetal protease mimic) but might exist singly as in carbonic anhydrases that would in turn reduce the number of degrees of freedom needed for design consideration. Of course, metal-chelant pairs need not be the only chemistries to investigate, amino acid residues or other functional groups could also be investigated for hydrolase/peptidase activity.

Of course, functional groups that might promote a reaction are only part of what would need to be considered. Of equal importance is how to bring the groups into the correct orientation relative to the substrate(s). For example, much work has been done on the functionalization of cyclodextrins whereby the cyclodextrin is adorned with “pendant” moieties that serve as reactive species to drive a desired reaction. These molecules have been synthesized and do work, albeit on less challenging ester substrates. Improvement of artificial enzymes for more challenging molecules containing amide bonds might require more degrees of freedom than a cyclodextrin can provide. As a scaffold, polypeptides might be considered, however for the desired function of creating passive bioactive surfaces one would have to be wary of bacterial or even human derived proteases that could quickly break them down. Toward that end peptidomimetic scaffolds might be tried or, perhaps use of standard polymers might be better suited for functionalization provided local environmental effects (hydrophobicity and pH) are taken into account.

In conjunction with functional group and scaffolding research and perhaps most important is the creation of the hydrophobic environment necessary to promote good charge separation in the substrate during bond breaking and making. In the case of AHL

substrates, not only does the hydrophobicity help in charge separation, but it can also be leveraged to aid in product clearance. That the product has two more exposed hydroxyl groups to the environment relative to the more closed nature of the lactone oxygens implies that the product is more hydrophilic than the reactant. This property difference could be even more accentuated if peptidase activities were introduced as a means to cleave the fatty alkyl tail. Toward that end, the use of a phase transfer catalyst might be best employed directly or in combination with the scaffolding structure and functionalization. These strategies are not new as catalytic polymers have already been already been synthesized as shown in Figure 31. These polymers might be better selected to enhance specific substrate partitioning.

**Figure 31. Artificial enzyme surface for AHL ring-opening or peptide hydrolysis**



This figure demonstrates how virtually any polymeric scaffold might be functionalized with the right moieties to provide for catalysis, however, achieving the correct geometries is what will ultimately provide catalytic turnover.

From an experimental standpoint, an interesting approach would be to mix both synthetic chemistry methodology (including biochemical assay) with quantitative structure activity relationship (QSAR) analyses to tune in to the most efficacious construct. For instance, it might be found that catalytic efficiency may be mapped to copolymer families versus stoichiometric ratios of various functional groups. Numerous mappings can be envisioned and may be limited only by resources available to create different construct combinations. Also, as already mentioned, further research into amide hydrolysis with respect to AHLs is recommended in order to bring dual catalytic function to the potential constructs.

## **CHAPTER 4: CONFORMATIONAL ANALYSIS OF PTEN:3L4F-F1 COMPLEX**

### **4.1 Introduction**

As mentioned before, drug dependence, or addiction ultimately described as the maladaptation of a neurochemical reward pathway that is normally used to reinforce behaviors that promote an organism's survival and/or propagation (e.g., high calorie food seeking and sexual drive). These pathways are mediated via dopaminergic VTA-NAc neuronal networks. The molecular basis for this interaction has been shown to rely on a 5HT-2cR:PTEN interaction via the 3L4F-F1 fragment. The nature of the interaction is unknown. In this work, the aim is to generate viable 3L4F-F1 conformations upon which exhaustive coarse-grained docking may be performed to find potential binding sites.

### **4.2 Computational methods**

#### **4.2.1 Molecular dynamics simulations – TIGER2**

Replica Exchange Molecular Dynamics (REMD)<sup>67</sup> simulations were performed on both the 3L4F and 3L4F-F1 peptides using the Temperature Intervals with Global Energy Reassignment (TIGER2)<sup>69</sup> strategy to sample the secondary structural content of the peptides. Molecular dynamics simulations for all simulation runs were carried out using the parallel molecular dynamics software NAMD 2.6 (Not Another Molecular Dynamics Program)<sup>98</sup> to compute dynamics trajectories utilizing the CHARMM 27 (Chemistry at Harvard Macromolecular Mechanics 27)<sup>66</sup> empirical force field. The 3L4F and 3L4F-F1 peptides were first generated in the extended conformation followed by solvation with TIP3<sup>91</sup> explicit water molecules after which counterions were inserted to



neutralize the system. Molecular dynamics trajectories were run for 50 ns each at 300K, 378K, 476K, and 600K with an exchange frequency of 3 ps.

#### **4.2.2 Clustering by dPCA**

The resultant 300K molecular dynamics trajectory was examined for secondary structure using dihedral principal component analysis (dPCA; implemented in Carma<sup>99</sup>) as the 300K conditions are the most physiologically relevant. In addition, the VMD plugin Timeline was used to graphically map characteristic peptide backbone  $\varphi$  and  $\phi$  angles to juxtapose against dPCA analytical results. The dPCA was conducted in two steps. First, the initial trajectory was subject to analysis for which the first two eigenvalues were identified. The contiguous regions of the trajectory corresponding to the first two eigenvalues were then extracted from the original trajectory after which it was subject to another round of dPCA analysis. Again, in this case the averaged structures for the first two eigenvalues were then extracted.

#### **4.2.3 Coarse, blind docking with GRAMM**

After using REMD to generate physically realizable conformations of the peptide, the candidate structures were then subject to coarse-grained cluster docking against the crystal structure of PTEN [Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank; [www.rcsb.org](http://www.rcsb.org); ID #1D5R] using GRAMM (Global Range Molecular Matching)<sup>75</sup> to predict potential peptide-protein interactions. All GRAMM docking runs were executed with the following parameterization:

Matching mode (generic/helix) ..... mmode= generic

Grid step ..... eta= 1.7

Repulsion (attraction is always -1) ..... ro= 5.0

Attraction double range (fraction of single range) ..... fr= 0.0

Potential range type (atom\_radius, grid\_step) ..... crang= atom\_radius

Projection (blackwhite, gray) ..... ccti= blackwhite

Representation (all, hydrophobic) ..... crep= hydrophobic

Number of matches to output ..... maxm= 1000

Angle for rotations, deg (10,12,15,18,20,30, 0-no rot.) ai = 10

In this parameterization set, matching mode (mmode) “generic” serves as a switch to enable exhaustive six-dimensional (translation and rotation) geometric searching. Along these lines, grid step (eta=1.7) and angle for rotations (ai=10 degrees) were set to their minima to allow for the highest degree of six-dimensional resolution. Repulsion, attraction double range, potential range type, projection, and representation parameters were set for high resolution hydrophobic docking using GRAMM.

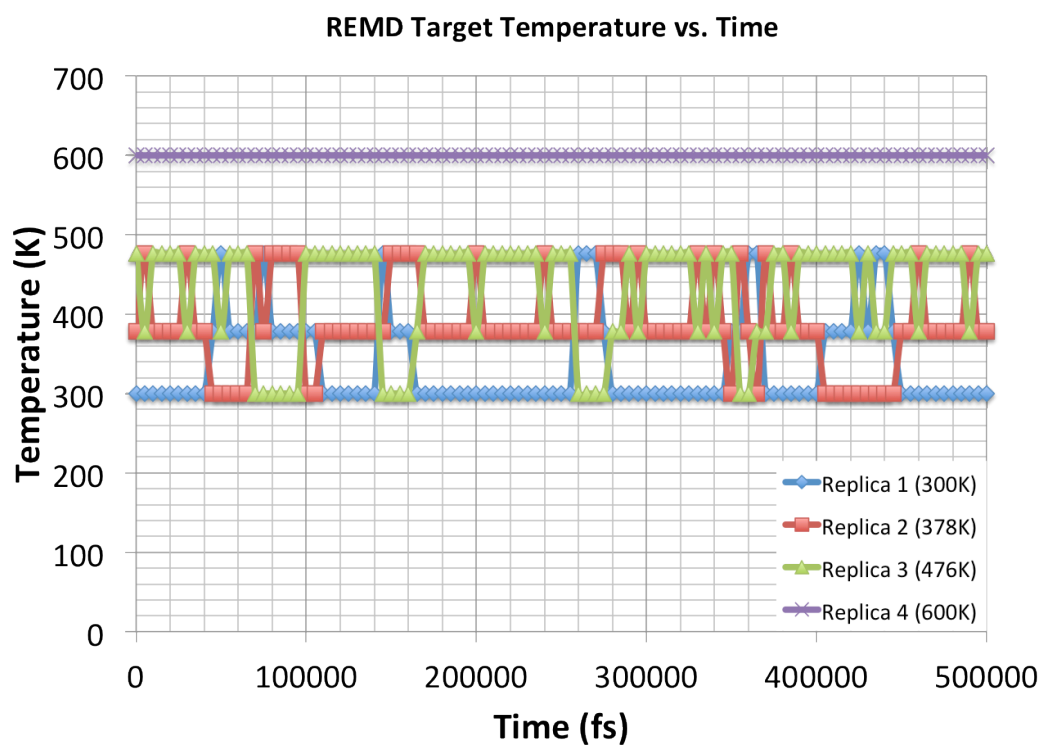
## **4.3 Results and discussion**

### **4.3.1 Molecular dynamics simulations**

In approaching this blind docking problem, the initial approach was to explore the preferred conformations the 3L4F-F1 peptide might assume using REMD. To better utilize computing resources the TIGER2 algorithm was used so as to minimize the number of replicas needed while also increasing the probability of exchange by allowing for global temperature reassignment of velocities (instead of exchange between only temperature adjacent replicas). A visual representation of the exchange data is shown in Figure 32 where only the first few picoseconds are shown so as to provide enough resolution to see the replica set point temperature oscillations. As can be seen, the lower temperature replicas at 300, 378, and 476K show favorable exchange whereas the 600K

replica remains isolated. Further quantification of exchange shown in Table 15 (as measured by the amount of time a replica spends at a given set point temperature) also shows the three lower temperature replicas undergoing significant exchange while the 600K replica remains isolated throughout the entire 50 ns trajectory. The data also show potential modal behavior whereby the 300K and 378K replicas tend to spend more time at 300K while the 476K replica tends to spend more time at 476K. This, coupled with the failure of the 600K replica to undergo exchange presents a very rough mapping of energy levels for 3L4F-F1. With regards to the accessibility of the 600K energy level, it may be a steep mesa in the landscape, or it may be that more intermediate replicas might be used to achieve exchange. Regardless, this is seen as an avenue for further exploration and speaks more to methodology than to the more immediate need to generate representative conformations at physiological conditions.

**Figure 32. Replica exchange vs. time**



---

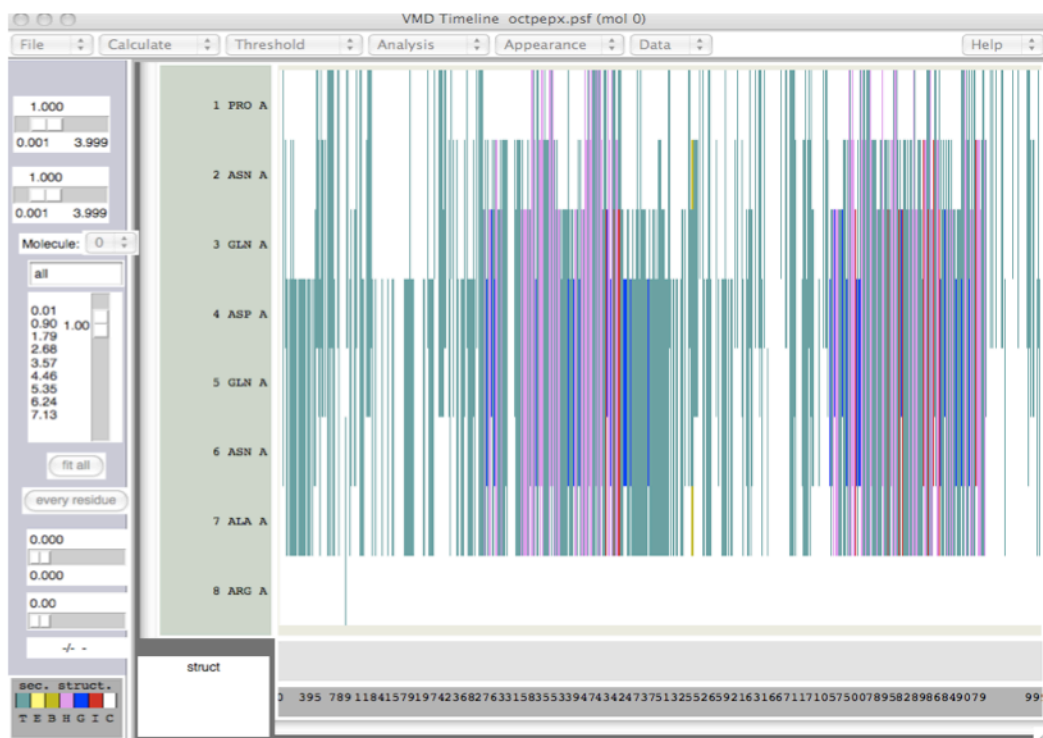
This is an example of trajectory replica exchange during a REMD simulation. The exchange is between velocities for replicas that meet energy exchange criteria.

**Table 145. Replica time at temperature**

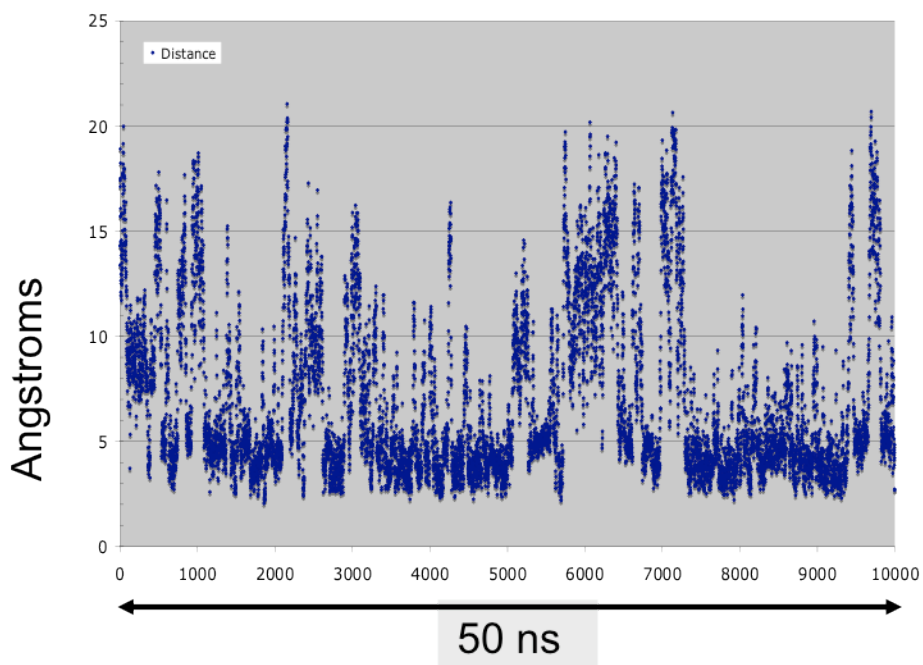
	300K	378K	476K	600K
Replica 1 (300K)	41%	34%	25%	0%
Replica 2 (378K)	48%	37%	15%	0%
Replica 3 (476K)	12%	29%	60%	0%
Replica 4 (600K)	0%	0%	0%	100%

Conformational space amongst the replicas was initially explored through the use of VMD Timeline and secondary structure tools to see how conformational space may have changed along a full trajectory. In these displays, it is important to note that all trajectories include conformations seeded with other replicas' velocities as designated by the REMD algorithm. In addition, salt bridge distance between ASP4 and ARG8 was tracked across all trajectories to see if there might be any correlation to secondary structure. The 600K trajectory appeared to have the most diffuse signal with regards to salt bridge distance and secondary structure (Figure 33). It was also remarkable that  $\alpha$ -helical structure was interspersed along the trajectory that seems appropriate behavior for conformations at that temperature (rapid conformation inter-conversion). Stability of structures at this temperature precluded their inclusion into any further docking study. Further molecular dynamics work would be needed to ensure replica exchange with the 600K replica to consider these structures (to make sure these conformations are realizable). The 476K and 378K replicas showed more stable behavior than the 600K replica with only turn-like structure being displayed (Figures 34 and 35). Also, the salt bridge distance roughly correlated to turn structures shown contiguously between residues four through seven as expected. The 300K trajectory revealed a significant region of stable  $\alpha$ -helical structure with corresponding correlation to salt bridge distance. Along with this observation, a region of turn- $\beta$ -turn structure was seen in the later part of the trajectory (Figure 36). As a means to extract meaningful prioritization of the structures observed, dPCA was conducted on the 300K trajectory to find candidate conformations for docking.

**Figure 33. 600K Replica secondary structure data**

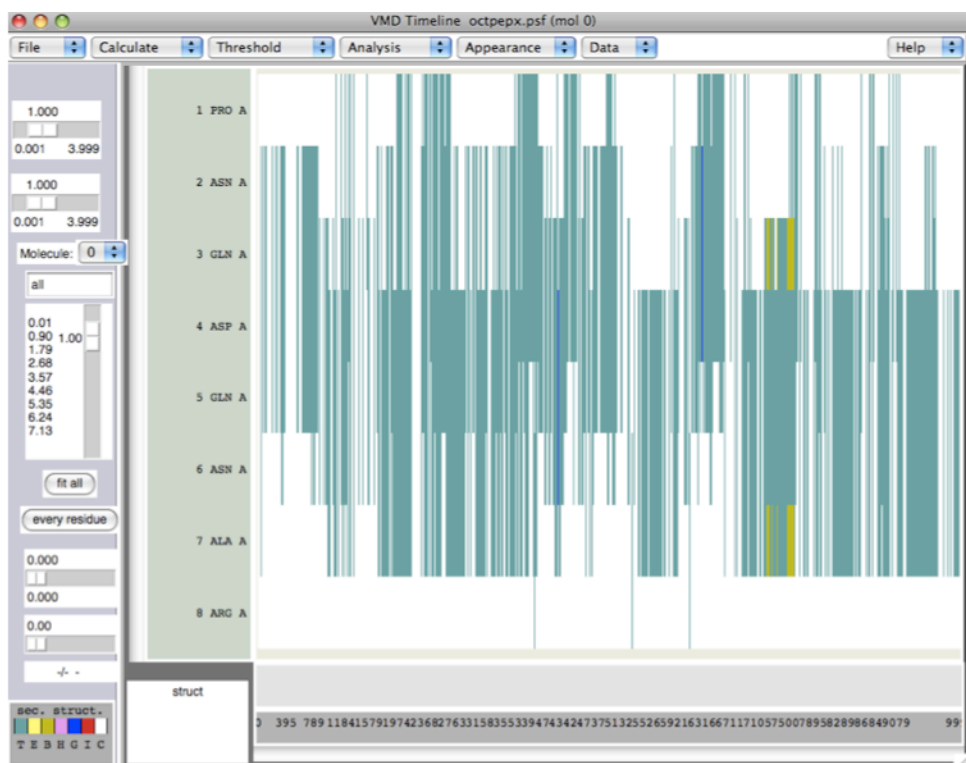


600K Replica Asp4 to Arg8 Salt Bridge Distance

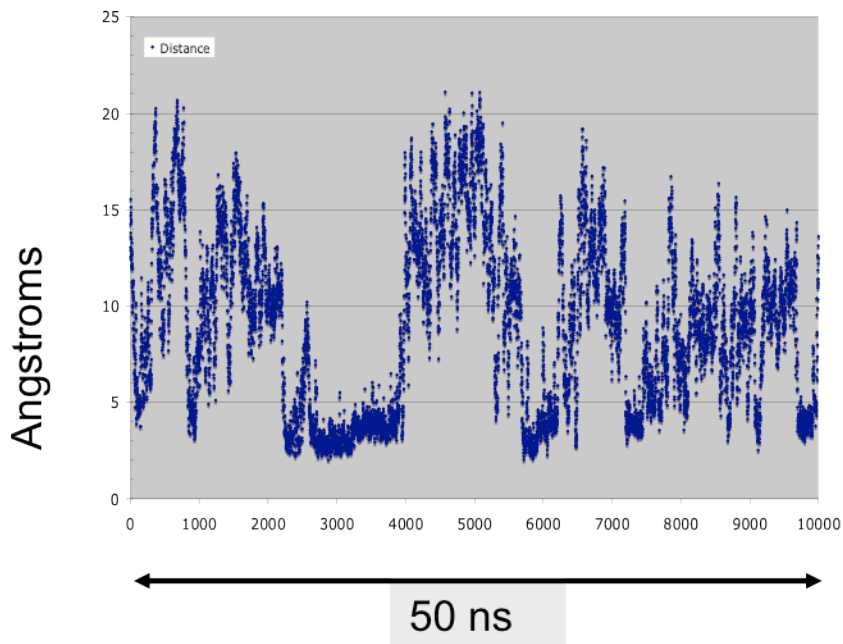


Of interest is the diffuse  $\alpha$ -helical content at the highest temperature replica. These conformations were inaccessible to lower temperature replicas during the REMD run.

**Figure 34. 476K Replica secondary structure data**



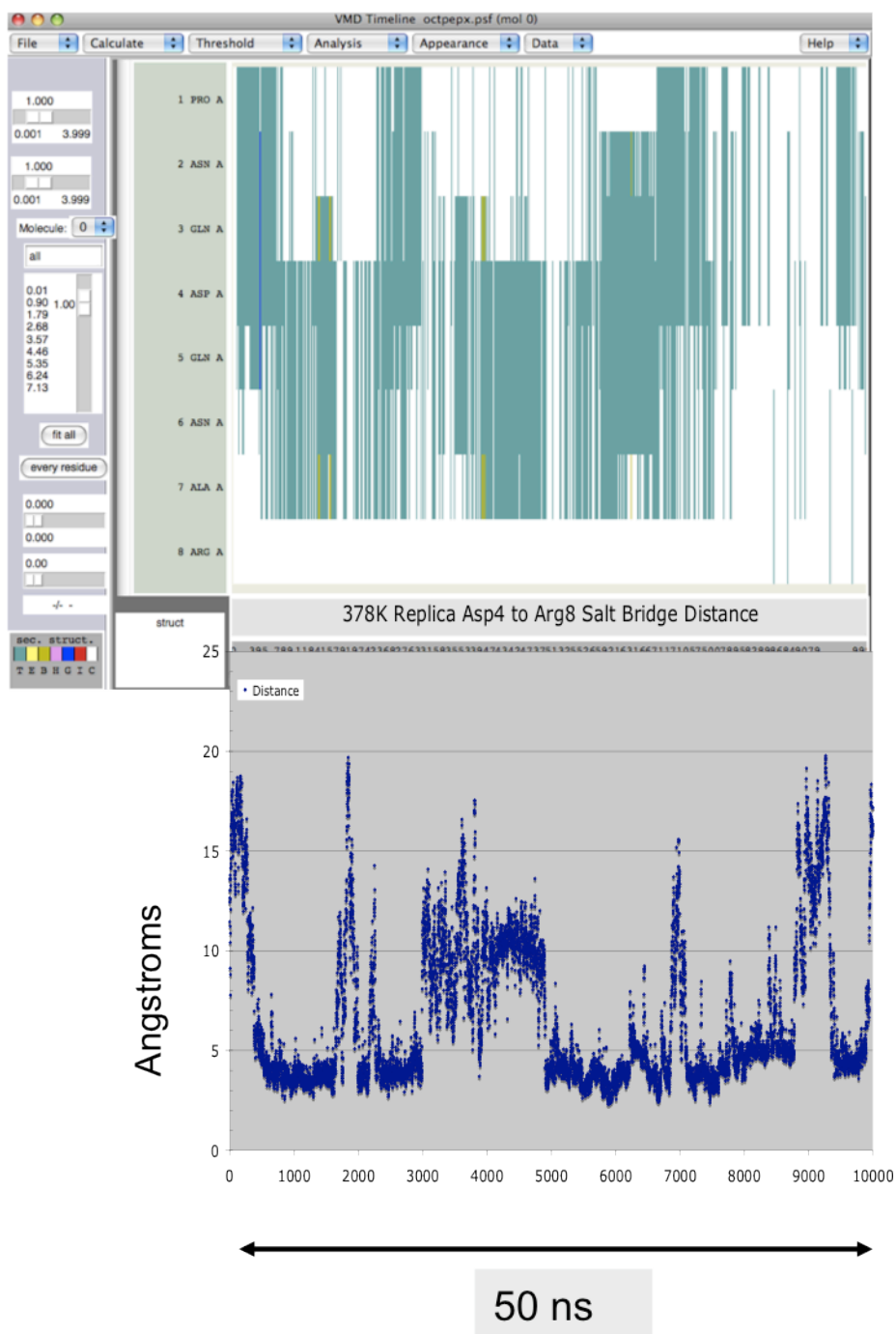
476K Replica Asp4 to Arg8 Salt Bridge Distance



Salt bridge distance correlates to turn structure seen between ASP4 and ARG8.

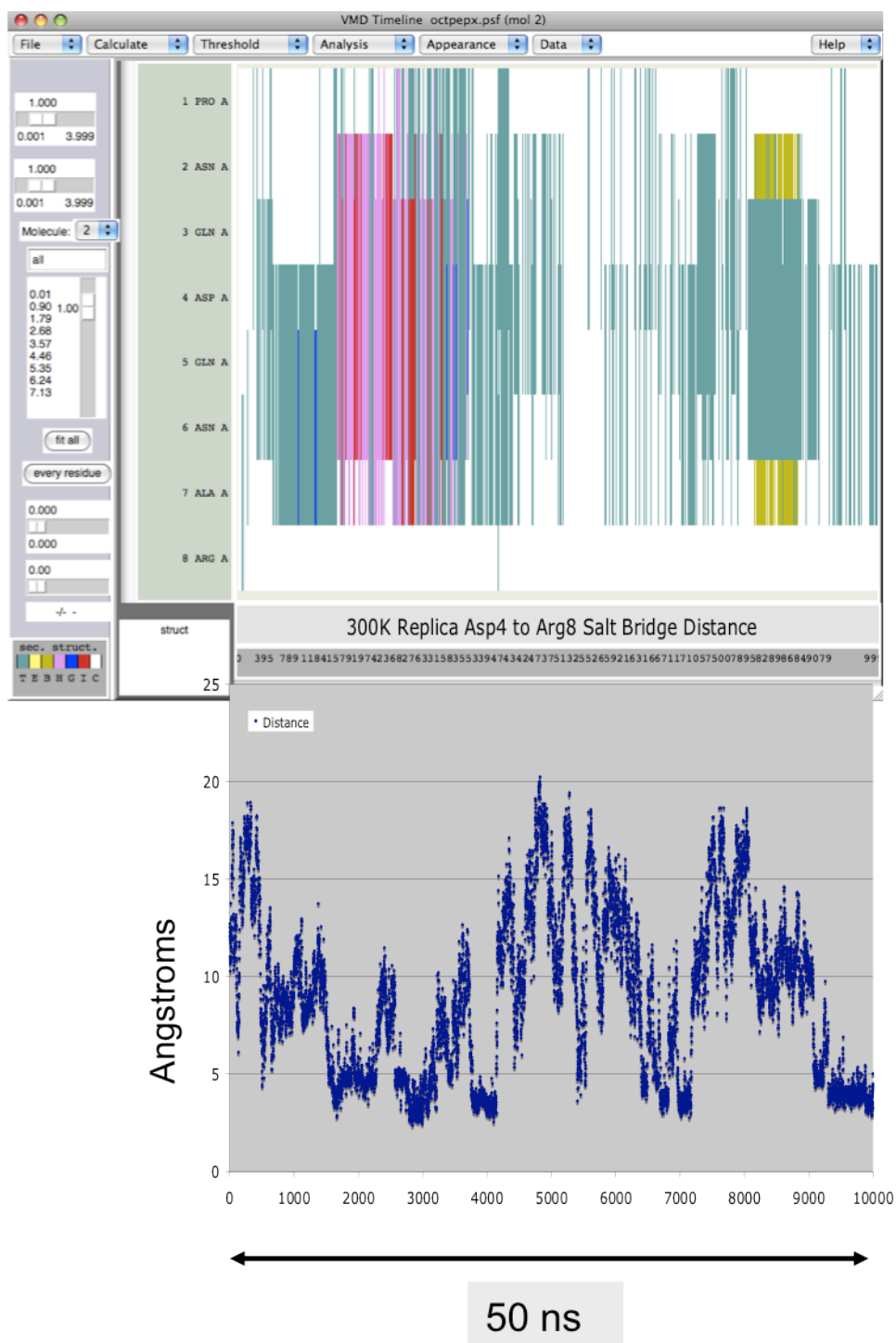


**Figure 35. 378K Replica secondary structure data**



Again, as in the 476K replica, salt bridge distance is seen to correlate to turn conformations seen between ASP4 and ARG8.

**Figure 36. 300K Replica secondary structure data**

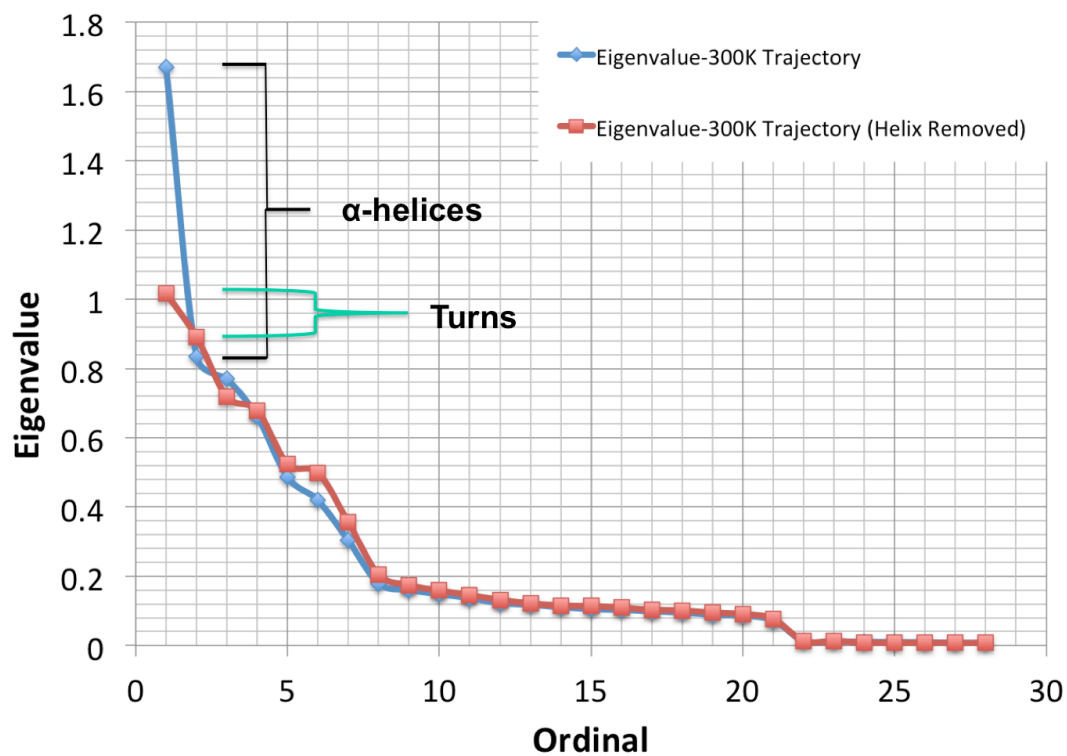


Relative to the 600K trajectory, the  $\alpha$ -helical content seen in this trajectory is less diffuse and also shows salt bridge distances between ASP4 and ARG8.

### 4.3.2 dPCA analysis

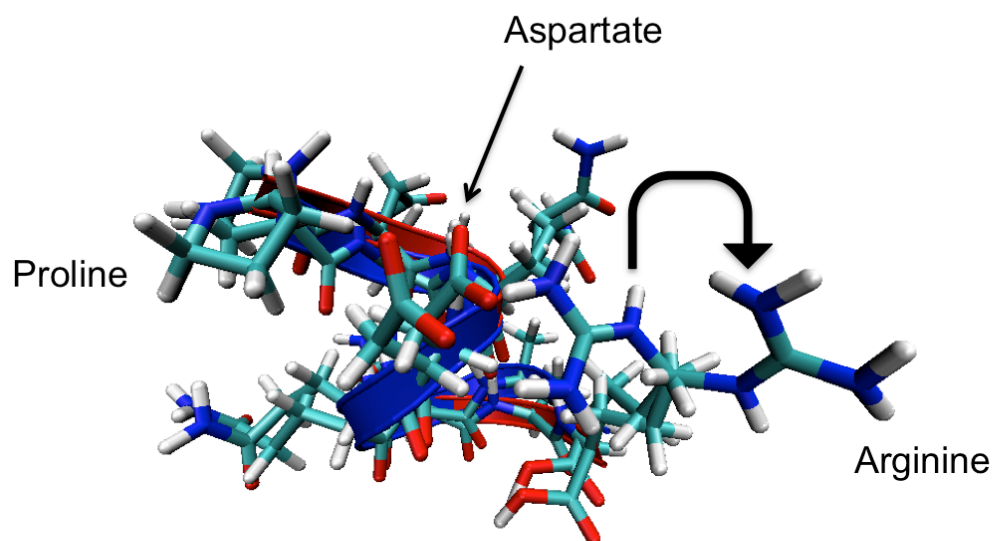
dPCA analysis was first performed on the whole 300K trajectory to quantify and rank secondary structure contributions of each residue. From this initial analysis, the averaged structures of the first two highest eigenvalues were extracted. Next, the regions in the molecular dynamics trajectory corresponding to the first two eigenvalues were removed. The resulting trajectory was then again subject to dPCA analysis from which the next two averaged structures from the highest eigenvalues were extracted. The eigenvalue analyses output is shown in Figure 37. The end result was a total of four candidate conformations from the most physiologically relevant trajectory to be subject to coarse-grained docking. The first two structures consist of  $\alpha$ -helices differentiated primarily by the position of the ARG8 residue whose side-chain angle differs by  $180^\circ$  (Figure 38). Close examination of the two conformations suggests that the ARG8 side-chain angle either promotes or disrupts a potential salt bridge interaction with upstream residue ASP4. The salt bridge interaction is nominally stronger than hydrogen bond interaction energy and may point to a stabilizing influence on the candidate conformation while also suggesting a location to covalently link a peptidomimetic compound. The other two structures (Figure 39) show two less similar, turn-like conformations with no apparent salt bridge interactions. As these structures were derived from the trajectory that is helix-free, a full set of prioritized conformations for the entire 300K trajectory were generated for later coarse-grained docking. Intervening water networks and their contributions were not considered in this analysis although they may carry some amount of importance in later studies. These are pilot conformations only from which later pharmacophore models might be derived.

**Figure 37. 300K replica dPCA analysis**



Top four 3L4F-F1 conformations shown as a function of relative eigenvalue score derived from dPCA.

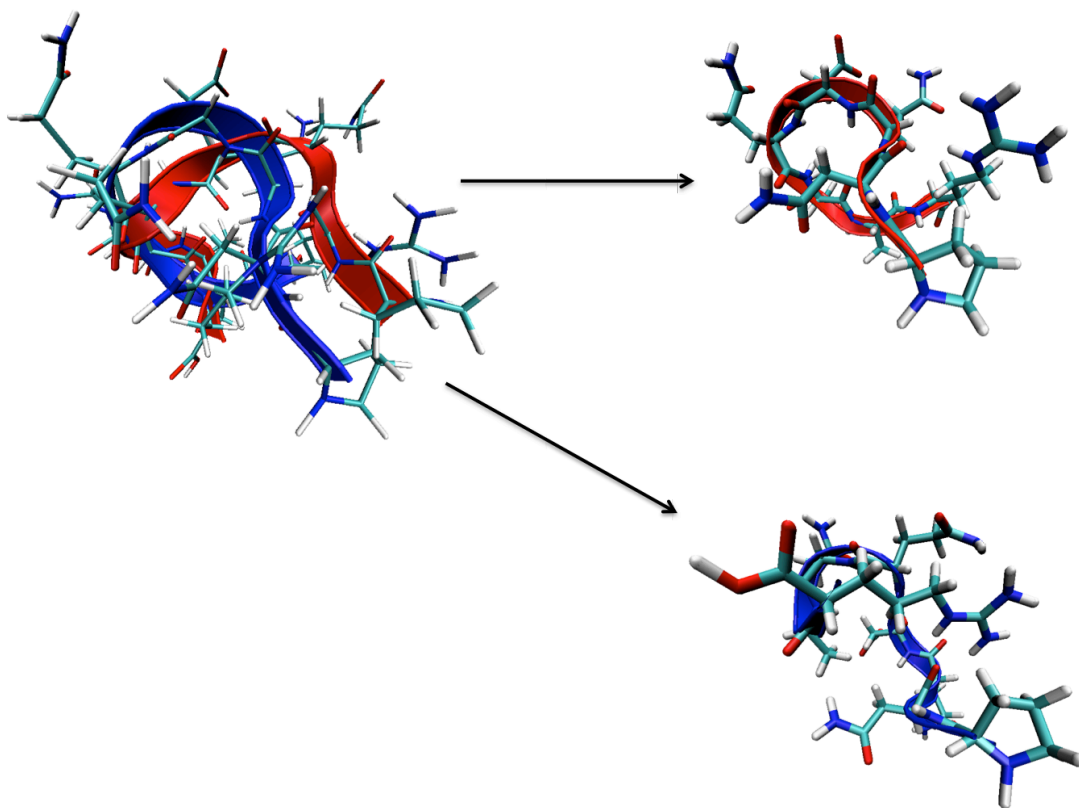
**Figure 38. Helical conformations per dPCA analysis**



---

Arrow is used to show swiveling of the ARG8 residue. Blue is used to designate helical conformation number 1. Red is used to designate helical conformation number 2.

**Figure 39. Turn conformations per dPCA analysis**



---

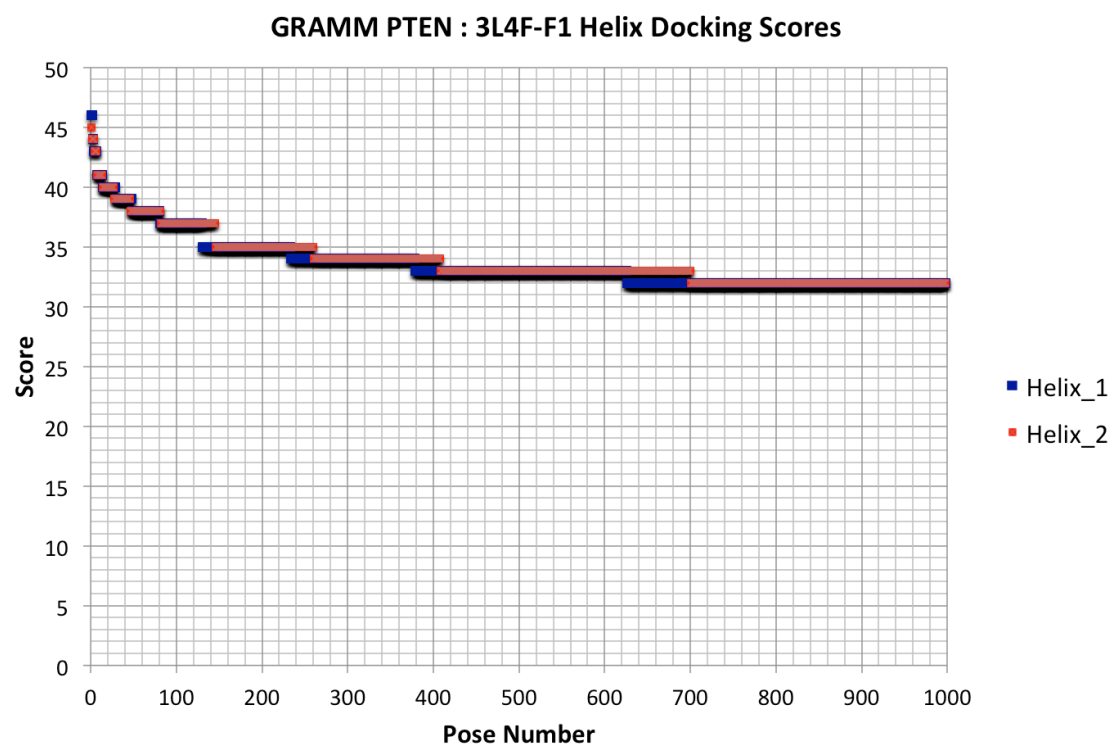
Blue is used to designate turn conformation number 1. Red is used to designate turn conformation number 2.

### 4.3.3 PTEN docking hot spots and CD validation

Prior to docking validation, investigation of both early 3L4F and 3L4F-F1 300K trajectories showed that the  $\alpha$ -helical conformation seen in 3L4F-F1 is preserved at least partially in the 3L4F trajectory (Figures 42A, 42B, and 42C). Coarse-grained docking using GRAMM was performed with each of the four candidate structures against the PTEN (1D5R PDB) protein. Amongst the top 100 poses, all candidates were observed to dock primarily within the cleft between the phosphatase/lipase catalyzing and membrane associating regions of the protein (Figures 42D and 42E). These 100 poses were selected for observation as they were the poses shown to undergo the steepest score (energy) decrement over the 1000 peptide-protein complexes considered (Figures 40 and 41). The putative interaction site location is based primarily on geometric considerations only and awaits further experimental validation; however, opportunities remain to perform such work readily.

A validation of the 3L4F-F1 and 3L4F peptide structures was found in the circular dichroism analyses performed by Anastasio, et al. (Figures 42F and 42G).<sup>100</sup> While not definitive, these analyses do show some amount of  $\alpha$ -helical character that is seen not only in the 3L4F-F1 300K trajectory, but also in pilot REMD runs undertaken on 3L4F. Not only is this structural characteristic observed experimentally, but  $\alpha$ -helical content is also seen in the same region of the 3L4F peptide as seen in 3L4F-F1. This strongly suggests that  $\alpha$ -helical structure should be considered when moving forward into peptidomimetic design and synthesis.

**Figure 40. 3L4F-F1 Helix GRAMM docking scores**

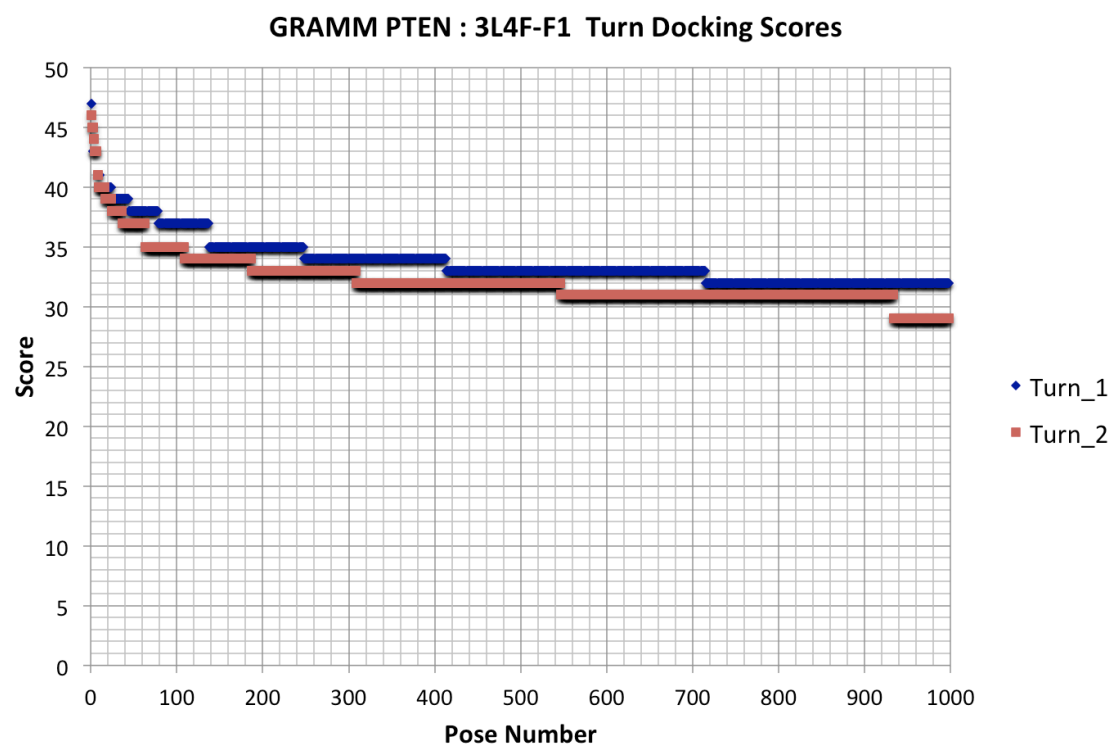


---

The score in this analysis is dimensionless and merely refers to the number of hydrophobic and non-hydrophobic contacts.



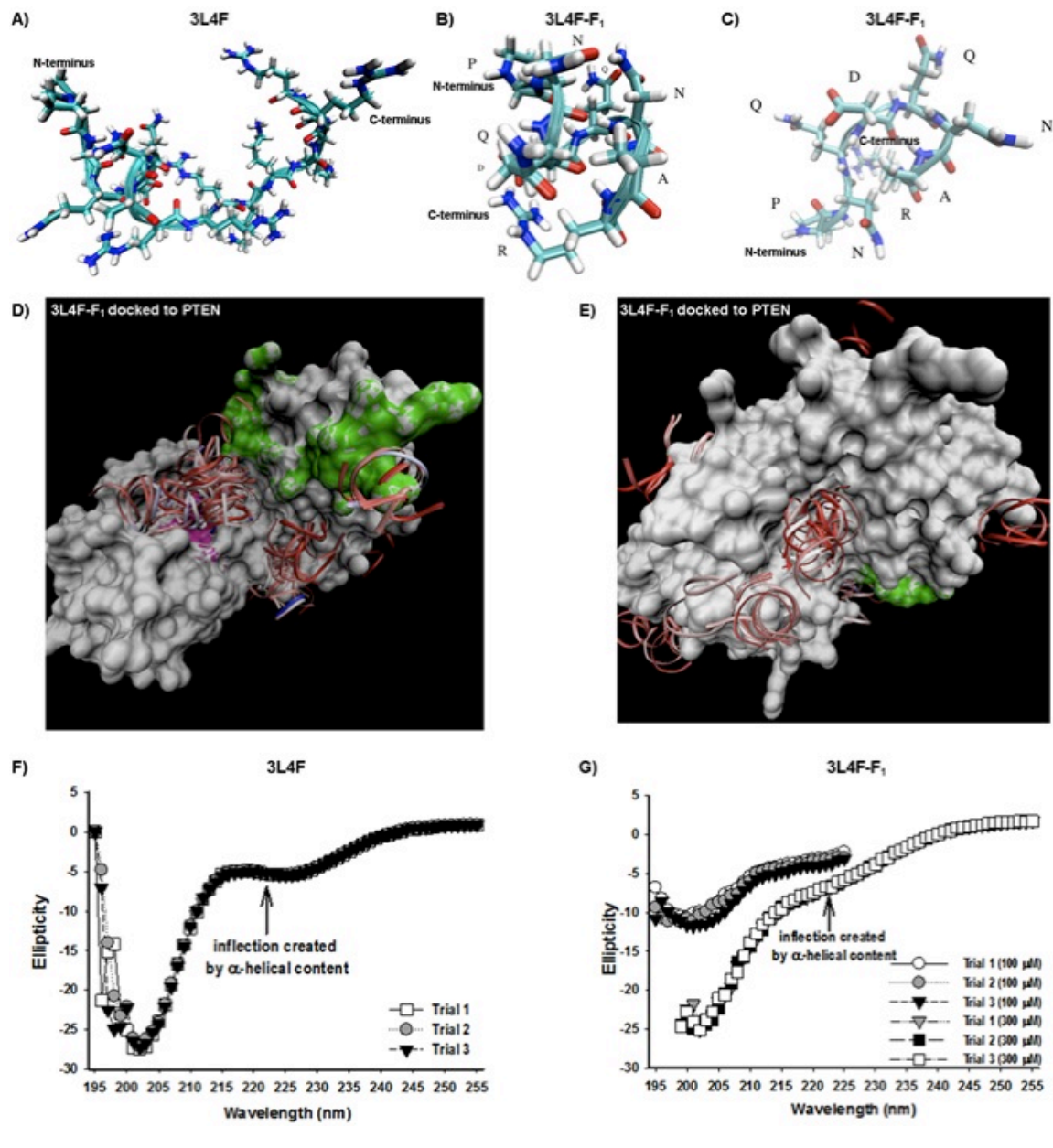
**Figure 41. 3L4F-F1 Turn GRAMM docking score**



---

3L4F-F1 turn conformations shown as per Figure 40.

**Figure 42. GRAMM and CD data**



A, B, and C, show the original 3L4F, 3L4F-F1 helix, and 3L4F-F1 turn conformations respectively. D and E show the putative binding sites as generated by GRAMM. F and G present CD validation of conformations A and B.

#### 4.4 Conclusion

Biologically active peptides are important leads for pharmaceuticals and bioprobe development. The process of first starting with a peptide, followed by pseudo-peptide (peptides with modified peptide bonds) synthesis succeeded by peptidomimetics and finally to analogs (small molecules that differ significantly from the initial peptide) has been used to develop important therapeutic medications such as angiotensin converting enzyme inhibitors,<sup>101,102</sup> Arg-Gly-Asp mimics,<sup>103,104</sup> and HIV protease inhibitors.<sup>105,106</sup>

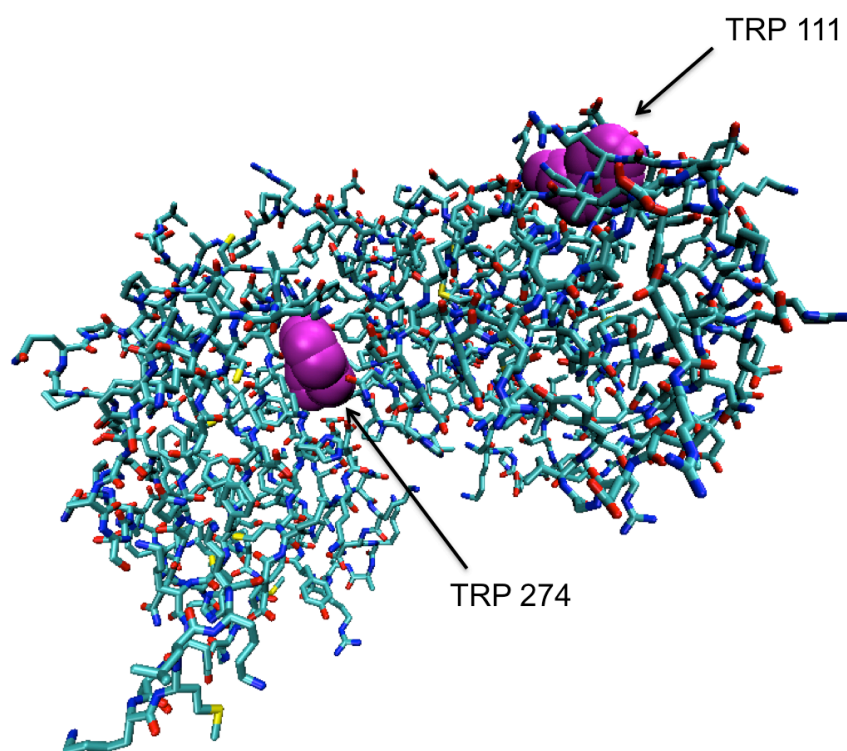
A process of synthesizing the lead peptide 3L4F and peptide fragments of 3L4F to identify the active portion of the 3L4F peptide was conducted. Following the discovery of the minimum active 8mer peptide 3L4F-F1, molecular modeling was employed to direct the future development of peptidomimetics that have greater stability and better activity than the small peptides. Toward that goal, 3L4F-F1 peptide conformations were computationally predicted to produce both  $\alpha$ -helix and  $\beta$ -turn secondary structures over the time course of the simulations. Moreover, coarse-grained docking of these peptide conformations against PTEN provide preliminary insight into potential interaction sites which cluster between the phosphatase active site and C2 membrane-associating domain of PTEN. Validation and refinement of 3L4F-F1 within the PTEN interaction site will be key in providing additional atomistic detail for pharmacophore model definition.

#### 4.5 Future direction

Close inspection of the PTEN sequence and structure revealed that the protein possesses only two tryptophan residues TRP111 and TRP274. While TRP111 is located near the surface of the protein, TRP274 is located well within the crevice region where the putative interaction site is as well (Figure 43). To validate the predicted 3L4F-F1

binding site, it is recommended that first, TRP111 be mutated to PHE. Following that, it is suggested that fluorescence spectroscopy be conducted to assess TRP signal quenching upon titration of a sample of the PTEN mutant with 3L4F-F1. Based on our predictions, it is hypothesized that upon titration, a decreased fluorescence should be observed. This decrease in signal is expected due to occlusion of TRP274 by 3L4F-F1 thus preventing both fluorophore excitation and emission. This approach should not prevent attempts to gain a crystal structure of the complex to allow for a better atomistic description of the PTEN:5HT-2cR interaction. This description could then be leveraged to provide for more detailed study toward a pharmacophore model or for a more nuanced reduction of the peptide to a peptidomimetic. Alternatively, further reduction of 3L4F-F1 to less than eight amino acids should also be attempted as computational docking becomes more attractive at five amino acids or less.

**Figure 43. PTEN tryptophan residues**



---

Tryptophan residues are shown for clarity. Specifically, W111F mutation might be leveraged for binding site validation via fluorescence spectroscopy.

## REFERENCES

- (1) Mok, K. C.; Wingreen, N. S.; Bassler, B. L. *EMBO J.* **2003**, *22*, 870.
- (2) Teplitski, M.; Mathesius, U.; Rumbaugh, K. *Chem. Rev.* **2011**, *111*, 100.
- (3) Wang, L. H.; Weng, L. X.; Dong, Y. H.; Zhang, L. H. *J. Biol. Chem.* **2004**, *279*, 13645.
- (4) Miller, M. B.; Bassler, B. L. *Annu. Rev. Microbiol.* **2001**, *55*, 165.
- (5) Pesci, E. C.; Pearson, J. P.; Seed, P. C.; Iglewski, B. H. *J. Bacteriol.* **1997**, *179*, 3127.
- (6) Tateda, K.; Ishii, Y.; Horikawa, M.; Matsumoto, T.; Miyairi, S.; Pechere, J. C.; Standiford, T. J.; Ishiguro, M.; Yamaguchi, K. *Infect. Immun.* **2003**, *71*, 5785.
- (7) Shiner, E. K.; Terentyev, D.; Bryan, A.; Sennoune, S.; Martinez-Zaguilan, R.; Li, G.; Gyorke, S.; Williams, S. C.; Rumbaugh, K. P. *Cell. Microbiol.* **2006**, *8*, 1601.
- (8) Jacobi, C. A.; Schiffner, F.; Henkel, M.; Waibel, M.; Stork, B.; Daubrawa, M.; Eberl, L.; Gregor, M.; Wesselborg, S. *Int. J. Med. Microbiol.* **2009**, *299*, 509.
- (9) Li, L.; Hooi, D.; Chhabra, S. R.; Pritchard, D.; Shaw, P. E. *Oncogene* **2004**, *23*, 4894.
- (10) Li, H. W., L.; Ye, L.; Mao, Y.; Xie, X.; Xia, C.; Chen, J.; Lu, Z.; Song, J. *Med. Microbiol. Immunol.* **2009**, *198-204*.
- (11) Kravchenko, V. V.; Kaufmann, G. F.; Mathison, J. C.; Scott, D. A.; Katz, A. Z.; Wood, M. R.; Brogan, A. P.; Lehmann, M.; Mee, J. M.; Iwata, K.; Pan, Q.; Fearn, C.; Knaus, U. G.; Meijler, M. M.; Janda, K. D.; Ulevitch, R. J. *J. Biol. Chem.* **2006**, *281*, 28822.
- (12) Vikstrom, E. T., F.; Magnusson, K.E. *FEBS Lett.* **2006**, *580*.
- (13) Sauer, K.; Camper, A. K.; Ehrlich, G. D.; Costerton, J. W.; Davies, D. G. *J. Bacteriol.* **2002**, *184*, 1140.
- (14) Pearson, J. P.; Van Delden, C.; Iglewski, B. H. *J. Bacteriol.* **1999**, *181*, 1203.
- (15) Teiber, J. F.; Horke, S.; Haines, D. C.; Chowdhary, P. K.; Xiao, J.; Kramer, G. L.; Haley, R. W.; Draganov, D. I. *Infect. Immun.* **2008**, *76*, 2512.
- (16) Chun, C. K.; Ozer, E. A.; Welsh, M. J.; Zabner, J.; Greenberg, E. P. *Proc. Natl. Acad. Sci. USA.* **2004**, *101*, 3587.

- (17) Draganov, D. I.; Teiber, J. F.; Speelman, A.; Osawa, Y.; Sunahara, R.; La Du, B. N. *J. Lipid. Res.* **2005**, *46*, 1239.
- (18) Xu, F.; Byun, T.; Deussen, H. J.; Duke, K. R. *J. Biotechnol.* **2003**, *101*, 89.
- (19) Dong, Y. H.; Zhang, L. H. *J. Microbiol.* **2005**, No. 43, 101.
- (20) Liu, D.; Lepore, B. W.; Petsko, G. A.; Thomas, P. W.; Stone, E. M.; Fast, W.; Ringe, D. *Proc. Natl. Acad. Sci. USA.* **2005**, *102*, 11882.
- (21) Liu, D.; Momb, J.; Thomas, P. W.; Moulin, A.; Petsko, G. A.; Fast, W.; Ringe, D. *Biochemistry* **2005**, *47*, 7706.
- (22) Momb, J.; Wang, C.; Liu, D.; Thomas, P. W.; Petsko, G. A.; Guo, H.; Ringe, D.; Fast, W. *Biochemistry* **2008**, *47*, 7715.
- (23) Breslow, R. *Artificial Enzymes*; Wiley: Weinheim, 2005; pp 145–158.
- (24) Kim, M. H.; Choi, W. C.; Kang, H. O.; Lee, J. S.; Kang, B. S.; Kim, K. J.; Derewenda, Z. S.; Oh, T. K.; Lee, C. H.; Lee, J. K. *Proc. Natl. Acad. Sci. USA.* **2005**, *102*, 17606.
- (25) Liao, R. Z.; Yu, J. G.; Himo, F. *Inorg. Chem.* **2009**, *48*, 1442.
- (26) National Data Intelligence Center. *The Economic Impact of Illicit Drug Umerican Society*; U.S. Department of Justice: Washington, DC, 2011; pp 137–161, 187–209.
- (27) Hartney, C.; National Council on Crime and Delinquency: 2006.
- (28) Webster, R. A. *Neurotransmitters, Drugs and Brain Function*; Wiley: New York, NY, 2001; pp 137–187.
- (29) Berridge, K. C. *Psychopharmacology (Berlin, Ger.)* **2007**, *191*, 391.
- (30) Bubar, M. J.; Cunningham, K. A. *Neuroscience* **2007**, *146*, 286.
- (31) Di Matteo, V.; Cacchio, M.; Di Giulio, C.; Esposito, E. *Pharmacol. Biochem. Behav.* **2002**, *71*, 727.
- (32) Gervais, J.; Rouillard, C. *Synapse* **2000**, *35*, 281.
- (33) Grottick, A. J.; Fletcher, P. J.; Higgins, G. A. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 1183.
- (34) Higgins, G. A.; Fletcher, P. J. *Eur. J. Pharmacol.* **2003**, *480*, 151.
- (35) Pierce, R. C.; Kumaresan, V. *Neurosci. Biobehav. Rev.* **2006**, *30*, 215.

- (36) Volkow, N. D.; Wang, G. J.; Fowler, J. S.; Tomasi, D.; Telang, F. *Proc. Natl. Acad. Sci. USA*. **2011**, *108*, 15037.
- (37) Alex, K. D.; Pehek, E. A. *Pharmacol. Ther.* **2007**, *113*, 296.
- (38) Xie, E.; Zhu, L.; Zhao, L.; Chang, L. S. *Genomics* **1996**, *35*, 551.
- (39) Berg, K. A.; Clarke, W. P.; Cunningham, K. A.; Spampinato, U. *Neuropharmacology* **2008**, *55*, 969.
- (40) Steck, P. A.; Pershouse, M. A.; Jasser, S. A.; Yung, W. K.; Lin, H.; Ligon, A. H.; Langford, L. A.; Baumgard, M. L.; Hattier, T.; Davis, T.; Frye, C.; Hu, R.; Swedlund, B.; Teng, D. H.; Tavtigian, S. V. *Nat. Genet.* **1997**, *15*, 356.
- (41) Teng, D. H.; Hu, R.; Lin, H.; Davis, T.; Iliev, D.; Frye, C.; Swedlund, B.; Hansen, K. L.; Vinson, V. L.; Gumpfer, K. L.; Ellis, L.; El-Naggar, A.; Frazier, M.; Jasser, S.; Langford, L. A.; Lee, J.; Mills, G. B.; Pershouse, M. A.; Pollack, R. E.; Tornos, C.; Troncoso, P.; Yung, W. K.; Fujii, G.; Berson, A.; Steck, P. A.; et al. *Cancer. Res.* **1997**, *57*, 5221.
- (42) Li, J.; Yen, C.; Liaw, D.; Podsypanina, K.; Bose, S.; Wang, S. I.; Puc, J.; Miliaresis, C.; Rodgers, L.; McCombie, R.; Bigner, S. H.; Giovanella, B. C.; Ittmann, M.; Tycko, B.; Hibshoosh, H.; Wigler, M. H.; Parsons, R. *Science* **1997**, *275*, 1943.
- (43) Di Cristofano, A.; Pesce, B.; Cordon-Cardo, C.; Pandolfi, P. P. *Nat. Genet.* **1998**, *19*, 348.
- (44) Li, D. M.; Sun, H. *Proc. Natl. Acad. Sci. USA*. **1998**, *95*, 15406.
- (45) Tamura, M.; Gu, J.; Matsumoto, K.; Aota, S.; Parsons, R.; Yamada, K. M. *Science* **1998**, *280*, 1614.
- (46) Podsypanina, K.; Ellenson, L. H.; Nemes, A.; Gu, J.; Tamura, M.; Yamada, K. M.; Cordon-Cardo, C.; Catoretti, G.; Fisher, P. E.; Parsons, R. *Proc. Natl. Acad. Sci. USA*. **1999**, *96*, 1563.
- (47) Eng, C. *Hum. Mutat.* **2003**, *22*, 183.
- (48) Radu, A.; Neubauer, V.; Akagi, T.; Hanafusa, H.; Georgescu, M. M. *Mol. Cell. Biol.* **2003**, *23*, 6139.
- (49) Lachyankar, M. B.; Sultana, N.; Schonhoff, C. M.; Mitra, P.; Poluha, W.; Lambert, S.; Quesenberry, P. J.; Litofsky, N. S.; Recht, L. D.; Nabi, R.; Miller, S. J.; Ohta, S.; Neel, B. G.; Ross, A. H. *J. Neurosci.* **2000**, *20*, 1404.



- (50) Lee, J. O.; Yang, H.; Georgescu, M. M.; Di Cristofano, A.; Maehama, T.; Shi, Y.; Dixon, J. E.; Pandolfi, P.; Pavletich, N. P. *Cell* **1999**, *99*, 323.
- (51) Maehama, T.; Dixon, J. E. *J. Biol. Chem.* **1998**, *273*, 13375.
- (52) Stambolic, V.; Suzuki, A.; de la Pompa, J. L.; Brothers, G. M.; Mirtsos, C.; Sasaki, T.; Ruland, J.; Penninger, J. M.; Siderovski, D. P.; Mak, T. W. *Cell* **1998**, *95*, 29.
- (53) Georgescu, M. M.; Kirsch, K. H.; Akagi, T.; Shishido, T.; Hanafusa, H. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10182.
- (54) Sun, H.; Lesche, R.; Li, D. M.; Liliental, J.; Zhang, H.; Gao, J.; Gavrilova, N.; Mueller, B.; Liu, X.; Wu, H. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6199.
- (55) Ji, S. P.; Zhang, Y.; Van Cleemput, J.; Jiang, W.; Liao, M.; Li, L.; Wan, Q.; Backstrom, J. R.; Zhang, X. *Nat. Med.* **2006**, *12*, 324.
- (56) Maillet, J. C.; Zhang, Y.; Li, X.; Zhang, X. *Prog. Brain. Res.* **2008**, *172*, 407.
- (57) Antosiewicz, J.; Briggs, J. M.; Elcock, A.H.; Gilson, M. K.; McCammon, J. A. *J. Comp. Chem.* **1996**, *17*, 1633.
- (58) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415.
- (59) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comp. Phys. Commun.* **1995**, *91*, 57.
- (60) Li, H. R.; A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704.
- (61) Olsson, M. H. M. S.; C.R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525.
- (62) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science: Oxford, England, 1987; pp 6–32.
- (63) Leach, A. R. *Molecular Modelling Principles And Applications*; Pearson Education: Harlow, England, 2001; pp 165–231.
- (64) Rapaport, D. C. *The Art of Molecular Dynamics Simulation*, 2nd ed.; Cambridge University Press: Cambridge, England, 2005; pp 45-82.
- (65) Scott, W. R. P.; Huenenberger, I. G.; Tironi, A. E.; Mark, S. R.; Billeter, S.R.; Fennel, J.; Torda, A. E.; Huber, T.; Krueger, P.; Van Gunsteren, W. F.; *Phys. Chem. A*, **2001**, *103*, 3596.

- (66) Mackerell Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher III W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Staub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B.* **1998**, *102*, 3586.
- (67) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.
- (68) Li, X.; Latour, R. A. *J. Comput. Chem.* **2010**, *0*.
- (69) Li, X.; Latour, R. A.; Stuart, S. J. *J. Chem. Phys.* **2009**, *130*
- (70) Tovchigrechko, A.; Wells, C. A.; Vakser, I. A. *Protein. Sci.* **2002**, *11*, 1888.
- (71) Vakser, I. A. *Protein. Eng.* **1995**, *8*, 371.
- (72) Vakser, I. A. *Biopolymers* **1996**, *39*, 455.
- (73) Vakser, I. A. *Protein Eng.* **1996**, *9*, 741.
- (74) Vakser, I. A. *Proteins* **1997**, *Suppl 1*, 226.
- (75) Vakser, I. A.; Aflalo, C. *Proteins* **1994**, *20*, 320.
- (76) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455.
- (77) Jain, A. K.; M. Narismha, N.; Flynn, P. J. *ACM Comput. Surv.* **1999**, *31*, 265.
- (78) Jain, A. K.; *Pattern Recog. Lett.* **2010**, *31*, 651.
- (79) Wold, S. E., K.; Geladi, P. *Chemometr. Intell. Lab.* **1987**, *2*, 37.
- (80) van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., Berendsen, H. J. C. *J. Comp. Chem.* **2005**, *26*, 1701.
- (81) Jarvis, R. A.; Patrick, E. A. *IEEE T. Comput.* **1973**, *C-22*.
- (82) Jolliffe, I. T. *Principal Component Analysis*, 2nd ed.; Springer: New York, NY, 2002; pp 32-55.
- (83) Shiens, J.; “A Tutorial on Principal Component Analysis”; Center for Neural Science, New York University; New York, NY, 2009.
- (84) Abdi, H. W.; L. J. *Wiley Interdiscip Rev. Comput. Stat.* **2010**, *2*, 2.

- (85) Han, T. X. "Computing Principal Components with the Eigen-Decomposition of a Low Dimensional Matrix." *Cognitive Computer Vision*. Electrical & Computer Engineering Department; University of Missouri; Columbia, MO, n.d. Web. 20 Apr. 2013.
- (86) Wavefunction, Inc.: Irvine, CA, 2006.
- (87) Schuettelkopf, A. W., van Aalten, D.M.F. *Acta. Crystallogr.* **2004**, *D60*, 1335.
- (88) Ferguson, D. M. *J. Comput. Chem.* **1995**, *16*, 501.
- (89) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comp. Chem.* **1997**, *18*, 1463.
- (90) Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- (91) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545.
- (92) Aqvist, J.; Marelus, J. *Comb. Chem. High. Throughput. Screen.* **2001**, *4*, 613.
- (93) Marelus, J.; Kolmodin, K.; Feierberg, I.; Aqvist, J. *J. Mol. Graph. Model* **1998**, *16*, 213.
- (94) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Wallingford, CT, **2004**.
- (95) Starkey, R. N., J; Hintzw, M. *J. Chem. Educ.* **1986**, *63*, 473.

- (96) Aubert, S. D.; Li, Y.; Raushel, F. M. *Biochemistry* **2004**, *43*, 5707.
- (97) Kaminskaia, N. V. H., C; Lippard, S.J. *Inorg. Chem.* **2000**, *39*, 3365.
- (98) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781.
- (99) Glykos, N. M. *J. Comput. Chem.* **2006**, *27*, 1765.
- (100) Anastasio, N. C.; Gilbertson, S. R.; Bubar, M. J.; Agarkov, A.; Stutz, S. J.; Jeng, Y.; Bremer, N. M.; Smith, T. D.; Fox, R. G.; Swinford, S. E.; Seitz, P. K.; Charendoff, M. N.; Craft, J. W., Jr.; Laezza, F. M.; Watson, C. S.; Briggs, J. M.; Cunningham, K. A. *J. Neurosci.* **2013**, *33*, 1615.
- (101) Meissner, R. S.; Perkins, J. J.; Duong le, T.; Hartman, G. D.; Hoffman, W. F.; Huff, J. R.; Ihle, N. C.; Leu, C. T.; Nagy, R. M.; Naylor-Olsen, A.; Rodan, G. A.; Rodan, S. B.; Whitman, D. B.; Wesolowski, G. A.; Duggan, M. E. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 25.
- (102) Urbahns, K.; Harter, M.; Vaupel, A.; Albers, M.; Schmidt, D.; Bruggemeier, U.; Stelte-Ludwig, B.; Gerdes, C.; Tsujishita, H. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1071.
- (103) Greenspoon, N.; HersHKoviz, R.; Alon, R.; Varon, D.; Shenkman, B.; Marx, G.; Federman, S.; Kapustina, G.; Lider, O. *Biochemistry* **1993**, *32*, 1001.
- (104) Jones, R. M. B., P. D.; Semple, G.; Shin, Y. J.; Tamura, S.Y. *Curr. Opin. Pharmacol.* **2003**, *3*, 530.
- (105) Breccia, P.; Boggetto, N.; Perez-Fernandez, R.; Van Gool, M.; Takahashi, M.; Rene, L.; Prados, P.; Badet, B.; Reboud-Ravaux, M.; de Mendoza, J. *J. Med. Chem.* **2003**, *46*, 5196.
- (106) Prabhakaran, E. N.; Rao, I. N.; Boruah, A.; Iqbal, J. *J. Org. Chem.* **2002**, *67*, 8247.