November 2019

Targeting Bacterial Resistance and Cancer Metastasis: A Structure Based Approach

Kyle Galen Kroeck
University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/etd
Part of the Biochemistry Commons, Biophysics Commons, and the Microbiology Commons

Scholar Commons Citation
https://digitalcommons.usf.edu/etd/8043

This Dissertation is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.
Targeting Bacterial Resistance and Cancer Metastasis: A Structure Based Approach

by

Kyle Galen Kroeck

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Department of Molecular Medicine
College of Medicine
University of South Florida

Major Professor: Yu Chen, Ph.D.
Robert Deschennes, Ph.D.
Wayne Guida, Ph.D.
Gloria Ferreira, Ph.D.
John Chodera, Ph.D.

Date of Approval:
April 30, 2019

Keywords: chemokines, lipid a biosynthesis, molecular docking, x-ray crystallography, temocillin, ticarcillin, beta lactamase, penicillin binding protein

Copyright © 2019, Kyle Galen Kroeck
Dedication

To my loving grandfather, Louis.
Acknowledgements

I first would like to acknowledge and thank my family for their constant support and always pushing me to strive for my highest potential. I would have never achieved all that I have without them. Next I would like to thank everyone in my lab who have helped me throughout my time in graduate school and were instrumental in my completion of this program: Dr. Emmanuel Smith, Dr. Eric Lewandowski, Dr. Xiujun Zhang, Dr. Orville Pemberton, Dr. Derek Nichols, Nicholas Torelli, Daniel Shoun, Michael Kemp, Michael Sacco, and Afroza Ahktar. I would also like to thank all of my committee members: Dr. Robert Deschenes, Dr. Wayne Guida, and Dr. Ferreira. They have all been incredible teachers and I am extremely grateful for their time and guidance throughout the process of getting through the Ph.D. program. I thank Dr. John Chodera for serving as my committee chair and offering his time for the special occasion. Additionally, I would like to thank all my friends in the program who have been unwavering in their emotional support, understanding, and offering of help when I needed it most: Dr. Adonis McQueen, Ahmad Jalloh, and Dr. Antwoine Flowers. Lastly, but perhaps most importantly, I would like thank Dr. Yu Chen. He has not only been the most incredible mentor I have ever had, but also an incredible friend. I have never met anyone quite like him and I doubt that I ever will. Without his help I would have never made it this far. He truly has helped mold me into the successful and adept scientist and person I am today.
Table of Contents

List of Tables ................................................................................................................................ iv
List of Figures .................................................................................................................................. v
Abstract .......................................................................................................................................... vi

Chapter 1: Introduction ....................................................................................................................1
  1.1 Bacterial Resistance to Antibiotic Treatments...............................................................1
  1.2 Cancer and Metastasis....................................................................................................3
  1.3 Structure-Based Drug Design ........................................................................................4
    1.3.1 Fragment-Based Drug Design.........................................................................7
  1.4 Molecular Docking ........................................................................................................8
    1.4.1 ZINC Database..............................................................................................11

Chapter 2: Dual Targeting the Lipid A Biosynthesis Pathway ......................................................13
  2.1 Overview ......................................................................................................................13
  2.2 Introduction ..................................................................................................................14
    2.2.1 The Problem of Antibiotic Resistance ..........................................................14
    2.2.2 LPS Components .........................................................................................14
    2.2.3 Lipid A Biosynthesis .....................................................................................15
    2.2.4 Targeting Lipid A Biosynthesis ....................................................................15
  2.3 Results ..........................................................................................................................16
    2.3.1 LpxD Crystallization .....................................................................................16
    2.3.2 Structure-Based Inhibitor Discovery ............................................................17
    2.3.3 Complex Crystal Structures of Novel Inhibitors ..........................................18
  2.4 Discussion ....................................................................................................................21
    2.4.1 Dual-Targeting Inhibitor Development ........................................................22
    2.4.2 Allosteric Regulation in LpxD Reaction.......................................................23
    2.4.3 Issues of Drug Permeability in Targeting Lipid A Biosynthesis .................24
  2.5 Conclusions ..................................................................................................................26
  2.6 Experimental Procedures .............................................................................................27
    2.6.1 Purification of Recombinant LpxA ...............................................................27
    2.6.2 Purification of Recombinant LpxD ...............................................................28
    2.6.3 Purification of Biotinylated Avidity-Avi Tagged LpxA and LpxD..............29
    2.6.4 LpxA Crystallization .....................................................................................30
    2.6.5 LpxD Crystallization .....................................................................................31
    2.6.6 Data Collection and Processing ....................................................................32
    2.6.7 Surface Plasmon Resonance .........................................................................32
List of Tables

Table 2.1: Surface Plasmon Resonance Assay Hits.................................................................34
Table 2.2: X-Ray Data Collection and Refinement Statistics..................................................35
Table 4.1: X-Ray Data Collection and Refinement Statistics..................................................87
Table 4.2: CCL21 NMR Titration Assay Hit Compounds ......................................................88
List of Figures

Figure 2.1: Lipid A Biosynthesis Pathway .................................................................36
Figure 2.2: Magnesium Ion and Coordinated Waters at the Core of LpxD Trimer .................37
Figure 2.3: LpxA Complex Structures with Novel Inhibitors .............................................39
Figure 2.4: LpxD Complex Structure with Compound 1 ..................................................40
Figure 2.5: Structural Alignments of P. aeruginosa LpxA & LpxD ........................................41
Figure 3.1: Diagram of Enzymatic Activity of PBP Enzymes ..............................................58
Figure 3.2: Diagram of Serine β-Lactamase Catalytic Mechanism ......................................59
Figure 3.3: Chemical Structures of Temocillin and Ticarcillin ............................................59
Figure 3.4: Thermal Shift Assay for P. Aeruginosa PBP3 Incubated with Temocillin and Ticarcillin .................................................................60
Figure 3.5: Unbiased Fo-Fc map of P. aeruginosa PBP3 Complex Structures .......................61
Figure 3.6: CTX-M-14 WT Complex Structure with Temocillin ........................................61
Figure 3.7: PBP3 Complexes Structure Alignment .............................................................62
Figure 4.1: Superimposition of Truncated CCL21 with Truncated CCL19 .........................89
Figure 4.2: Chemokine Receptor Binding Mechanism .......................................................90
Figure 4.3: Truncated CCL21 (Residues 1-79) Crystal Structure .......................................91
Figure 4.4: CCL21 Sulfotyrosine Recognition Site NMR Analysis ......................................92
Figure 4.5: Superimposition of CCL21 Crystal Structure with CCL21 NMR Structure ..........93
Figure 4.6: Superimposition of CCL19 Crystal Structure with CCL19 NMR Structure ..........94
Abstract

Current research in pharmaceutical development commonly utilizes a profusion of methods in molecular modeling in order to probe intricate biological problems. Many original and promising compounds have been identified and developed by integrating experimental and computational methods. Structural biology utilizes many different research techniques including x-ray crystallography, NMR, and electron microscopy in order to develop molecular models of macromolecules that are of biological interest. Such techniques can be used in conjunction with molecular docking, which utilizes those molecular models in order to target macromolecules of therapeutic interest by computationally analyzing the conformations adopted by ligands upon interaction with a desired binding site and estimating the free energy of binding. This technique allows for the screening of millions of compounds with great variety in terms of structure and chemotype. The initial hits of such drug discovery efforts generally consist of low affinity small molecules, but by developing complex structures of these compounds with the macromolecular target they can be optimized through the addition of functional groups and enlarging the compound structure in order to take advantage of the chemical space surrounding the inhibitor within the binding site, which leads to higher affinity compounds. This is the process of a structure based drug design effort and the work herein utilizes this process in order to develop and optimize small molecule inhibitors that offer the potential to be utilized in battling bacterial resistance to current antibiotics and preventing the metastasis of different cancers.
The first project seeks to target the lipid A biosynthetic pathway which is a highly conserved biochemical process seen in all Gram-negative bacterial cells and has shown to be essential for the growth and survival of Pseudomonas aeruginosa. Two of the enzymes involved in this biochemical pathway, LpxA and LpxD, display high structural similarity and catalyze highly similar reactions which offers the unique chance to potentially develop dual activity inhibitors that bind to both enzymes. Such small molecules should lead to an increase in inhibitory effects and lower the possibility of mutational resistance to develop against such activity. By developing x-ray crystallographic structures of the LpxA and LpxD enzymes we were able to use molecular docking to develop several initial hit compounds that bind to the active sites of both these proteins. The binding affinity for these compounds was determined using surface plasmon resonance and complex structures of the compounds with LpxA and LpxD were developed. These structures revealed a previously undetermined magnesium at the core of the LpxD trimer which may have biological relevance to the structure of the protein. Additionally, one of the inhibitors displayed allosteric effects upon binding to LpxD by inducing conformational changes in the C-terminal alpha-helical domain of the protein, which is responsible for forming substantial contacts with acyl carrier protein (ACP), a protein that functions by depositing the necessary catalytic substrate within the active site of LpxD.

The second project also focused on targeting bacterial resistance in Pseudomonas aeruginosa but wanted to exploit the bacteria’s required functioning of its PBP3 enzyme for growth. PBP3 is a penicillin binding protein, enzymes involved in the catalytic formation of the bacterial cell wall. However, rather than developing new inhibitors with this project we sought to figure out why a known antibiotic, temocillin, displays broad spectrum activity against most resistant strains of bacteria but is limited when it comes to the ability to combat Pseudomonas
aeruginosa infections. We did so by analyzing how the 6-α-methoxy group on the compound’s β-
lactam ring influences the ability to bind to PBP3. Furthermore, we wanted to analyze the
mechanism behind temocillin’s ability to counter antibiotic resistance. Looking at the influence
of the 6-α-methoxy group involved looking at the binding kinetics of temocillin and ticarcillin (a
parent compound of temocillin which varies structurally only in missing that 6-α-methoxy
group) through fluorescence polarization assays, along with thermal shift assays, and the
development of complex crystal structures between PBP3 and both compounds. In order to
determine the mechanism behind temocillin’s ability to counter bacterial resistance we
developed a complex structure of temocillin with CTX-M-14, a β-lactamase enzyme which is
responsible for the catalytic degradation of many known antibiotic compounds. These studies
will allow for the ability to improve and possibly develop new broad spectrum β-lactamase
inhibitors and also hopefully lead to the development of new therapeutics that target clinically
significant antibiotic resistant strains of Pseudomonas aeruginosa.

The final project sought to develop small molecule inhibitors that targeted the CCR7
signaling axis by targeting the protein-protein interfaces (PPIs) of the receptors cognate
chemokines, CCL19 and CCL21. Both of these signaling proteins induce their effects by binding
to the CCR7 receptor, which induces signaling cascades that result in cell motility and thus plays
an important role in the trafficking of immune cells from the bloodstream into the lymphatic
system. This signaling axis gets hijacked by cancer cells and plays a role in supporting tumor
growth, angiogenesis, and metastasis of various cancers. Uniquely, sulfation of the N-terminal
tyrosine residues of the CCR7 receptor plays a critical role in the binding of the chemokines by
impacting large amounts of affinity for the interaction. The recognition sites for these sulfated
tyrosines on the chemokines are highly conserved across this family of proteins and make for
ideal hot spots in the development of small molecule inhibitors to prevent these protein-protein interactions. In order to identify these recognition sites we used NMR perturbation assays to outline the binding pockets on the protein structure. We then developed x-ray crystallographic structures to utilize in molecular docking studies that screened for compounds that would bind to these surfaces. Further NMR perturbation assays were used to determine the binding constants of our selected compounds from the molecular docking experiments. The compounds developed by our efforts represent the first CCL21-specific inhibitors and support the hypothesis that the conserved sulfotyrosine recognition sites found on the chemokine family make for good small molecule drug targets in an effort to develop inhibitors that target these various signaling axes. Additionally, our crystal structures of these chemokines shed light on various structural mechanisms that play a role in these molecules ability to bind to their target receptor and give further insight into a basic understanding of potential regulatory mechanisms for altering their signaling.
Chapter 1:

Introduction

1.1 Bacterial Resistance to Antibiotic Treatments

In the United States alone, 2 million people annually develop severe bacterial infections that are resistant to one or more antibiotics currently used to treat that infection. Out of those 2 million people, 23,000 cases prove to be fatal as a direct result of their resistant infection.\(^1\) In addition to the risk of mortality, these incidents of infection result in huge financial and economic costs as well. Each case of antibiotic resistant bacterial infection adds an estimated $1,383 to the cost of treating a patient, which amounted to over $2 billion in added costs in the year 2014.\(^3\) This figure is double the amount spent in treating such infections in 2002, which can be seen when comparing the incidents of resistant infections between these years. While the number of bacterial infections remained relatively similar (13.5 million in 2002 and 14.3 million in 2014), the percentage of these infections that were resistant to antibiotic treatment went from just over five percent in 2002 (700,000 infections) to eleven percent in 2014 (1.6 million infections).\(^3\) The reason behind the spike in these types infections is multifactorial. One of the factors is the lack of novel antibiotic research, which has seen no new developments since carbapenems and fluoroquinolones were first introduced in the 1980s.\(^4,5\) The number of approved novel antibiotic treatments has been steadily decreasing each decade since the 1980s when over 30 new therapeutics were released for use, dropping to just a little over 20 in the 1990s, and
under 10 between 2000 and 2012. The reason for the lack of innovation is that financial investment into research and development of new antibacterial medications is viewed by the pharmaceutical industry to not be a worthwhile endeavor. The lack of financial viability for developing new antibacterial medications has been supported by economic research that indicated the average net value of a novel antibiotic drug for a company is significantly less than that of the value of developing pharmaceuticals for chronic diseases. It isn’t just a lack of new drugs though that is causing this spike in untreatable infections; it is also due to the enormous and irresponsible overuse of antibiotics, which has been shown to be the major driving force behind the development of resistant strains. Small epidemiological studies in hospitals have identified a direct correlation between antibiotic use and increased prevalence for bacterial resistance amongst the patients. Part of this overuse problem is the result of inappropriate prescriptions by doctors, as shown by studies where medical practitioners’ choice of antibiotic, length of antibiotic treatment, or antibiotic indications were incorrect 30%-50% of the time. These misdiagnoses can frequently lead to sub-inhibitory concentrations of the antibacterial in the patient’s bloodstream and this results in the introduction of alternative forms of mutagenesis, genetic expression, and horizontal gene transfer; all of which induce and propagate antibiotic resistance. Another major contributing factor to rise in rates of antibiotic resistance is the administration of antimicrobial agents to livestock, where they are used in prophylactic treatment of the animals and also as growth promoters. It is estimated that as much as 80% of the antibiotics sold in the United States are used to treat livestock. These antibiotics and the resistant bacteria found in the animals, as a result of their prophylactic treatment with the antibiotics, are then transferred to humans via the consumption of these animals and their byproducts. This combination of increased rates for development of antibiotic resistance and
lack of new innovation for drugs combatting these resistant strains of infection is poised to become a major health crisis unless money is invested into research to counteract resistance immediately.

1.2 Cancer and Metastasis

Cancer is the second leading cause of death worldwide and continues to elude a cure.\textsuperscript{14} This is largely because cancer is actually an extensive group of diseases with great degrees in variability with regards to the specific mechanisms of cellular deregulation by which they occur.\textsuperscript{15} In broad terms, all cancers are the result of mutations within the regulatory genes responsible for cell proliferation and apoptosis, which causes normal cell functioning to cease and continual cell replication to occur in conjunction with a blocking of proliferative inhibition and auto-apoptosis.\textsuperscript{16} As the tumor grows it may also spread to different areas of the body, and this spread is the cause of 90\% of cancer related deaths.\textsuperscript{17} Metastatic cancers are significantly more difficult to get rid of due to the need to remove as many of the malignant cells throughout the body in order to enter remission and prevent recurrence of tumor formation. This along with the increased risk of complications caused by the spread of cancer cells can often limit treatment options.\textsuperscript{18}

Globally, there were a total of 17.2 million cases of cancer in 2016 and it caused a total of 8.9 million deaths that same year. This number has continued to rise as the estimated number of deaths caused by cancer in 2018 was put at 9.6 million.\textsuperscript{15,19} In 2016 alone, cancer caused a loss of 213.2 million disability-adjusted life-years worldwide (the sum of years lived with disability and the years of life lost as a the result of a disease), which is a prevalent measure of the burden of cancer.\textsuperscript{19} The economic burden of cancer is also large; in 2015 the total annual medical costs caused by these diseases in the United States alone was $80.2 billion.\textsuperscript{20} The estimated cost
worldwide of cancer treatment in 2010 was $1.16 trillion dollars and as incidence rates rise (28% increase from 2006 to 2016) this number is sure to continue growing.\textsuperscript{15,19}

### 1.3 Structure-Based Drug Design

Since the 1960s when the world saw the development of the first protein structures solved using x-ray crystallography, structural data has been used to help design more effective pharmaceuticals in an effort to more specifically target the pathways known to be involved with various disease states.\textsuperscript{21,22} These pioneering efforts to understand protein folding, the evolution of molecules, and the relationship between the structure and function of these proteins also led to the structural identification of many potential binding sites for the development of various inhibitors, and laid the essential groundwork for the now pullulating field of structure-based drug design (SBDD).\textsuperscript{21-24} There are many reasons why SBDD has grown so rapidly over the past few decades, but it is primarily due to the exponential improvement of all the research techniques used in the implementation of SBDD research. Perhaps the most evident of these improvements can be seen in the three main techniques used for the development of structural models: (1) x-ray crystallography with now widely available access to high intensity x-ray sources and highly sensitive detectors capable of sub-angstrom resolution detection\textsuperscript{25} (2) nuclear magnetic resonance (NMR) with improvements to sample preparation and access to high powered NMR spectrometers\textsuperscript{26} (3) cryogenic electron microscopy (cryo-EM) that has also seen vast improvements to sample preparation techniques, along with increased power and stability of the electron gun and higher sensitivity of detectors in the electron microscopes as well.\textsuperscript{27,28} These advances have resulted in a massive surge of structure depositions into the Protein Data Bank, with nearly 150,000 structures deposited to date and over half of those have been deposited in the past eight years.\textsuperscript{29} The improvement of these structures quality and the massive increase in
total number of structures has greatly helped increase the accuracy of virtual screening by providing much more precise and accurate data and in greater amounts, thus allowing for significant improvements to be made to *in silico* algorithms for predicting where and how compounds will bind to desired targets.\textsuperscript{30-32} Additionally, the exponential increase seen in availability of computing power has also allowed for more accurate prediction of compound binding with little sacrifice to speed of the virtual screen, due to less need for approximations to be made in the docking parameters for the sake of time efficiency.\textsuperscript{31,33} This acceleration of computing speed has also helped increase the speed of the overall process of finding successful lead candidates, along with the rapid growth of virtual compound libraries which allow for far greater chemical sampling than ever before.\textsuperscript{34,35}

The process of SBDD is a seemingly straightforward multistep process that is complicated by common hindrances in the process that require reiterative troubleshooting techniques to overcome them.\textsuperscript{34,36} The first step in this process is always identifying a target protein of interest and, either through previous research or experimental methods, obtaining a structural model of the protein. If experimental methods are required, the first common hindrance is usually the purification of the protein. The procedures used depend on the type of protein and its various properties, but most often problems arise as a result of solubility issues.\textsuperscript{36,37} Upon obtaining a purified sample the process of determining the protein structure begins; this typically entails using x-ray crystallography, NMR, or cryo-EM.

Upon having structural information for a protein of interest the next step is identifying a target site for drug development. Most commonly this is an enzymatic active site, but can also be protein-protein interfaces, natural ligand binding sites, sites that induce allosteric regulation of protein function, or computationally determined hotspots that are favorable for the binding of
small molecules.\textsuperscript{34} Once a target site for drug development has been identified, screening for hit compounds can begin. The process of screening is traditionally carried out through either high throughput screening (HTS) or virtual screening (VS). When using HTS researchers screen large compound libraries (typically >1000 compounds) with an assay that can appropriately measure the activity of these compounds against the target of interest in order to establish binding affinities for the compounds.\textsuperscript{38,39} The strength of using HTS is the ability to obtain biochemical or biophysical data for the binding of compounds to the target of interest. The drawbacks to this method are the costs associated with it and the relatively low success rates of lead hits making it through the optimization process due to poor ADME (absorption, distribution, metabolism, excretion) and toxicity profiles, especially given the amount of time, effort, and resources needed when carrying out such experiments.\textsuperscript{39} As previously mentioned, the other option is virtual screening which employs using the computational model of the target of interest, outlining the previously identified binding site, and using computer simulations to screen massive digital compound libraries (typically > 4 million compounds) for their ability to bind to that site. These simulations are carried out using force field calculations and simplified quantum mechanical algorithms to determine the free energy of ligand binding and rank the ligands in order of likelihood of being able to bind to the target protein.\textsuperscript{40} The drawbacks to this method are the lack of biochemical and biophysical data and the error rate resulting from the need to ignore certain variables in the calculations due to the complexity of calculating protein-ligand interactions and the relatively limited computational power of modern day computers.\textsuperscript{40-42}

After having identified lead hit compounds, the next goal is to optimize these compounds in order to improve their specificity and affinity for the protein of interest at the target binding site.\textsuperscript{34} This is accomplished first by determining the actual structure of the ligand-protein
complex in order to analyze all the molecular interactions that are involved in the binding of the ligand to the target protein. Generally these structures are determined using x-ray crystallography, but with recent increases in the resolution capabilities of cryo-EM, complex structures have been determined using this technique as well. The process of optimization is carried out by making additional increases to affinity, along with increases to specificity of the ligand for the target binding site by adding functional groups to the ligand that can interact with other nearby residues thus increasing the number of favorable interactions with the protein. Functional groups may also be swapped out from the initial hit in an effort to improve ADME profiles, reduce toxicity of the compound, to ensure it is not a PAINS (pan-assay interference compound), to reduce potential off target effects, or any other reason that could help improve the quality of the hit compound based on the structural and known data about the target protein that is relevant. Another way chemical space may be explored is simply by searching for analogues of the hit compound through various compound libraries and then testing those analogues against the target protein either using an assay or through VS. Regardless of how the novel compound hits are found, the reiterative process of drug optimization then generally starts all over again and this process is repeated until one of the compounds obtained reaches a satisfactory level of affinity and specificity for the target of interest.

1.3.1 Fragment-Based Drug Design

Fragment-based drug design (FBDD) is a form of SBDD but utilizes small fragment molecules (compounds with molecular weights generally < 300 Da) as the initial starting point for drug discovery. The benefits of FBDD are: the reduction in necessary computational power when using virtual screening due to significantly reduced conformational sampling required, precision in targeting specific subpockets within the desired ligand target site, thus capitalizing
on valuable chemical spaces, and providing acutely potent compounds through combining multiple fragments into a single high affinity compound.\textsuperscript{48} FBDD is restricted to NMR, Surface Plasmon Resonance (SPR), and x-ray crystallographic techniques for implementation and generally cannot be carried in conjunction with HTS. This is due to the high level of sensitivity necessary to use fragments (which is possible with NMR and the more advanced SPR systems) or the ability to use high concentration of compound (which is possible when developing complex structures with protein crystals) to overcome the generally weak affinity of the initials hits.\textsuperscript{47,48} FBDD has often become favored by many researchers over more traditional methods such as HTS for development of lead compounds due to the cost effectiveness of virtual screening, the increased amount of chemical diversity that can be screened using the FBDD approach, and the relatively high hit rate through rational prioritization and early elimination of false positives typically found in screening.\textsuperscript{49,50}

1.4 Molecular Docking

Molecular docking is a computational method of predicting the binding pose of molecules with various macromolecular structures.\textsuperscript{33,34,46} It allows for the screening of massive compound libraries against targets of interest and has become an invaluable tool for the process of SBDD.\textsuperscript{34,35} Docking programs are able to predict the binding poses of large compound libraries by using atomic models of the macromolecule that is being targeted in conjunction with simplified molecular mechanics to calculate force fields, which are parameters and functions that allow for the calculation of the potential energy of molecular systems.\textsuperscript{51} The potential energy calculations are used to probe which ligand conformations have the highest probability of binding to the target molecule and those calculations are used to rank the most likely conformations for all the ligands.\textsuperscript{52,53}
The first problem the docking program must confront is the flexibility of the millions of ligands in the compound library. In order to tackle the issue of flexibility, the docking program must find a way to sample all the possible conformations of ligands and interacting residues. All docking programs accomplish this by employing sampling algorithms. There are three main types of sampling algorithms that can be employed by docking programs: systematic, stochastic, and deterministic searches. Systematic searches sample all the degrees of freedom for the ligand; however, this methodology by itself can be incredibly computationally taxing so typically specific types of systematic searches are employed to improve efficiency, such as hierarchal searches. Hierarchal searches take advantage of the tautology of massive multi-conformation compound libraries by organizing ligands into hierarchal groups that allow the docking software to quickly eliminate common conformations, which would introduce steric clashes with the target and minimize redundant calculations. Stochastic search algorithms generate different conformations of the ligands by applying randomized alterations to the positioning and torsion angles of the compounds and then evaluate these alterations by using energy functions to determine if the position and torsion angles of each subsequently generated conformation is higher or lower in energy when compared to the previous one. These types of searches include the Monte Carlo algorithm with the Metropolis criterion and the genetic algorithm. Deterministic search algorithms attempt to increase the accuracy of virtual screening by not only accounting for ligand flexibility but receptor flexibility as well. These types of searches use molecular dynamics (MD) simulations and energy minimization, which make them incredibly time consuming to carry out. Due to the time consuming nature and computationally intensive requirements of MD simulations researchers have proposed using them in conjunction with
hierarchical search algorithms by running MD simulations on selected compounds found from an initial hierarchical search in order to determine which ligands to investigate further.\textsuperscript{60,61}

The problem of conformational sampling is not the only issue docking programs must overcome, they also face the challenge of estimating the binding affinity for the compounds in all of their possible conformations, and this is accomplished through the use of scoring functions.\textsuperscript{40,41,46} Scoring functions allow the docking software to determine the most energetically favorable ligand binding pose when interacting with the target site on the macromolecule of interest by assessing the estimated binding affinity of each sampled conformation.\textsuperscript{31,52,62} There are several categories of scoring functions, each with its own strengths and weaknesses: physics-based, empirical, knowledge-based, and descriptor-based.\textsuperscript{63} Physics-based scoring functions rely on molecular mechanics and force field parameters to estimate the binding affinity of the compound library.\textsuperscript{62,64} The force fields used in the potential energy calculations include the stretching of bonds, the energy associated with bond angle bending, the torsion or the energy associated with the rotation of all the bonds, all the non-bonded terms, which includes van der Waals and electrostatic interactions, the desolvation energy of the target molecule, and potential hydrogen bonds involved in the ligand-receptor interaction.\textsuperscript{51-53} Empirical scoring functions use various molecular mechanical components such as hydrophobic interactions, hydrogen bonds, and intramolecular strain but then weight these components with coefficients and other additional terms from experimentally determined data in order arrive at a final score.\textsuperscript{31,52,63} Knowledge-based scoring functions employ the use of statistics-based force field parameters to calculate the energy potentials of the ligand and the target molecule and then use an inverse Boltzmann analysis to determine the effect of the interatomic distance on these energy potentials.\textsuperscript{31,52,63} In descriptor-based scoring functions the
docking software utilizes previously determined quantitative structure-activity relationship (QSAR) data by giving descriptors to the properties of the individual ligands found in the compound library, the target molecule, and the interaction patterns found between these two. These descriptors can then be plugged into a machine-learning algorithm that will generate the most statistically probable binding scores for these ligand-target complexes.63

Molecular docking has become a critical part of any SBDD effort through its low cost, high rates of success, and ability to overcome complex problems through the use of conformational sampling algorithms and scoring functions.31,43,50 Molecular docking is not without its drawbacks and weaknesses, despite the rapid advances and surmounted obstacles within the field of computational chemistry. Due to the overwhelming complexity of any molecular interaction, especially when considering macromolecular interactions with ligands, accurate prediction of the behavior of these molecules using computational models remains a challenge of the highest order and the results of a molecular docking run still cannot be used as definitive evidence for said interactions.31,40,52 However, as molecular docking algorithms advance and computers become more powerful in terms of their speed and processing abilities this branch of research will flourish.

1.4.1 ZINC Database

The ZINC database is an online database of compounds that are able to be used for the purpose of molecular docking and are freely available for purchase. It is an open-access platform and used regularly by researchers for ligand discovery. The database is replete with over 500 million protomers in biologically relevant forms and ready for use in molecular docking simulations. ZINC is made up of compounds from 266 different vendors and 122 annotated catalogs and is frequently updated to ensure all compounds are still purchasable.65 It is also
functionally organized into tranches, which allow all of the compounds to be separated from one another based on wait time for delivery and various properties including molecular weight, log P values, reactivity, pH representations, and overall charge of the molecule. ZINC also offers predefined subsets consisting of common value ranges for the molecular weight and log P values, which include shards, fragments, lead-like, and drug-like. The most common ranges are the fragment (200-300 Da and log P values between -1 and 3.5) and lead-like (300-375 Da and log P values also between -1 and 3.5). The fragment subset offers lower affinity compounds that are ideal starting points for FBDD efforts that implement x-ray crystallography, NMR, and SPR due to the ability to use high levels of compound concentration that are normally unattainable when using bioactivity assays. The lead-like subset offers higher affinity compounds that are typically suited for use in inhibitor discovery efforts, but can also be used as initial points of inquiry for SBDD projects. In an effort to assist non-experts in chemoinformatics and computational biology, ZINC also allows the user to search for compounds with known bioactivity at varying stages of research (i.e., in vitro, in vivo, in trials, in man, in cells, and FDA approved), compounds with varying degrees of biogenic origins (i.e., endogenous, metabolite, or biogenic), and analogs of compounds of interest.
Chapter 2:

Dual Targeting the Lipid A Biosynthesis Pathway

2.1 Overview

The lipid A biosynthesis pathway is essential in *Pseudomonas aeruginosa*. LpxA and LpxD are the first and third enzymes in this pathway respectively, and are regarded as promising antibiotic targets. The unique structural similarities between these two enzymes make them suitable targets for dual-binding inhibitors, a characteristic that would decrease the likelihood of mutational resistance and increase cell-based activity. We report the discovery of multiple small molecule ligands that bind to *P. aeruginosa* LpxA and LpxD, including dual-binding ligands, demonstrating the feasibility of developing such small molecules targeting lipid A biosynthesis. The binding poses were determined for select compounds by X-ray crystallography. The new structures revealed a previously uncharacterized magnesium ion residing at the core of the LpxD trimer. In addition, ligand binding in the LpxD active site resulted in conformational changes in the distal C-terminal helix-bundle, which forms extensive contacts with acyl carrier protein (ACP) during catalysis. These ligand-dependent conformational changes suggest a potential allosteric influence of reaction intermediates on ACP binding, and vice versa. Taken together, the novel small molecule ligands and their crystal structures provide new chemical scaffolds for ligand discovery targeting lipid A biosynthesis, while revealing structural features of interest for future investigation of LpxD function.
2.2 Introduction

2.2.1 The Problem of Antibiotic Resistance

Antibiotic resistance is a worldwide threat that challenges our ability to successfully treat bacterial infections, reducing treatment success and exhausting health care resources.\(^1\) Infections caused by antibiotic-resistant Gram-negative bacteria (GNB) are particularly intractable due to the presence of an additional outer membrane that protects the bacterial cell from harsh environments and antibiotics.\(^6^7\) The outer leaflet of this membrane primarily consists of lipopolysaccharide (LPS).\(^6^8\) Although LPS is nonessential in some organisms (i.e. *A. baumannii*, *N. meningitidis*), it is essential in many important GNB such as *P. aeruginosa*, a common and potentially life-threatening nosocomial pathogen that is naturally resistant to many antibiotics.\(^1,6^7-7^1\) Compounds targeting LPS synthesis provide an excellent opportunity for the development of new antibiotics with a novel mechanism of action.

2.2.2 LPS Components

LPS is made of three components: 1) a linear chain of repeating saccharide units known as the O-antigen, 2) an oligosaccharide core domain, and 3) lipid A, a core glucosamine disaccharide that is connected to multiple fatty acid chains of various length.\(^6^8,7^1\) Lipid A has several interesting properties that underscore its importance to GNB; for one, lipid A is the minimal component of LPS required for cellular viability in most Gram-negative bacteria.\(^6^8,6^9,7^2\) Additionally, lipid A is the primary antigenic determinant of LPS and is the offending chemical species that precipitates septic shock.\(^6^8,7^3,7^4\)
2.2.3 Lipid A Biosynthesis

The Lipid A biosynthetic pathway, also known as the Raetz pathway, is highly conserved amongst all GNB. LpxA, LpxC and LpxD make up the first three enzymes in the Raetz pathway (Figure 2.1). In *P. aeruginosa*, LpxA catalyzes the first reversible first step, transferring a 10C hydroxydecanoate fatty acid to the uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) substrate through an acyl carrier protein (ACP). In the second step, LpxC catalyzes the zinc-dependent irreversible deacetylation of the LpxA product, producing UDP-3-O-(3-hydroxydecanoyl) glucosamine and committing the molecule to this pathway. LpxD is responsible for the reversible third step, in which a β-hydroxydodecanoate is transferred to the 2’ amine of UDP-3-O-(3-hydroxydecanoyl) via ACP. Six additional enzymatic steps are required before the completed lipid A product can be incorporated into the LPS molecule through attachment to the core component (Figure 2.1). While LpxC is completely dissimilar to LpxA and LpxD, the latter two enzymes share several unique structural features, consistent with their functional similarities in catalyzing the transfer of a 10 or 12 carbon chain fatty acid from ACP to UDP-GlcNAc through a concerted acid-base mechanism. Both proteins form biological homotrimers that contain a left-handed helix fold comprised of multiple parallel β-sheets. Although *P. aeruginosa* LpxA and LpxD only have 27% sequence identity, they exhibit highly conserved protein backbone and side chain features, particularly at the junctions of adjacent β-helix monomers which form the acyl chain binding pocket (Figure 2.5).

2.2.4 Targeting Lipid A Biosynthesis

LpxC has been extensively targeted in novel antibiotic discovery and many potent inhibitors with bactericidal properties currently exist. Whereas LpxA and LpxD are promising
drug targets themselves and are known to be essential in *E. coli*, inhibitor discovery against these enzymes has remained largely unexplored with no known small molecule inhibitors having ever been identified.\textsuperscript{69,76} The unique, shared structural similarities also make LpxA and LpxD amenable to dual-targeting inhibitors, which offers the advantages of increased potency and reduced likelihood of resistance formation.\textsuperscript{88} The concept of dual targeting early steps of the lipid A biosynthetic pathway has previously been supported with a peptide molecule RJPXD33, which was found to inhibit both LpxA and LpxD when expressed in *E. coli*.\textsuperscript{89} However, the clinical utility of such peptide inhibitors is limited due to their inability to penetrate the cell envelope. Herein we report the discovery of several small molecules that bind to both LpxA and LpxD with μm affinity, identified using a targeted structure-based methodology that utilizes molecular docking, Surface Plasmon Resonance (SPR) bioanalysis, and high-resolution crystallography. The structural analysis has offered valuable insights into not only inhibitor binding hot spots of LpxA and LpxD, but also allosteric effects of ligand binding in LpxD active site, which may play an important role in the LpxD reaction mechanism.

2.3 Results

2.3.1 LpxD Crystallization

The crystal structures of both Pae LpxA and LpxD have been determined previously.\textsuperscript{81,90} However, in the published Pae LpxD structure, the thrombin protease recognition sequence of the N-terminal His-tag linker is located in the active site, particularly, the uracil binding pocket. This obstructs the diffusion of small molecules into the active site, preventing the use of this construct in both functional and structural studies of ligand binding. After failed attempts to crystallize untagged LpxD following protease cleavage, we tested a few variations of this LpxD
construct by changing the thrombin protease cleavage site to TEV protease recognition sequence, and by excising the first two residues of the protein, His1 and Met2, which are located immediately after the N-terminal 20-AA protease site/hexahistidine tag. The resulting crystal structure is similar to the published one, with each asymmetric unit of the H3 space group containing one monomer which forms a biologically relevant homotrimer through 3-fold crystallographic symmetry operation. Importantly, the density of the His-tag linker is no longer observed in the active site. In fact, whereas the His-tag and the linker is ordered in the previous structure, this entire region is disordered in the current structure. However, our failure to crystallize untagged LpxD suggests the His-tag may still contribute to the stability of the crystal packing interface, despite the lack of an ordered conformation.

An interesting observation in our new LpxD structure is a well-defined magnesium ion in the core of the trimer (Figure 2.2). This magnesium ion, likely from the crystallization buffer, coordinates six water molecules and appears to be critical to the stability of the crystal, as removing magnesium from the crystallization buffer or chelating the magnesium with EDTA eradicates diffraction. A similar but slightly weaker density corresponding to this magnesium ion was observed in the previously published structure crystallized in the absence of magnesium in the crystallization buffer (PDB ID: 3PMO). Although it was modeled as water, the surrounding density peaks suggest that it may be a low-occupancy divalent metal ion coordinating six water molecules as observed in our new structure.

2.3.2 Structure-Based Inhibitor Discovery

In the absence of a functional high-throughput assay and no previously discovered small molecule inhibitors, virtual screening of the ZINC small molecule database was performed using
DOCK to identify potential compounds that would bind to the LpxA active site. Specifically, the acyl chain binding pocket was targeted because of key, shared residues between LpxA and LpxD. Multiple iterations of docking simulations were performed targeting LpxA, each with minor alterations to sampling space and side chain partial charge to favor conformations that would engage the targeted residues. Top scoring compounds (0.5%) were visually inspected. 25 of these compounds were selected and experimentally assessed against both LpxA and LpxD with SPR bioanalysis, a label-free biophysical technology capable of detecting ligand binding affinity and quantifying binding kinetics. The results from this study indicated 6 of the 25 selected compounds (24%) demonstrated micromolar affinity for LpxA, and 2 of these 25 (8%) retained similar affinity against LpxD (Table 2.1). R-3-hydroxydecanoyl and UDP-GlcNAc were included in the SPR assay as positive controls. While R-3-hydroxydecanoyl was a potent binder of LpxA (7.6 μM) and LpxD (41.1 μM), the affinity of UDP-GlcNAc for these two enzymes could not be determined by the SPR assay due to poor data quality.

2.3.3 Complex Crystal Structures of Novel Inhibitors

To elucidate the molecular interactions between these novel inhibitors and LpxA/D, complex crystal structures were determined for two inhibitors, including one crystallized with both proteins. The crystal structures of Pae LpxA in its apo and substrate-bound form have previously been solved in our laboratory at 1.8 Å and 2.15/2.3 Å, respectively. LpxA belongs to the P2₁2₁2₁ space group with an asymmetric unit consisting of six monomers that form two separate biologically relevant homotrimers through a noncrystallographic 3-fold symmetry (Table 2.2). This results in six actives sites per asymmetric unit, along with 6 copies of the ligand; one at each of the dimer interfaces present in the homotrimers. Interestingly, both ligands
have similar chemical characteristics, with two, bridged six-membered rings substituted with a 4-6 atom acyl chain and a terminal carboxyl group, a chemical species closely resembling its fatty acid substrate (Table 2.1).

For compound 1, two binding conformations for the ligand can be identified, differing from one another mainly in the location of the carboxyl group (Fig. 2.3). In one conformation (pose 1), the carboxyl group hydrogen bonds with Gln157 main chain and His156 side chain (Fig. 2.3A). This appears to be the primary conformation for the ligand since 4 out of the 6 copies present in the asymmetric unit assume this pose (Figure 2.3A). In the other conformation (pose 2), the ligand carboxyl group hydrogen bonds with a new conformation of Gln157 side chain (Figure 2.3B). The rest of the ligand, including the carbonyl group and the naphthalene ring, is positioned identically between the two conformations. In both conformations, compound 1 forms hydrogen bonds with both monomers at the dimer interface, including the aforementioned contacts with one monomer and a HB with Gly151 side chain from the other monomer. The naphthalene ring establishes multiple van der Waals interactions with residues Asn133, Tyr152, His156, Phe166, and Met169, pi-sigma interactions with Val132 and Ala138, a pi-stacking interaction with His118, and a pi-alkyl interaction with Ala136. The second observed binding conformation is relatively similar. Compared with the apo LpxA structure, one minor conformational change in the protein active site is also observed. This involves the so called “hydrocarbon ruler” Met169, a methionine residue that confers substrate specificity based on acyl chain length, which shifts upward in order to accommodate the naphthalene ring.

The compound 2 complex structure with LpxA was solved at 2.0 Å (Figure 2.3C). Corresponding ligand density is seen at all 6 active sites within the asymmetric unit, all adopting
the same conformation. Many of the interactions formed by compound 1 with LpxA are also seen with compound 2, including two hydrogen bonds with Gln157 and Gly151. The carboxyl group of compound 2 interacts with the backbone and side chain of His156, which shifts inwards to form a HB with its Nπ atom. The benzoxazine group of compound 2 forms pi-alkyl interactions with Ala136, Val132 and Ala138, and a pi-sulfur interaction with the hydrocarbon ruler, Met169. Unlike the naphthalene ring of compound 1 which is buried in the acyl chain pocket, the benzoxazine ring of compound 2 is partially drawn outwards, likely due to the H-bonding interactions between the substituted amide group of this ring and the catalytic residues His121 and Asp70.

Our sole LpxD complex structure with compound 1 (2.7 Å) showed distinguishable electron density corresponding to the ligand at the dimer interface formed via crystallographic symmetry (Figure 2.4). The fitted binding pose suggests that there are two critical hydrogen bonds for the protein-ligand complex formation, both with the backbone nitrogen of two glycine residues Gly272 and Gly278 from the adjacent monomers forming the active site. Additionally, compound 1 forms extensive van der Waals interactions with residues Met239, Ser259, Gly257, Gly272, Gly275, Val277, and Gly290, which are also critical for the hydrophobic contacts with the substrate acyl chain. Indeed, structural alignments reveal the decanoyl acyl chain of the substrate superimposes closely with the naphthalene ring of compound 1. Additional pi-alkyl interactions are seen with Ala253 and Ala254, amide-pi stacking interactions with Ile258 and Leu276, and pi-sulfur interactions with Met291 and Met293.

Two conformation changes are observed in LpxD upon ligand binding, one being Ser259. In its apo form, the Ser259 side chain normally points inwards towards the active site, however,
the naphthalene ring of compound 1 forces this residue to rotate inwards to avoid steric clash (Figure 2.4B). An additional, unexpected shift is also seen outside of the active site, in the distal α-helical ACP recognition domain (ARD) domain (Figure 2.4A), comparable to the changes observed between the apo structure of *E. coli* LpxD (PDB 3EH0) and the ACP bound forms (PDB 4IHF, 4IHG, and 4IHH). Given the functional role of this domain in binding and releasing ACP bearing one key substrate, it seems likely that this structural shift may play a role in the proper functioning of the enzyme.

Given the similarity of the binding pockets in LpxA and LpxD, it is not surprising that the binding modes for compound 1 are overall similar between the two proteins. Minor differences include the overall positioning of compound 1, which is slightly deeper in the LpxA active site with the naphthalene ring oriented horizontally, in contrast to LpxD, where the napthal ring faces vertically. Additionally, the acyl chain and terminal carboxyl group of the ligand in LpxD protrudes slightly out of the pocket (Figure 2.4), rather than upwards, as seen in LpxA (Figure 2.3). Similar intermolecular interactions are seen between structures, with three hydrogen bonds and five hydrophobic interactions formed in LpxA and two hydrogen bonds and seven hydrophobic interactions formed in LpxD.

### 2.4 Discussion

Despite the importance of LpxA and LpxD in Lipid A biosynthesis and their potential as antibiotic targets, small molecule inhibitor discovery has been lacking against these two proteins. Many details of their enzymatic reactions also remain unclear. Our results not only provide the first examples of small molecule inhibitors targeting LpxA and LpxD, but also demonstrate the possibility of designing dual-binding compounds active against both enzymes. In addition, our
new LpxD crystal structures shed new light on previously uncharacterized structural features, including an intramolecular network, that will deepen our understanding of LpxD catalysis.

2.4.1 Dual-Targeting Inhibitor Development

Both LpxA and LpxD catalyze the transfer of a 10- or 12-carbon acyl-chain from ACP to GlcNAc. Despite their low sequence identity, the two proteins share high structural similarity in the overall architecture of a homotrimer and in the active site residues, especially in the acyl-chain binding pockets that reside in the dimer interface of the β-helix trimer core. The similarities between the acyl-chain binding pockets of LpxA/D are highlighted by the discovery of a dual-inhibitor peptide for \textit{E. coli} LpxA and LpxD (with $K_d$ of 20 μM and 6 μM respectively),\textsuperscript{89} whose complex crystal structure with LpxA reveals potential hydrogen bond (HB) interactions with the protein backbone amide groups on the β strands as well as non-polar interactions with protein backbone and side chains (Fig. 2.3).\textsuperscript{92} The same backbone functional groups, as well as similar side chain moieties, can be found in the acyl-chain binding pocket of LpxD. Our recent determination of the \textit{P. aeruginosa} LpxA and LpxD crystal structures further demonstrates the conservation of binding hot spots between the two enzymes at the atomic level (Figure 2.5B). Similar to \textit{E. coli} proteins, the acyl-chain binding pockets present the same backbone functional groups in both enzymes from \textit{P. aeruginosa}. Many side chains are also conserved between the proteins (e.g., Gly151/Gly272 (from LpxA and LpxD respectively, same order below), Phe166/Phe287, His121/His242 (the catalytic histidine), or present the same functional groups (e.g., Cβ atom of Ser150/Ala271, non-polar hydrocarbons of Val132/Ala253 and Ile148/Met269).
The dual-binding activity of several of our novel inhibitors has demonstrated the feasibility of inhibitor discovery targeting LpxA and LpxD simultaneously. As shown by the complex structures of compound 1 bound to LpxA and LpxD, the inhibitor exhibited overall similar binding poses in the two acyl-chain binding pockets, revealing both structural similarities and differences that can guide future dual inhibitor development. The charged and polar functional groups of compound 1 interact with backbone amide groups shared by both LpxA and LpxD. The naphthalene ring is nestled in the hydrophobic pocket that recognizes the acyl-chain. Compared with the LpxA structure, the ligand aromatic ring goes deeper in the LpxD structure, consistent with a slightly larger acyl-chain binding pocket in LpxD. The two complex structures also shed light on additional shared binding hot spots that can be exploited for future lead optimization. Particularly, Phe166/Phe287, Val132/Ala253 and Ile148/Met269 provide large hydrophobic binding surfaces that can be very valuable for enhancing ligand binding affinity for both proteins.

2.4.2 Allosteric Regulation in LpxD Reaction

The reaction catalyzed by LpxD follows a sequential ordered mechanism where acyl-ACP binds first, and holo-ACP dissociates last after the acyl chain is transferred to UDP-GlcNAc. The recent determination of a series of E. coli ACP crystal structures in various complexes with LpxD provided important insights into the complex interactions between ACP and LpxD, and ligand conformations in the acyl-chain binding pocket crucial to ACP release. Aside from a β-coil motif, the ARD domain consists mostly of the C-terminal α-helix. One of the most significant observations in the LpxD inhibitor complex is how ligand binding in the acyl-chain binding pocket triggers conformational change in the distal C-terminal helix,
suggesting crosstalk between the catalytic center and ACP binding site. This suggests potential allosteric regulation of substrate binding in the active site by protein-protein contacts between ACP and LpxD, and vice versa, which can possibly play a role in dictating the order of the catalytic events in LpxD reaction.

Our new LpxD crystal structures also shed light on a previously uncharacterized magnesium ion at the center of the trimer core. Although the presence of magnesium in the crystallization buffer undoubtedly facilitate its presence in our crystal structure, electron densities corresponding to this magnesium and surrounding water molecules was observed in the previously determined LpxD structure as well. We hypothesize that magnesium ion can potentially stabilize the trimer. Interestingly, research has shown that low magnesium conditions in P. aeruginosa leads to alternative lipid A species with unique patterns of acylation through the removal of the R-3-hydroxydecanoic acyl chain. These alternate lipid A species are found present in isolates of patients with cystic fibrosis and seem to be an adaptation to infecting the airways of these hosts. Similar undiscovered lipidation patterns could occur in response to low magnesium conditions in different environments through the mechanism of destabilizing the LpxD trimer. However, further investigation is needed to probe the exact role of magnesium in LpxD structure and function.

2.4.3 Issues of Drug Permeability in Targeting Lipid A Biosynthesis

Both LpxA and LpxD are located on the cytoplasmic side of the inner membrane of the gram negative bacterial cell, despite the final destination of lipid A being the outer membrane. Bacteria overcome the membrane barriers between the location of LPS synthesis and final cellular localization by having LPS export proteins that transport the intermediates and final
product across the inner and outer membranes. These export proteins include MsbA, the lipid A core flippase. They also include the Lpt proteins, which are a group of proteins that form multiple protein complexes in order to act as a multifaceted transenvelope protein complex capable of transporting LPS across both membranes. Due to subcellular localization of LpxA and LpxD, and the resultant protection from the multiple cell layers separating the cytosol from the extracellular space, utilizing these proteins as drug targets poses a unique challenge. Research has shown that even when compared to other Gram negative bacteria, 

The challenge of permeating the gram-negative bacterial cell can be overcome though through various techniques that improve the uptake, or decrease the efflux, of bactericidal compounds. One method is utilizing rational drug design, such as in the recent study that analyzed the penetration of carbapenem in 

Using this information, researchers were able to create a series of analogues for carbapenem that allowed for the probing of new structural features to add to the molecule that would allow for it to maintain inhibitory activity while reducing dependence on OccD1 proteins for translocation, and thus increase permeability towards 

Another method for increasing drug permeability in bacterial cells is through the use of Siderophore-conjugated antibiotic systems. Siderophore is a small, high affinity, iron chelator which serves to transport iron across cell membranes. By attaching antibiotics to Siderophore, the bacterial transport systems used for iron-uptake can be taken advantage of in order to transport the conjugated antibiotic into the periplasmic or cytoplasmic spaces of the cell. One other method for tackling the issue of drug permeability is through the use of efflux pump inhibitors (EPIs) given in conjunction with
antibiotic medications, such as PAβN and other similar arylamines, which act as broad-spectrum EPIs, or the *P. aeruginosa* MexB specific pyridopyrimidines.\textsuperscript{100-102} By using EPIs the need for high concentrations of compound to reach the intended target is reduced, rather than increasing the ability of the drug to enter the system of the bacterial cell more effectively.

While utilization of the methods for bypassing low permeability in bacterial cells are beyond the scope of this project, their application in future research may be worth considering should drug permeability become an issue with the newly discovered compounds targeting LpxA and LpxD. This would first require cell-based assay testing of the inhibitors and determining if the negative results are from a lack of inhibitory properties or from a lack of compound uptake into the bacterial cells. Upon establishing the reason for such results, steps can be taken to improve the current compounds and make them more effective in the treatment of *P. aeruginosa*.

### 2.5 Conclusions

LpxA and LpxD are both promising targets for new antibiotic agents against GNB bacteria such as *P. aeruginosa*. Our studies have identified novel scaffolds that can serve as starting scaffolds for future inhibitor discovery including dual-binding compounds that underscore the structural and functional similarities shared by these two proteins. Furthermore, the reported X-ray structures for LpxA and LpxD represent the first published structures of these enzymes complexed with non-substrate/product small molecule ligands. Though it was not the primary goal of this project, these complex structures reveal unique conformational changes that occur in these enzymes. Most notably is the α-helical shift in the ARD of LpxD following ligand binding; it is possible that the perturbation of intermolecular interactions between monomers at
the interface of the acyl-chain binding pocket induces this conformational shift, an event that would elucidate the mechanism of substrate binding and release.

2.6 Experimental Procedures

2.6.1 Purification of Recombinant LpxA

The plasmid pET28b (Achaogen) containing the N-terminal His-tagged \( P. aeruginosa \) LpxA sequence was transformed into Rosetta (DE3) pLysS cells (Novagen). The cells were incubated in 50 ml of LB media supplemented with 35 μg/mL chloramphenicol and 50 μg/mL kanamycin at 20°C overnight. Then 10 ml of overnight culture was added to 1 L of LB media containing 35 μg/mL chloramphenicol and 50 μg/mL kanamycin and incubated at 37 °C until the \( \text{OD}_{600} \) reached 0.6–0.8. The protein expression was induced with 0.5 mM IPTG, and incubation continued at 20 °C overnight. The culture is then centrifuged at 5,000 g for 10 minutes and the pellet is resuspended in 10 ml the lysis buffer (20 mM Tris-HCl, pH 8.4, 250 mM NaCl, 20 mM imidazole, and 10% glycerol). The resuspended pellet was then sonicated on a 10 second sonication/15 second rest cycle for a total of 15 minutes at an amplitude of 6. This is followed by centrifugation at 40,000g for 40 minutes at 4°C. The supernatant was then loaded to a HisTrap affinity column and eluted with a linear concentration gradient spanning 10-500 mM imidazole. The fractions containing LpxA were pooled and concentrated. The sample was loaded to a HiLoad 16/60 Superdex 75 column for further purification in thrombin cleavage buffer (20 mM Tris, pH 8.4, 150 mM NaCl, and 10% glycerol). The peak fractions containing the His-tagged LpxA were pooled, and the concentration of the protein was determined by \( \text{OD}^2_{280} \). The protease thrombin (Roche) was added at a ratio of one unit per milligram of the protein. Thrombin was selected as the protease due to it having high specificity and fidelity for the given
cleavage sites. After overnight incubation at room temperature, the samples were then loaded onto a HisTrap column to remove any uncleaved protein. The flow-through was collected and concentrated, followed by gel filtration with the HiLoad 16/60 Superdex 75 column. The protein eluted at a peak consistent with the size of the trimeric form. The untagged LpxA was stored at −80 °C at 20.0 mg/mL concentration in a buffer containing 20 mM potassium phosphate pH 8.6 and 250 mM NaCl. The purity of the protein was determined by SDS-PAGE to be >95%.

2.6.2 Purification of Recombinant LpxD

The pETMHL plasmid (Addgene) containing the N-terminal His-tagged *P. aeruginosa* LpxD with the 2 amino acid deletion sequence was transformed into cells, which were then grown in a 50 ml overnight culture of LB media with 35 µg/ml of chloramphenicol and 50 µg/ml of kanamycin at 20°C. Then 10 ml of the overnight culture was added to 1 L of LB media containing 1 ml of 35 µg/ml of chloramphenicol and 50 µg/ml of kanamycin each. These cells were incubated at 37°C for 4 hours until they reached an OD$_{600}$ of 0.6 to 0.8. Then induction was carried out by adding 1 ml of 0.5 M IPTG and further incubating at 20°C again overnight. The culture is then centrifuged at 5,000 g for 10 minutes and the pellet is resuspended in 10 ml of the lysis buffer (20 mM Tris-HCl pH 8.4, 300 mM NaCl, 10% glycerol v/v, and 20 mM imidazole). The cells are thawed on ice and transferred to a 40 ml beaker. The cells are then sonicated on a 10 second sonication/15 second rest cycle for a total of 15 minutes at an amplitude of 6. The lysate is then centrifuged at 40,000 rpm for 40 minutes at 4°C. The supernatant is then filtered and loaded onto a HisTrap affinity column and eluted with a linear concentration gradient spanning 10-500 mM imidazole. The fractions containing LpxD were then collected, pooled, and concentrated down using an Amicon filter. The concentration of the protein was checked using
The protein was loaded to a HiLoad 16/60 Superdex 75 column for additional purification along with the storage buffer (20 mM Tris-HCl pH 8.6 and 250 mM NaCl). The protein eluted at a peak consistent with the size of the trimeric form. The LpxD was stored at -80°C at 35.8 mg/ml. The purity of the protein was determined by SDS-PAGE to be >95%.

**2.6.3 Purification of Biotinylated Avidity-Avi Tagged LpxA and LpxD**

The LpxA and LpxD genes were inserted into BamHI and HindIII site of pAvibir plasmid (Achaogen). The plasmids were transformed into Rosetta DE3 pLysS cells, which were then grown in an overnight culture of 50 ml LB media with 100 µg/ml ampicillin and 35 µg/ml chloramphenicol at 20°C. Then 10 ml of the overnight culture was added to 1 L of LB media again with 100 µg/ml ampicillin and 35 µg/ml chloramphenicol and incubated at 37°C until OD<sub>600</sub> was at least 0.6 but less than 0.8. Protein expression was induced by adding 1 ml of 0.5 M IPTG, along with 10 ml of fresh biotin solution at a concentration of 5 mM. The culture was then incubated at 20°C with vigorous shaking overnight. Cells were harvested by centrifuging the solution for 10 minutes at 4000 g at 4°C. The supernatant was discarded and the cell pellet was then resuspended in 10 ml of the lysis buffer (10 mM Tris pH 8.4, 300 mM NaCl, 20 mM Imidazole, and 10% glycerol v/v) and a single dissolved protease inhibitor tablet. The resuspended cell pellet was then transferred to a 50 ml beaker for sonication. The cells were then sonicated on a 10 second sonication/15 second rest cycle for a total of 15 minutes at an amplitude of 6. The lysate was then transferred to a centrifuge tube and centrifuged at 35,000 rpm for 35 minutes at 4°C. The supernatant was then filtered and loaded onto a HisTrap affinity column and eluted with a linear concentration gradient spanning 10-500 mM imidazole. The fractions containing protein were then collected, pooled, and concentrated down using an
Amicon filter. The concentration of the protein was checked using OD$_{280}$. The protein was the loaded to a HiLoad 16/60 Superdex 75 column for additional purification along with the storage buffer (20 mM Tris pH 8.6, 250 mM NaCl, and 1 mM EDTA). The protein eluted at a peak consistent again with the trimeric form of the protein. The protein was stored at -80°C at 7.9 mg/ml for LpxA and 16.1 mg/ml for LpxD. The purity of the protein was determined by SDS-PAGE to be >95%.

### 2.6.4 LpxA Crystallization

Qiagen crystallization screens JCSG suites I–IV, AmSO$_4$, MPD, and Core I-II, were screened using a Phoenix nanodispenser, and 0.2 and 0.4 μl aliquots of protein solution (14.3 mg/ml), each with 0.2 μl of well solution, were used to search for crystallization conditions. *P. aeruginosa* LpxA readily crystallized under many conditions, producing cuboidal crystals of poor X-ray diffraction quality. Crystals with an almond-like morphology emerged in 20% (w/v) PEG 1000, 0.2 M calcium acetate, and 0.1 M imidazole, pH 8.0, which diffracted to high resolution but inconsistently. The crystallization condition was optimized to 12% (w/v) PEG 1000, 0.2 M calcium acetate, and 0.1 M imidazole, pH 7.0 with a drop ratio of 1 μl of protein to 2 μl of well solution and 0.5 μl of seed stock. The crystals appeared within 2–4 days and measured up to 0.1 mm in length. An additional 750 nl of protein is added to the drop after the crystals have formed to increase the size of the crystal, which requires an addition 2-4 days of growth.

Compound 1 and 2 complex structures with LpxA were obtained by transferring apo crystals into crystallization solution containing 20 mM of each respective compound and 10% DMSO and then they were soaked for 24 hours. After the 24 hour soak the crystals were
transferred to another crystallization buffer containing 25% glycerol and immediately flash frozen in liquid nitrogen.

2.6.5 LpxD Crystallization

Qiagen crystallization screens Core I-IV, JCSG suites I–IV, and AmSO₄ were screened using a Phoenix nanodispenser, and 0.2 and 0.4 μl aliquots of protein solution (29 mg/ml), each with 0.2 μl of well solution, were used to search for crystallization conditions. *P. aeruginosa* LpxD readily crystallized under many conditions, producing crystals of poor X-ray diffraction quality or of poor stability making them unable to be mounted into loops for diffraction analysis. Crystals with a cuboidal-like morphology emerged in 20% (w/v) PEG 3350 and 0.2 M magnesium acetate which diffracted to high resolution consistently. The crystallization condition was optimized to 12% (w/v) PEG 3350 and 0.2 M magnesium acetate with a drop ratio of 2 μl of protein to 3 μl of well solution and 0.5 μl of seed stock. The crystals appeared within 21-28 days and measured up to 0.1 mm in length.

The apo crystals were transferred to another crystallization buffer containing 25% glycerol and immediately flash frozen in liquid nitrogen. Compound 1 complex structure with LpxD was obtained by transferring apo crystals into crystallization solution containing 20 mM compound 1 and 10% DMSO and then they were soaked for 51 hours. After the 51 hour soak the crystals were transferred to another crystallization buffer containing 25% glycerol and immediately flash frozen in liquid nitrogen. The compound 2 complex structure with LpxD was obtained by transferring apo crystals into crystallization solution containing 20 mM compound 2 and 4% DMSO and then they were soaked for 24 hours. After the 24 hour soak the crystals were
transferred to another crystallization buffer containing 25% glycerol and immediately flash frozen in liquid nitrogen.

**2.6.6 Data Collection and Processing**

X-ray diffraction data for the LpxA complex structures with both compounds 1 & 2 were collected at the SER-CAT BM beamline at the Advanced Photon Source (APS) within Argonne National Laboratory (ANL). X-ray diffraction data for the apo crystal and compound 2 complex structure with LpxD were collected at the SER-CAT ID beamline at the same synchrotron. X-ray diffraction data for the compound 1 complex structure with LpxD was collected at the SBC BM beamline which is also at APS within ANL. All data sets were indexed and integrated using iMOSFLM\(^\text{103}\), scaled with Scala, and structures solved using Molrep\(^\text{104}\); all of these programs can be found in the CCP4 suite.\(^\text{105}\) Molrep uses the method of molecular replacement to solve structures so for the LpxD crystals the structures were solved using the previously solved *P. aeruginosa* LpxD structure (PDB ID: 3PMO) as the model. All LpxA crystals structures were solved using the previously solved *P. aeruginosa* LpxA structure (PDB ID: 5DEM) as the model. All model rebuilding of the solved structures was done using Coot\(^\text{106}\), while all the refinements were carried out in the program Refmac5\(^\text{107}\) also found in the CCP4 suite.

**2.6.7 Surface Plasmon Resonance**

All experiments were conducted using a Biacore 4000 instrument with a CM5 chip at 25°C. Avi-LpxA (31.6 kDa, 7.9 mg/ml stock concentration) and Avi-LpxD (39.8 kDa, 16.1 mg/ml stock concentration) were used as ligands to capture onto the Neutravidin immobilized CM5 chip surface. Eighty-two different compounds were used as analytes. Neutravidin (60 kDa,
10 mg/ml) was diluted in 10 mM sodium acetate buffer at pH 4.5 (1:50 dilution, 200 µg/ml
diluted concentration) and immobilized to a level of ~18000 RU, using standard amine coupling
chemistry. This Neutravidin was immobilized onto all spots (Ss) of all flow cells (FCs). PBS-P
(20 mM Phosphate buffer, pH 7.4, 2.7 mM KCl, 137 mM NaCl, 0.05% v/v surfactant P20) was
used as the immobilization running buffer. S3 off all FCs were used as the reference spot. Avi-
LpxA was diluted (1:75 dilution, 105.3 µg/ml diluted concentration) in PBS-P and injected onto
S1 of all FCs (FCs 1-4). Avi-LpxD was diluted (1:75 dilution, 214.7 µg/ml diluted
centration) in PBS-P and injected onto S5 of all FCs. The ligands were captured in the
presence of PBS-P. Based on the captured response values, theoretical R_max values were
calculated for the lowest and highest MW analytes. The R_max values assum 1:1 interaction
mechanism. Overnight kinetics were performed for all compounds in the presence of PBS-P+1%
DMSO buffer. Contact time and dissociation time used in screening experiments were 60
seconds and 300 seconds respectively. Injected analyte concentrations were 0 µM, 2.5 µM, 5
µM, 10 µM, 20 µM, and 40 µM. The compounds were injected in triplicate for each
concentration. Data from overnight kinetics were evaluated by steady state affinity or 1:1
kinetics models fitting.
Table 2.1. *Surface Plasmon Resonance Assay Hits.* The eight compounds we were able to confirm bound to LpxD through using an SPR assay.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
<th>LpxA (µM)</th>
<th>$\chi^2$ (LpxA)</th>
<th>LpxD (µM)</th>
<th>$\chi^2$ (LpxD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>19.5</td>
<td>0.2196</td>
<td>36.7</td>
<td>0.2282</td>
</tr>
<tr>
<td>3</td>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>16.7</td>
<td>4.124</td>
<td>NB</td>
<td>NB</td>
</tr>
<tr>
<td>4</td>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>13.6</td>
<td>0.3219</td>
<td>NB</td>
<td>NB</td>
</tr>
<tr>
<td>5</td>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>2.1</td>
<td>0.4179</td>
<td>NB</td>
<td>NB</td>
</tr>
<tr>
<td>6</td>
<td><img src="image6.png" alt="Structure 6" /></td>
<td>64.6</td>
<td>0.2646</td>
<td>21.6</td>
<td>0.1569</td>
</tr>
<tr>
<td>7</td>
<td><img src="image7.png" alt="Structure 7" /></td>
<td>7.6</td>
<td>0.3253</td>
<td>41.1</td>
<td>0.1254</td>
</tr>
</tbody>
</table>

A NB = No Binding; NA = $K_d$ could not be determined via fitting due to poor data quality

B $\chi^2$: The average deviation of the experimental data from the fitted curve, where lower numbers indicate a better fit. Each compound concentration was tested in triplicates and all data were fitted onto one dose-response curve.
### Table 2.2. X-Ray Data Collection and Refinement Statistic

#### Data Collection

<table>
<thead>
<tr>
<th>Structure (PDB ID)</th>
<th>LpxD Apo</th>
<th>LpxD +236</th>
<th>LpxA+236</th>
<th>LpxA+251</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space Group</td>
<td>H3</td>
<td>H3</td>
<td>P2₁2₁₂₁</td>
<td>P2₁2₁₂₁</td>
</tr>
<tr>
<td>Cell Dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a, b, c$ (Å)</td>
<td>104.78</td>
<td>105.07</td>
<td>80.28</td>
<td>80.16</td>
</tr>
<tr>
<td></td>
<td>104.78</td>
<td>105.07</td>
<td>82.5</td>
<td>82.52</td>
</tr>
<tr>
<td></td>
<td>94.12</td>
<td>96.54</td>
<td>221.9</td>
<td>223.45</td>
</tr>
<tr>
<td>$a, b, g$ (°)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>120</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>52.39-1.52</td>
<td>52.51-2.72</td>
<td>77.33-2.00</td>
<td>57.50-2.00</td>
</tr>
<tr>
<td>No. Reflections</td>
<td>59298</td>
<td>10690</td>
<td>96184</td>
<td>85168</td>
</tr>
<tr>
<td>$R_{merge}$ (%)</td>
<td>9.8</td>
<td>20</td>
<td>12.5</td>
<td>11.8</td>
</tr>
<tr>
<td>$I / \langle I \rangle$</td>
<td>13.6 (2.0)*</td>
<td>5.1 (2.1)*</td>
<td>8.0 (2.0)*</td>
<td>10.9 (2.3)*</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100.0</td>
<td>100.0</td>
<td>96.2</td>
<td>84.6</td>
</tr>
<tr>
<td>Redundancy</td>
<td>11.4 (11.0)</td>
<td>5.5 (5.6)</td>
<td>4.4 (4.4)</td>
<td>5.2 (5.5)</td>
</tr>
</tbody>
</table>

#### Refinement

| Resolution (Å)     | 52.39-1.52| 52.51-2.72| 77.33-2.00| 57.50-2.00|
| $R_{work}/R_{free}$ (%) | 17.86/19.3 | 20.79/29.36 | 20.43/24.49 | 19.63/24.54 |
| No. Heavy Atoms    |           |           |          |          |
| Protein            | 2606      | 2458      | 11942    | 11886    |
| Ligand/Ion         | 1         | 18        | 108      | 137      |
| Water              | 162       | 22        | 159      | 116      |
| $B$-Factors (Å²)   |           |           |          |          |
| Protein            | 19.99     | 49.53     | 28.97    | 25.58    |
| Ligand/Ion         | 26.6      | 45.5      | 29.31    | 21.55    |
| Water              | 29.45     | 24.34     | 25       | 16.66    |
| Ramachandran Plot  |           |           |          |          |
| Most Favored Region (%) | 98.8     | 90.5      | 95.3     | 94.8     |
| Additionally Allowed (%) | 1.2      | 7.1       | 4.6      | 4.8      |
| Generously Allowed (%) | 0.0      | 2.4       | 0.1      | 0.3      |

*Values in parentheses represent highest resolution
**Figure 2.1. Lipid A Biosynthesis Pathway.** In *Pseudomonas aeruginosa* LpxA catalyzes the first step in the lipid A biosynthetic pathway (Raetz pathway) by mediating the reversible transfer of β-hydroxydecanoate from the Acyl Carrier Protein (ACP) onto the 3-OH position of UDP-GlcNAc and forming an ester bond. LpxD catalyzes the third step in the Raetz pathway, the reversible transfer of a β-hydroxydodecanoate from ACP onto the 2-NH$_2$ of the UDP-3-O-(3-hydroxydecanoyl) glucosamine through the formation of an amide bond.
Figure 2.2. Magnesium Ion and Coordinated Waters at the Core of LpxD Trimer. The 2Fo-Fc electron density map is shown at 2.0 $\sigma$. (A) Overhead view of the LpxD trimer (green) with well defined electron densities of a magnesium ion (green sphere) coordinating with six water molecules (red spheres) at the core of the trimer. (B) Side view of LpxD with the C-terminal helix bundle shown at the top.
Figure 2.2. Magnesium Ion and Coordinated Waters at the Core of LpxD Trimer.
(Continued) The 2Fo-Fc electron density map is shown at 2.0 $\sigma$. (C) Zoomed in side view of LpxD and the magnesium ion. Hydrogen bonds are indicated by black dashed lines.
Figure 2.3. LpxA Complex Structures with Novel Inhibitors. Compounds 1 (orange) and 2 (purple) both bind to LpxA (green). Hydrogen-bonds are shown as dashed lines. The unbiased Fo – Fc map (contoured at 2.0 σ) allows for the identification of the binding poses of compound 1 in the active site. Two unique binding poses are observed for compound 1 in different LpxA active sites of the same trimer. (A) Pose 1 of compound 1 has the carboxyl tail forms two hydrogen bonds with Gln157 main chain and His156 side chain. (B) Pose 2 of compound 1 hydrogen bonds with a new conformation of Gln157 side chain. (C) The binding pose of compound 2.
Figure 2.4. LpxD Complex Structure with Compound 1. The compound is colored in orange. HBs are shown as dashed lines. (A) Superimposition of the apo (purple) and complex (green) protein structures shows a shift of the C-terminal alpha helix region of the protein in response to ligand binding. (B) The unbiased Fo – Fc map (contoured at 2.0 σ) identifies the binding pose of the compound 1 in the active site.
Figure 2.5. Structural Alignments of P. aeruginosa LpxA & LpxD. (A) Superimposition of the overall LpxA (green) and LpxD (magenta) trimers shows high levels of conservation for the backbone structure. (B) Similarities in the binding hot spots of LpxA (green) and LpxD (magenta). The acyl-chain of the LpxA-product complex (yellow) is shown to show the active site. His121 and His242 are the catalytic histidine in LpxA and LpxD respectively. (C) Compound 1 exhibits a comparable binding pose in LpxA (green) and LpxD (magenta).
Chapter 3:

Analyzing Mechanisms of Action for β-Lactamase Resistant Antibiotics

3.1 Overview

There is a lack of understanding in the literature about the countering effects towards antibiotic resistance displayed by temocillin, why despite temocillin’s broad spectrum activity against other resistant strains it appears to be ineffective in the treatment of Pseudomonas aeruginosa, and the role it’s 6-α-methoxy group plays in binding to Penicillin Binding Proteins (PBPs). In order to solve these problems, structural analyses were carried out comparing complex structures between PBPs with temocillin and ticarcillin (the parent compound of temocillin), which only varies from temocillin in its lack of a 6-α-methoxy group. Such studies will allow for improved broad spectrum β-lactamase inhibitors and ideally lead to new antibiotic treatments for resistant strains of P. aeruginosa seen in the clinic.

3.2 Introduction

3.2.1 Penicillin Binding Proteins and Traditional β-Lactam Antibiotics

PBPs are bacterial membrane-bound enzymes located in the periplasm of the cell that are responsible for the formation of the peptidoglycan cell wall.108,109 Bacteria synthesize their cell walls through the formation of N-acetylglucosamine-N-acetyl muramic polymers.108,109 After being synthesized, these polymers get polymerized through the linking of the disaccharide peptides (glycosyltransferase activity) and through covalent bond formation at the D-alanine
residues of the muramyl pentapeptides (transpeptidase activity) (Figure 1). This enzymatic activity is carried out by the PBPs in order to form the highly-ordered peptidoglycan cell wall. Every species of bacteria expresses multiple types of PBPs and these enzymes can be classified into two subgroups: high-molecular mass (HMM) PBPs and low-molecular mass (LMM) PBPs. The HMM PBPs can be further divided into two subclasses: class A enzymes which catalyze both the glycosyltransferase activity and the transpeptidase activity and class B enzymes, which are capable of only transpeptidase catalysis. The necessity, for cell survival, of having functional copies of these different types of PBPs varies from species to species. β-lactams are compounds that irreversibly inhibit the catalytic function of these enzymes by mimicking the D-alanine-D-alanine substrate and then forming of a covalent bond between their β-lactam ring and the active site serine of the PBPs. By preventing the PBPs from cross-linking the polymers, they induce disruptions in the cell wall integrity which causes the bacterial cell to lyse and results in cell death.

3.2.2 The Rise of β-Lactamases and Antibiotic Resistance

Since the introduction of antibiotics, bacteria have evolved various mechanisms of resistance to combat these drugs and their ability to kill the cells. The most common mechanism seen in bacteria is the encoding and expression of β-lactamase proteins. These enzymes degrade the activity of β-lactam antibiotics by forming an acyl-enzyme complex with the β-lactam ring. The enzymes then release a hydrolyzed compound, which can no longer interact with the PBPs (Figure 2). Such enzymatic activity has rendered many known antibiotics clinically useless, due to their high susceptibility to degradation by these enzymes. There are four classes of β-lactamase enzymes, separated into classes A, B, C, and D. Classes A, C, and D share a conserved catalytic serine within their active site that is responsible for the hydrolysis
of the β-lactam compounds, whereas class B is made up of the metallo-enzymes. Class B enzymes rely on the presence of a catalytic metal ions to properly function and do not form an acyl-enzyme complex during their mechanism of action. Additionally, there are some β-lactamase proteins encoded by different strains of bacteria that have the ability to hydrolyze a wide range of different β-lactam compounds called extended-spectrum β-lactamases (ESBLs). β-lactamases and the many other mechanisms of antibiotic resistance pose great challenges for researchers and clinicians to overcome in order to prevent this epidemic from becoming a prevalent crisis.

3.2.3 Renewed Clinical Relevancy of Temocillin

Due to the rise in antibiotic resistance in the clinical setting many healthcare providers have turned to older antibiotics, which saw a decline in clinical use prior to the evolution of newer antibiotic resistant strains, in an effort to find treatments that will work against these novel forms of resistance; one of these clinically forgotten antibiotics with new relevancy is temocillin. Temocillin is a broad spectrum carboxypenicillin used in the treatment of gram-negative bacteria that display multidrug resistance in the clinical setting due to carboxypenicillin’s ability to resist hydrolyzation by β-lactamase proteins (Figure 3). Temocillin’s mechanism of action, like other β-lactam antibiotics, is carried out by covalently binding to the active site of PBPs. Temocillin does not, however, show clinically significant activity against *P. aeruginosa*. The only PBP that has shown to be essential for cell growth and viability of *P. aeruginosa* is PBP3, which has made it an important drug target for treating these infections.
3.2.4 Comparative Analysis of Temocillin and Ticarcillin

One of the currently known inhibitors of PBP3 is ticarcillin, considered a parent compound of temocillin, because they share the exact same structure with the only difference being ticarcillin lacking a 6-α-methoxy group (Figure 3).\textsuperscript{120} Ticarcillin is clinically approved for treating \textit{P. aeruginosa} when in conjunction with clavulanic acid, a known β-lactamase inhibitor.\textsuperscript{121,122} The two compounds are administered in combination with one another because ticarcillin is able to be degraded by β-lactamase activity.\textsuperscript{123} Temocillin though is not hydrolyzed by many β-lactamase enzymes commonly seen in the clinic, including AmpC cephalosporinases.\textsuperscript{124,125} This has particular clinical relevancy due to the fact that \textit{P. aeruginosa} encodes the gene for AmpC and can be induced to express this β-lactamase protein.\textsuperscript{125} Due to the difference between temocillin and ticarcillin, in terms of ability to be hydrolyzed by β-lactamases, it can be inferred that this difference must be the result of the presence of the 6-α-methoxy group. This group can be observed, on the β-lactam ring of temocillin, to be interfering with the ability of these enzymes to carry out their catalytic activity (Figure 3).

Interestingly, temocillin has shown to be a weak inhibitor of \textit{P. aeruginosa} growth (MIC = 128-256 mg/L), even though it is resistant to β-lactamase activity and has high structural similarity to ticarcillin, which displays strong activity against this particular species (MIC = 16-32 mg/L).\textsuperscript{118,126} It is not currently understood why temocillin displays weak inhibition of \textit{P. aeruginosa} growth or how temocillin’s molecular mechanism to resist β-lactamase activity works. Research has suggested that MexAB-OPrM efflux pump encoded by \textit{P. aeruginosa} is what allows the bacteria to remain resistant to the bactericidal effects of the compound.\textsuperscript{118} The efficacy of temocillin is also thought to be affected by a weak binding affinity for PBPs when
compared to other β-lactam compounds. In either case, the difference in MIC values between temocillin and ticarcillin must be related to the presence of the 6-α-methoxy group seen on the β-lactam ring of temocillin.

In order to determine the exact role that the 6-α-methoxy group plays in the binding of temocillin to PBPs, comparative analysis with ticarcillin was carried out through the development of complex structures, SPR analysis, and fluorescence polarization assays for *P. aeruginosa* PBP3. Further structural information was determined by solving a complex structure of CTX-M-14 β-lactamase with temocillin in order to compare modes of binding and studying the mechanism by which temocillin prevents serine β-lactamase activity. The combination of the experiments carried out in this research were done in an effort to further the goal of designing better drugs that counteract the mechanisms of antibiotic resistance employed by *P. aeruginosa* seen in the clinic.

### 3.3 Results

#### 3.3.1 Fluorescence Polarization Assays

Due to the essential role of PBP3 in the growth of *P. aeruginosa*, it is essential to look at and compare the binding kinetics of ticarcillin versus temocillin in order to determine why these compounds have such different efficacies in this particular species. In order to resolve the various binding constants, fluorescence polarization assays were carried out using the commercially available fluorescence labelled penicillin called bocillin. This compound binds to PBP3 with strong affinity but also has an attached fluorophore that absorbs light at a wavelength of 490 nm and emits at 520 nm. When bound to PBP, the fluorescence anisotropy of bocillin increases and can be calculated by measuring the polarization properties observed when
exciting the fluorophore on the molecule with plane-polarized light. The binding constants can be determined from the fluorescence anisotropy of the molecule due to the fact that when bocillin is in solution it has a higher rotational mobility and thus has an increased likelihood of becoming depolarized between the times of excitation and emission. The differences in bocillin’s fluorescence anisotropy when bound versus unbound allows for PBP inhibitor binding kinetics to be measured using a plate reader. Both compounds displayed similar high $K_{\text{on}}$ values with temocillin at 77,000 M$^{-1}$s$^{-1}$ and ticarcillin at 21,000 M$^{-1}$s$^{-1}$. Temocillin displayed a 20-fold difference in its $K_{\text{off}}$ (0.001 s$^{-1}$) when compared to ticarcillin ($K_{\text{off}} < 0.00005$ s$^{-1}$). Due to the fact that both temocillin and ticarcillin are covalent inhibitors there are two potential explanations for seeing a significant difference in the $K_{\text{off}}$ values but not in the $K_{\text{on}}$ values. It could be that temocillin is displaying noncovalent inhibitor kinetics due to it binding in a conformation that is catalytically incompatible with the PBP3 active site, thus resulting in the dissociation of the compound from the protein before forming a covalent complex within the active site of PBP3. Another possibility for why a significant difference in the $K_{\text{off}}$ values is observed but not in the $K_{\text{on}}$ values is that PBP3 is displaying low levels $\beta$-lactamase activity against temocillin by forming an acyl-enzyme complex with the compound and then hydrolyzing the molecule to release it from the product complex. Due to an inability to determine which hypothesis is correct further studies were required to determine what the cause of the observed $K_{\text{off}}$ values was.

### 3.3.2 Thermostability Studies of PBP3

Previous research looking at the binding kinetics of temocillin with *E. coli* PBP3, suggested that either temocillin has weak efficiency in forming an acyl-enzyme complex with PBP3 or that it forms an unstable complex with the protein.\textsuperscript{127} The $K_{\text{on}}$ value observed in the
previously discussed fluorescence polarization assay suggests that it is unlikely due to weak efficiency in binding to PBP3 and thus it seemed pertinent to probe the stability of the acyl-enzyme complex of PBP3 with temocillin using circular dichroism in a thermal shift assay. Both compounds, temocillin and ticarcillin, were added in molar excess ratios to PBP3 and incubated with the protein in order to allow for the covalent product complex to form. After incubation the differences in stability of the protein were measured (Figure 4). The melting temperature \( T_m \) for apo PBP3 was observed to be 49.1°C. By incubating PBP3 with temocillin a \( T_m \) of 53.52°C was observed marking a 4.42°C shift in the melting temperature of the protein; however, when incubating the protein with ticarcillin an increase in melting temperature of 7.93°C was observed \( (T_m=57.03°C) \). These results in conjunction with the binding kinetics observed in the fluorescence polarization assay suggest that temocillin is forming a destabilized complex with PBP3 that results in the dissociation of the compound from the active site.

3.3.3 Crystallographic Complex Structure Determination with PBP3

The biochemical and biophysical data from the previous experiments has suggested that due to the destabilizing effects of temocillin binding to PBP3 the compound was dissociating from the active site of the protein, but this does not allow for the determination of a mechanism by which the deacylation of the compound occurs. In order to develop a mechanistic understanding of the deacylation process, x-ray crystallographic studies were performed so that complex structures of these compounds with PBP3 could be developed and used to determine what molecular interactions were involved in this process. These studies produced complex structures of temocillin at 2.1 Å (Figure 5a) and ticarcillin at 1.76 Å (Figure 5b), which provided unambiguous, single conformation densities for each of the compounds. Both compounds bind
with high similarity in terms of their conformations upon acyl-enzyme complex formation and this can be seen upon superimposing the two structures with an RMSD = 0.321. The carboxylate moiety, seen on the thiazolidine ring, interacts in the traditional β-lactam fashion with the Ser485 of conserved PBP K-T/S-G motif. These structures also allowed for the observation of hydrogen bond formation between Ser485 and Thr487 and an electrostatic interaction with Lys484. All of these contribute to the binding of temocillin and ticarcillin by positioning their β-lactam rings within covalent bond formation distance of the catalytic residue of PBP3, Ser294. A continuous density can be distinctly seen between the β-C of both compounds and the Ser294 residue of PBP3. This unambiguously rejects the hypothesis that temocillin’s binding conformation to PBP3 is incompatible for covalent bond formation of the acyl-enzyme complex. In looking further at the superimposition of both complex structures, the highly conserved conformations of the sidechains can be observed. The hydrophobic pocket formed by PBP3 residues Tyr503, Tyr532, and Phe553 is clearly seen interacting with the thiophene rings of both compounds. Both compounds also form salt bridges between their carboxylate groups and the residue Arg489. Additionally, hydrogen bonds can be seen forming between their carboxylate groups and Tyr409, along with another hydrogen bond being formed between the Asn351 of the conserved PBP SXN motif and the amide oxygen of both compounds’ sidechains. Temocillin forms an extra hydrogen bond with the Asn351 residue by utilizing the oxygen on its 6-α-methoxy group.

3.3.4 Structure Determination of Temocillin Complexed with CTX-M

Compounds, such as cefoxitin, utilize structures of previously identified β-lactam antibiotics and have functional groups added to the carbons at positions 6 and 7 of the β-lactam ring, in order to prevent the hydrolysis of the β-lactam ring by the β-lactamase enzymes. These
functional group additions serve the purpose of displacing or obstructing the catalytic water necessary for the deacylation process that is necessary for the proper functioning of serine β-lactamase enzymes. This allows the antibiotics themselves to act as effective β-lactamase inhibitors. In order to confirm this functional role of the 6-α-methoxy group on temocillin, further structural studies were carried out by solving a complex crystal structure with CTX-M-14 wild type, which was able to be determined at a 1.3 Å resolution (Figure 6A).

PBP5s and serine β-lactamases share a high degree of sequence homology, most likely due to an evolutionary linkage between the two enzymes. This is further evidenced by the conserved active site motifs seen in both such as the previously mentioned substrate recognition motifs SXN and K-T/S-G, along with the catalytic sequence motif SXXXK which is required for the formation of the acyl-enzyme complex. Where these two enzyme families begin to delineate is the presence of the Glu166 in the serine β-lactamases. The Glu166 residue is necessary as it acts as the catalytic base by activating the catalytic water that hydrolyzes the β-lactam ring, thus severing the acyl-enzyme complex and freeing up the active site of the β-lactamase enzyme to degrade additional β-lactam compounds (Figure 2). Given these high similarities in the chemical and structural composition of these active sites, the binding conformation of temocillin when in complex with CTX-M-14 matching that of the one seen in PBP3 is of no surprise. The carboxylate group attached to the thiazolidine ring forms ideal electrostatic interactions with Lys234 and Arg276, along with hydrogen bonds to a concerted water and Thr235. The two binding conformations differ in the positioning of the temocillin side chain, as in CTX-M-14 this sidechain forms hydrogen bonds with Asn104 and Asn132 which pulls it to the opposite side of the active site. The most interesting interaction that was observed is the oxygen atom of the 6-α-methoxy group forming a hydrogen bond with Lys73. The lysine residue acts as the general base
in the activation the catalytic Ser70 by abstracting a hydrogen from the Ser130 sidechain allowing the protein to form the acyl-enzyme complex with β-lactam rings (Figure 6B). In the complex structure, the catalytic water can be observed to have an unobstructed path to the β-C that forms the covalent bond with Ser70. This suggests that unlike other C6 and C7 altered β-lactams, such as cefoxitin, temocillin’s 6-α-methoxy group prevents hydrolytic degradation of the compound by preventing the proton transfer carried out by Lys73, rather than through the steric hindrance of the catalytic water molecule.

3.4 Discussion

3.4.1 The Role of the 6-α-Methoxy Group in β-Lactam Interactions

The results from the fluorescence polarization assay that showed temocillin having a measurable $K_{\text{off}}$ rate seemed to point to PBP3 having β-lactamase activity against temocillin given the previous research on the compound found in the literature. PBPs have been shown in previous studies to have small amounts of β-lactamase catalytic capabilities.\textsuperscript{129,130} Additionally, PBP3 has proven to be able to degrade some β-lactam compounds, as evidenced by the hydrolyzed products of penicillin, azlocillin, and cefoperazone seen in complex with \textit{P. aeruginosa} PBP3.\textsuperscript{131,132} It is counter intuitive, however, to think that the compound with proven evasion of hydrolytic activity in β-lactamases, due to its 6-α-methoxy group, would display an increased $K_{\text{off}}$ rate when compared to its parent compound, which lacks this important functional group and has been shown to be susceptible to hydrolytic degradation. If this were the case, it would have implied that the hydrolytic water in PBP3’s β-lactamase activity is approaching from a direction that does not get obstructed by the 6-α-methoxy group of the compound like in some β-lactamase active sites.
The results of the thermal shift assays, however, suggested an alternative hypothesis that temocillin is forming a destabilized complex with PBP3 that results in the dissociation of the compound from the active site. The crystallographic studies allowed for an understanding of the mechanism by which the destabilization was occurring. Although the similarities between the two PBP3 complex structures were seen largely throughout both structures, the distinction can be made when looking at the 6-α-methoxy group, which gets buried into the active site and points towards the Asn351 residue of the highly conserved PBP SXN motif. When looking at the ticarcillin complex structure, a hydrogen bond can be seen forming between ticarcillin’s side chain amide oxygen and the side chain nitrogen of Asn351. In comparison the temocillin complex structure has the same Asn351 side chain nitrogen forming a bifurcated hydrogen bond with both the side chain amide oxygen of temocillin and the oxygen of the 6-α-methoxy group. The 6-α-methoxy group, however, is a poor hydrogen bond acceptor when compared to the amide oxygen. Additionally, this interaction moves temocillin’s amide oxygen 0.4 Å further away from Asn351 residue (hydrogen bond distance of 3.2 Å) when compared to the hydrogen bond distance seen between ticarcillin and the Asn351 side chain nitrogen which is 2.8 Å in length (Figure 7). In competing with the strong amide oxygen hydrogen bond, seen forming with Asn351, by introducing a significantly weaker hydrogen bond acceptor and pushing the stronger hydrogen bond acceptor further away from the nitrogen donor, the 6-α-methoxy group destabilizes the PBP3-temocillin complex. There was no density resolved for a coordinated water molecule in the active site of either PBP3 complex structure, which suggests that it is a non-coordinated water carrying out the catalytic activity necessary to break apart the acyl-enzyme complex. This supports the hypothesis that rather than the 6-α-methoxy group acting to protect the ligand as thought to be seen in serine β-lacatamase enzymes, the 6-α-methoxy group is
simply destabilizing the active site which reduces the entropic requirements for the hydrolysis of
the product complex and allows for non-coordinated waters to be effective in their ability to
carry out this reaction. This notion is also supported by the previous study done by Labia et. al
where temocillin inhibition was restored in a radiolabeled penicillin assay by reducing the
incubation time and temperature at which the assay was carried out, thus showing the instability
of the PBP3-temocillin complex.\textsuperscript{127}

The predicted role of the 6-\(\alpha\)-methoxy group in preventing the hydrolysis of temocillin in
\(\beta\)-lactamases has been understood to be due to a steric hindering of the coordinated water
molecule, which would prevent the enzyme from degrading the acyl-enzyme complex. However,
the complex crystal structure of temocillin with CTX-M-14 wild type suggests an alternative
mechanism for temocillin’s evasion of \(\beta\)-lactamase activity. Given the hydrogen bond formation
between the 6-\(\alpha\)-methoxy group and Lys73 and the unobstructed path of the catalytic water to the
\(\beta\)-C of temocillin, it would suggest that it is actually carrying out this function by preventing
Lys73 from acting as a general base, which would prevent the enzyme from hydrolyzing the
covalent bond it forms with the \(\beta\)-lactam compound and thus irreversibly inhibiting the enzymes
ability to catalyze further reactions. Additional studies are necessary to support this hypothesis. It
would be beneficial to obtain an ultra-high resolution CTX-M-14 structure in complex with
temocillin in which the hydrogen atoms could be resolved. This would allow for the
determination of whether or not Lys73 was in fact prevented from acting as a general base in the
catalytic mechanism.
3.5 Conclusions

Due to the rise of antibiotic resistance and the contribution of *P. aeruginosa* in this emerging health crisis, renewed relevancy for the once ignored antibacterial pharmaceuticals has begun to take hold in the clinical setting in an effort to combat bacterial infections that are displaying resistance to the current antibiotic regimens. Given its resistance to many β-lactamase enzymes including ESBLs and AmpCs, both of which are produced by clinically observed strains of *P. aeruginosa*, temocillin has once again become a desirable therapeutic for use in hospitals. However, unlike temocillin’s parents compound ticarcillin, temocillin has shown to have weak inhibitory properties when it comes to *P. aeruginosa* infections. There has been a relative lack of exploration in the literature with regards to understanding why temocillin lacks clinically useful inhibition of *P. aeruginosa* infections and has up until this point been thought to be related to the expression efflux pumps and poor PBP binding properties. By gathering biochemical, biophysical, and structural data support has been gathered for the hypothesis that although the 6-α-methoxy group successfully prevents hydrolysis of the β-lactam ring when binding to many β-lactamases, it also acts in reducing the stability of the PBP3 product complex, which causes a reduction in the entropic requirements for a non-coordinated water to carry out a hydrolysis reaction and results in β-lactamase activity being observed when binding to PBP3. This resulted in the observed $K_{\text{off}}$ rate and lower thermal stability of the protein seen in the in vitro assays. Although there is an exponential increase in the number of potentially confounding variables when looking at in vivo data, these results may explain the why there is observed weakness in temocillin’s activity against *P. aeruginosa*, given the critical role PBP3 plays in the growth of the organism. While these findings may show a limitation of this particular β-lactams usefulness in the clinic, other β-lactam compounds that substitute the 6-α-methoxy side chain
with other functional groups that form more favorable interactions with the highly conserved Asn351 residue of PBP3 may allow for the preservation of the compound’s ability to resist serine β-lactamase activity, while also making it effective in the treatment of *P. aeruginosa* infections.

### 3.6 Experimental Procedures

#### 3.6.1 Cloning, Expression, and Purification of CTX-M-14 and PBP3

CTX-M-14 WT was purified as previously described. The recombinant *P. aeruginosa* PBP3 construct (residues 50-579) was cloned into a pET15MHL vector (Addgene). The cloned plasmid was then transformed into Rosetta (DE3) PlysS cells and cultured in 2XYT media at 37°C until an OD<sub>600</sub> of 0.6-0.8 was reached. Induction of protein expression was done overnight at 20°C using 0.5 mM IPTG. The cell culture is then pelleted down using the centrifuge at 5,000 x g for a total of 10 minutes and resuspended. The resuspended cells are then lysed using a sonicator set to an amplitude of 6 with cycles of 10 seconds active and 15 seconds resting for a total of 15 minutes. The sample is then put through additional centrifugation and the supernatant is collected and loaded onto a HisTrap affinity column and eluted using a linear concentration gradient of imidazole into a fractionation tray. Fractions corresponding to the peak were then pooled and buffer exchanged in order to remove any imidazole and sodium chloride. The sample is then cleaved overnight at 4°C using TEV protease in order to remove the His-tag. TEV was selected as the protease due to the lab having an established protocol for simple purification with high yield and the protease having high specificity and fidelity for the given cleavage sites. The sample is then again loaded onto the HisTrap affinity column in order to remove the His-tag and protease completely from the sample. The flow through is collected, concentrated, and then loaded onto a gel-filtration column (HiLoad 16/60 Superdex 75) for additional purification.
polishing. Fractions corresponding to the expected peak for PBP3 were pooled and concentrated to 6 mg/ml. The purity of the protein was assessed using SDS-PAGE and determined to be >95%.

### 3.6.2 Protein Crystallization

*P. aeruginosa* PBP3 crystals were grown in a crystallization buffer conditioning containing 20% PEG 3350 and 0.2 M Calcium Acetate using the hanging drop vapor diffusion method at 20°C with a drop ratio of 1:1 for protein (6 mg/ml) to crystallization buffer. In order to form complex structures the fully grown PBP3 crystals were transferred to a drop of crystallization buffer containing 2mM temocillin or ticarcillin and soaked in solution for 2 hours. All crystals were transferred to a cryoprotectant buffer containing 27.5% PEG3350, 0.2 M Calcium Acetate, and 15% glycerol prior to flashing freezing with liquid nitrogen.

CTX-M-14 crystals were obtained by setting up drops with a crystallization buffer condition of 1.2 M Potassium Phosphate pH 8.3 and using the hanging drop vapor diffusion method at 20°C. The drop was setup with a ratio of 1:1 for protein (10–15 mg/ml) to crystallization buffer. In order to form complex structures the fully grown CTX-M-14 crystals were transferred to a drop of crystallization buffer containing 2mM temocillin for 2 hours. All samples were soaked in a solution of cryoprotectant containing 1.2 M Potassium Phosphate pH 8.3 and 30% Sucrose, prior to being flash frozen with liquid nitrogen.

### 3.6.3 Structure Determination

X-ray diffraction data for the PBP3-temocillin complex was collected at the Advanced Photon Source (APS) within the Argonne National Laboratories using the 19-BM beamline and
processed using the software suite HKL2000.\textsuperscript{134} Diffraction data for the PBP3-ticarcillin complex was also collected at APS on the 19-BM beamline but was processed using the CCP4 software suite\textsuperscript{105} with iMOSFLM.\textsuperscript{103} Diffraction data for CTX-M-14 complexed with temocillin was collected at APS using the 22-BM beamline and processed with iMOSFLM. All data was further processed using CCP4 with Phaser\textsuperscript{135} and Molrep\textsuperscript{104} for the purpose of molecular replacement with the previously solved structure of apo-PBP3 (PDB ID: 3PBN) and apo CTX-M-14 (PDB ID: 1YLT) as the input models. Refmac\textsuperscript{107} was utilized for the purpose of further refinement of the models and these outputs were visually inspected and refined using Coot.\textsuperscript{106}

3.6.4 Thermal Shift Assay

Changes in secondary structure in response to temperature changes were monitored using circular dichroism with a Jasco J-815 CD spectropolarimeter coupled to a Peltier cell holder. A sample of PBP3 was diluted down to 2 µg/ml in a buffer of 50 mM sodium phosphate pH 7.0. Those samples looking at the effects of temocillin and ticarcillin on protein stability were incubated with 15 µM temocillin or ticarcillin for 20 minutes. Melting curves from 40-65°C were carried out while measuring at 222 nm within the CD spectra. Data from the experiments was analyzed using SigmaPlot (Systat Software, San Jose, CA) with a two-state fitting program.
**Figure 3.1. Diagram of Enzymatic Activity of PBP Enzymes.** The glycosyltransferase activity (GT) of PBP enzymes serves to polymerize the N-acetylglucosamine-N-acetyl muramic polymers, while the transpeptidase activity (TP) covalently links the D-alanine residues of the muramyl pentapeptides to give rise to the highly ordered peptidoglycan network.

**Note to Reader** 3.1

Figure 1 in this chapter was previously published 2 by Sauvage & Terrak in Antibiotics, 2016 Feb 17;5(1). pii: E12. and has been reproduced with permission (see Appendix 1).
**Figure 3.2. Diagram of Serine β-Lactamase Catalytic Mechanism.** The reaction mechanism for the hydrolytic catalysis of β-lactams in serine β-lactamases starts with the ground-state Michaelis complex (I), and then continues with the formation of a high-energy intermediate transition state (II), before formation of the acyl-enzyme complex (III). The process of deacylation then begins when a high-energy intermediate transition state is formed after a catalytic water attacks the acyl-enzyme complex (IV), which then results in the release of the hydrolyzed β-lactam product and a renewal of the β-lactamase active site (V).

**Figure 3.3. Chemical Structures of Temocillin and Ticarcillin.** The only structural difference between the two compounds is the presence of a 6-α-methoxy group (red) on the β-lactam ring of temocillin.
Figure 3.4. Thermal Shift Assay for P. Aeruginosa PBP3 Incubated with Temocillin and Ticarcillin. Apo PBP3 had an observed $T_m$ of 49.1°C. Incubation of PBP3 with ticarcillin ($T_m=57.03$°C) displayed a 3.51°C difference in protein stability when compared to the results seen for incubation of PBP3 with temocillin ($T_m=53.52$°C).
Figure 3.5. Unbiased Fo-Fc map of P. aeruginosa PBP3 Complex Structures. (A) PBP3 (green) in complex with temocillin (orange) at 2.1 Å resolution with hydrogen bonds (red) formed between temocillin and PBP3. (B) PBP3 (cyan) in complex with ticarcillin (magenta) at 1.76 Å resolution with hydrogen bonds (red) formed between ticarcillin and PBP3.

Figure 3.6. CTX-M-14 WT Complex Structure with Temocillin. (A) Unbiased Fo-Fc map of CTX-M-14 β-lactamase (pink) in complex with temocillin (orange) at 1.3 Å resolution. (B) Catalytic core of CTX-M-14 in complex with temocillin. The hydrolytic water (magenta) is normally activated by Glu 166 (cyan), allowing it to carry out a nucleophilic attack on the acyl-enzyme complex. However, the 6-α-methoxy function prevents the formation of the deacylation transition state by interfering with proton transfer between Lys73 and Ser70. It thus prevents the release of the compound from the acyl-enzyme complex and acts as a covalent inhibitor of the β-lactamase enzyme.
Figure 3.7. PBP3 Complexes Structure Alignment. Comparison of ligand binding poses of temocillin (orange) with ticarcillin (magenta). The 6-α-methoxy group disrupts the stronger hydrogen bond formed between temocillin’s amide oxygen and Asn351, increasing the hydrogen bond distance of it by 0.4 Å.
Chapter 4:

Structural Analysis of the CCR7 Signaling Molecules

4.1 Overview

CCL19 and CCL21 are both chemokines that bind to the CCR7 chemokine receptor and thus play important roles in the trafficking of immune cells, as well as metastatic tumor cells, from the bloodstream to the lymphatic system. Both CCL19 and CCL21 have structures that were previously solved by NMR. Expanding on the structural data that is available for these proteins x-ray crystal structures for both CCL19 and CCL21 were determined. Using the chemical shift mapping data that identifies the critical CCR7 receptor binding sites on both of these proteins in conjunction with the protein crystal structures provides the opportunity to carry out structure-based drug design with these proteins as the targets and thus develop anti-metastatic compounds.

4.2 Introduction

4.2.1 Chemokines

Chemokines are a protein family consisting of around 50 small cytokines in humans that are generally 8-17 kDa in size and have varying degrees of sequence homology (some as little as 20%) but all share conserved secondary motifs and a conserved tertiary fold. These proteins are grouped into 4 subfamilies based on the spacing of their conserved cysteine residues (CC,
CXCR, CX3CR, and XC. They play a critical role in the directing of homeostatic and proinflammatory immune responses by binding with high-affinity and activating their cognate seven transmembrane G-protein coupled receptors (GPCRs) that are responsible for inducing cell migration along a gradient of increasing chemokine concentration.

4.2.2 Chemokines and Cancer

Chemokines primarily function as signaling molecules for immune cell homing and migration; hence, they also pose a pivotal role in inflammation, which is now known to be critical to the progression and spread of many different types of cancer. Different types of cancer take advantage of the chemokine receptor signaling axes by inducing the primary tumor cells to express the receptors, in order to cause tumor cell migration. In addition to playing a crucial role in metastasis chemokines also appears to be critical to tumor growth and progression through their recruitment of proinflammatory leukocytes (macrophages, neutrophils, fibroblasts, and mesenchymal stem cells), to the primary tumor site. The leukocytes then modulate the microenvironment for optimal tumor growth. Individual microenvironments reflect the biochemical space created as a result of leukocytes releasing various proteases, growth factors, angiogenic factors, and immunosuppressive cytokines; the combination of signaling molecules then complement tumor cell proliferation, invasion of the extracellular matrix, and eventually angiogenesis. It is because of the chemokine’s crucial involvement in the development and progression of the cancer disease state that chemokines have become of particular interest to researchers.
4.2.2.1 CCR7 Signaling Axis and Disease

The homeostatic chemokine receptor CCR7 plays a role in the trafficking of lymphocytes cells across high endothelial venules.\textsuperscript{152,153} CCR7 is activated by both CCL19 and CCL21, the two cognate chemokines of the receptor; however, each chemokine induces differential effects and these effects also vary across the different cell types that express the CCR7 receptor.\textsuperscript{154} For example, white blood cell cancer studies conducted by Lopez-Giral et al. (2004) and Harris et al. (1994) have shown that the CCR7 signaling axis plays a particularly significant part in the metastasis of lymphomas and leukemias by causing these tumors to spread to the lymph nodes and secondary lymphoid organs.\textsuperscript{155,156} CCL19 appears to play a pivotal role in the metastasis of T-cell leukemias when invading the central nervous system (CNS).\textsuperscript{157,158} In addition to playing a role in the metastasis of certain cancers, CCL19 also induces signaling of the CCR7 receptor when HIV-1 infects resting T-cells, which results in a latent infection due to the prevention of reproduction of the viral genome but increase in the viral DNA being localized to the nucleus and integrated into the host cell genome.\textsuperscript{159,160} Prevention of this latent infection would be of great interest for therapeutics since this capability of the virus is one of the primary mechanisms by which it escapes eradication by the currently available antiretroviral therapies for treatment of HIV-1.\textsuperscript{159-161} Furthermore, it is CCL21 that acts as the primary signaling molecule for the metastatic patterns seen in many different types of cancer (e.g., skin, colon, cervical, and breast cancers).\textsuperscript{162-166} Additionally, some cancers overexpress CCL21 in order to evade host immune responses and create lymphoid-like microenvironments that are ideal for tumor growth and progression.\textsuperscript{167}
4.2.3 Structure and Function Comparison of CCL19 and CCL21

In addition to both playing a role in cancer metastasis, both CCL19 and CCL21 share the canonical chemokine tertiary fold, the same spacing of the conserved cysteine residues, and have a sequence identity of 35% (Figure 4.1). However, CCL21 (M.W. = 14.6 kDa) is significantly larger than CCL19 (M.W. = 10.9 kDa), which is due to the extended C-terminus seen on CCL21. The 40 amino acid C-terminal tail has proven to be essential for CCL21’s ability to bind the glycosaminoglycan heparin sulfate, and thus plays a vital role in CCL21’s functionality as an adhesive motility protein. This role also in part explains the differential responses seen when the two chemokines (CCL19 and CCL21) bind to CCR7, despite having similar binding affinities for the receptor. Additionally, CCL21 is unique in that it has a total of six cysteine residues instead of the typical four, like most other chemokines. Sequence analysis predicts the existence three disulphide bonds for this protein but only two have been established in the literature, due to the difficulty in solving a structure of the wild type form with its highly flexible C-terminal tail. Further explanation for the differential functionality of CCL19 and CCL21 is that they both bind to other receptors. CCL19 binds to the atypical chemokine receptor ACKR4, which allows for stromal cells to scavenge the CCR7 signaling axis during skin inflammation. CCL21 is capable of binding to CXCR3 where it plays a role in regulating Th1 immune responses. The other way by which the body maintains functional differences for these two proteins is through altered expression and localization of the proteins throughout various tissue regions.

4.2.4 Receptor Binding of the Chemokines and Sulfotyrosine Recognition

Chemokines and corresponding receptors form protein-protein interactions which follow a general mechanism of receptor activation. The chemokine is initially recognized by the
extracellular, flexible N-terminal tail domain of the receptor. After recognition, the flexible N-terminus of the chemokine gets inserted into the transmembrane channel of the receptor and binds to this portion of the protein. This second binding event triggers receptor internalization and intracellular downstream signaling cascades such as calcium influx, all eventually resulting in a chemotactic response by the cell expressing these receptors (Figure 4.2). \(^{177-179}\) Studies done by Farzan et al. (2002) and Seibert & Sakmar (2008) have shown that sulfation of the tyrosine residues located on the N-terminus of chemokine receptors are essential for binding to their cognate chemokines by imparting a large amount of the affinity involved in these interactions.\(^ {180,181}\) Additionally, this specific functional significance of tyrosine sulfation has been shown to be conserved across a multitude of chemokines.\(^ {182-184}\) Veldkamp et al. (2008) showed that an equivalent sulfotyrosine-binding pocket appears to be conserved across the chemokine superfamily.\(^ {183}\) Traditionally protein-protein interaction interfaces have been considered poor candidates for small molecule drug discovery due to their large, shallow, and solvent exposed binding sites.\(^ {185,186}\) However, these sulfotyrosine recognition sites have proven to be exploitable for designing inhibitors that block the binding of the chemokine for the receptor due to the high amount of affinity they impart to this interaction.\(^ {187,188}\)

4.3 Results

4.3.1 Solving the Crystal Structure of CCL19

As previously mentioned, crystallographic protein complex structures are invaluable to structure based drug design efforts as they afford the ability to determine what specific molecular interactions are involved in the binding event of small molecules to their target protein. Additionally, these structures tend to have higher resolution than other structural techniques and
thus provide more accurate results when carrying out molecular docking experiments. Thus crystal structures for CCL19 and CCL21 were solved in order to obtain this information.

CCL19 was crystallized with a P4$_3$ space group and a unit cell length of a=87.31 Å, b=87.31 Å, c=143.49 Å, and unit cell angles of $\alpha=90^\circ$, $\beta=90^\circ$, and $\gamma=90^\circ$ (Table 4.1). The x-ray crystal structure was resolved at a resolution of 2.5 Å. The asymmetric unit is made up of 18 protein monomers but the observed symmetry does not appear to be biologically relevant as none of the orientations taken on by the protein monomers to form dimer interfaces match with the previously solved CC chemokine dimers.\textsuperscript{189,190} There is no significant variation in the conformations of the different monomers which can be observed upon superimposition of all the monomers onto one another. Upon superimposing all the monomers onto monomer N, the monomer with the lowest B-factor (25.53 Å$^2$), the average RMSD is 0.4 Å, aligning an average of 383 atoms. The low average RMSD value confirms the visual analysis of the superimposition. The only observed conformational variation between the monomers lies in the flexible N-terminus of the protein. The secondary and tertiary structure observed in all the monomers of the crystal structure correspond to the previously solved NMR structure of CCL19.\textsuperscript{137} Conserved cysteines (residues 8 and 9) form disulfide bonds with the conserved Cys34 and Cys50. The N-terminal tail is followed by the N-loop (Cys9-Val17) of the protein. The N-loop connects to the $\beta_1$ strand (Arg18-Lys23), which is joined to the $\beta_2$ strand (Ala33-Thr37) via the 30’s loop (Lys24-Pro32). In between the $\beta_2$ strand and the $\beta_3$ strand (Gly41-Ala46) lies the small 40’s loop (Thr38-Gly41). The third $\beta$ strand is connected to the C-terminal alpha helix (Pro51-Thr63) with the protein ending in the truncated C-terminus.
4.3.2 Solving the Crystal Structure of CCL21

CCL21 was crystallized with a P2₁ space group and a unit cell length of a=65.75 Å, 
b=58.24 Å, c=66.0/5 Å and unit cell angles of α=90°, β=119.94°, and γ=90° (Table 4.1). The x-ray crystal structure was resolved at a resolution of 1.9 Å, which yielded novel structural information not afforded by the previously determined NMR structure. In addition, this further enhances the literature with regards to the critical molecular interactions that play a role in the structure and function of CCL21. The asymmetric unit is made up of six monomers and has a 6-fold pseudo-symmetry (Figure 4.3A). This symmetry does not appear to be biologically relevant, like in CCL19, as none of the dimer interfaces formed between the various monomers corresponds to previously determined dimeric CC chemokine structures; it has been previously established by past research that CCL21 is monomeric. The average B-factor of each monomer ranges from 22.6 to 24.3. When superimposing all of the monomers onto monomer B, the monomer with the lowest B-factor (22.6 Å²), the observed average RMSD is 0.2 Å, aligning an average of 417 atoms (Figure 4.3B). This shows there is little variation between the individual monomers within the asymmetric unit and this can be seen upon visual inspection as well of the superimposition. The primary source of conformational variation is observed in the 30’s loop, composed of residues Glu29 through Pro37, with some mild conformational variation in the 40’s loop (residues Arg44-Glu50) and N-loop (Cys9-Val22) as well. As expected there is also a tiny amount of variation in the truncated but still flexible C-terminus, as well as in the N-terminus. The secondary and tertiary structure of each monomer matches the NMR structure previously determined in the literature with two adjacent cysteines (Cys8 and Cys9). Together they form disulphide bonds with the other conserved cysteines (Cys34 and Cys52). The disulfide bonds are followed by the N-loop, then the three β-strands (β1 [residues Val22-Gln28], β2 [residues Ala38-
Pro43], β3 [residues Leu51-Ala53]), ultimately forming an anti-parallel β-sheet. In between the β1 and β2 strand is the 30’s loop made up of residues Glu29-Pro37; in between the β2 and β3 strand is the 40’s loop. The third β-strand is then followed by a ten amino acid long α-helix (residues Glu57 through Leu67) and the protein ends with the truncated but still flexible C-terminal tail (Figure 4.3C).

4.3.3 Determining the Sulfotyrosine Recognition Sites on CCL21

It was first sought to confirm that, like other chemokine receptors, sulfation of the N-terminal tyrosine residues imparted CCR7 with an increased binding affinity for CCL21. Upon confirming this the next step was the identification of the potential binding sites that could be used to carry out structure-based drug discovery, so it needed to be established where the CCR7 sulfotyrosine recognition hotspots were located on CCL21. Previous research has been published showing that a unsulfated N-terminal fragment peptide of the CCR7 receptor (residues 1-30) bound to full length CCL21 using a 2D $^1$H-$^{15}$N HSQC assay which allowed for the analysis of chemical shift perturbations in the $^1$H-$^{15}$N spectrum of CCL21. By mapping the chemical shifts onto the structure of CCL21, the residues that exhibited specific binding with the titrated molecule were determined. This assay showed that the receptor binds to the N-loop, the 40’s loop, and the β3-strand of CCL21 with a $K_d$ value of 150 ± 30 µM (Figure 4.4B). The 2D $^1$H-$^{15}$N HSQC assay was repeated with a similar CCR7 synthetic fragment peptide (residues 5-30), but this fragment had sulfated tyrosine residues at positions 8 and 17; this resulted in over a 2000-fold increase in the estimated binding affinity, with the $K_d$ value being 72 nM (Figure 4.4B). The 2D $^1$H-$^{15}$N HSQC assay established that similar to other chemokines and their cognate receptors, the affinity of CCL21 for CCR7 is greatly affected by the sulfation of the N-
terminal tyrosines. However, in order to determine specifically where these high affinity hotspots were on CCL21 it was necessary to carry out the same 2D $^1$H-$^{15}$N HSQC assay again.

The subsequent titration was carried out with small CCR7 fragment peptides containing the two separate sulfotyrosine residues (sY8 and sY17) found on the N-terminus of the receptor. When titrating with the sY17 containing fragment (CCR7 residues 11-30) the significant chemical shifts outlined a binding pocket which was composed of the N-loop, 40’s loop, and β3-strand, giving an estimated binding affinity with a $K_d$ of 480 ± 80 µM (Figure 4.4). These residues corresponded to what was found in the previously published data with the single unsulfated CCR7 N-terminal fragment peptide and the conserved sulfotyrosine binding pocket found on other chemokines. However, when titrating with the sY8 containing fragment (CCR7 residues 5-11) significant chemical shifts outlined a novel binding pocket between the N-loop and the α-helix of CCL21. This resulted in a much higher estimated binding affinity ($K_d = 150 ± 40$ µM) when compared to the sY17 binding pocket (Figure 4.4).

Upon having established that sulfation was indeed critical for the binding affinity of CCL21 to CCR7, the next logical step was to probe if sulfation potentially had any effect on specificity for the receptor as well. This was accomplished by carrying out the same chemical shift perturbation assay with the same CCR7 N-terminal peptide fragments but lacking sulfation on the tyrosine residues 8 and 17. The results of this showed that the same residues are perturbed upon titration with the unsulfated peptides but as predicted a significant decrease in binding affinity estimates are seen for both fragments (Figure 4.4), suggesting that sulfation does not play a role in binding specificity. Comprehensively, the results from the NMR perturbation assays support the current literature on the existence of conserved sulfotyrosine recognition sites across the chemokine superfamily, along with the understanding that sulfation of the CCR7 N-terminal
tyrosine residues contributes heavily to the binding affinity, but not the binding site specificity, seen for CCL21 towards its cognate receptor.

4.3.4 Structure Based Drug Design Against CCL21

Using the identification of two discreet sulfotyrosine recognition sites on CCL21, molecular docking experiments were carried out in order to screen for compounds that could potentially bind to the chemokine. The first target was the sY8 site given its significantly higher affinity contribution towards the binding free energy of the CCR7 receptor. Multiple iterations of docking were carried out against this site using DOCK 3.5.4 and the ZINC database as the compound library. Initially the conformational space being sampled in the docking experiments was determined by selecting target residues. These target residues were previously shown in the NMR perturbation studies to interact with the CCR7 fragments and thus used to generate an outline of the pocket of interest. After the initial round of docking the predicted ligand poses were visually inspected, and the top compounds were selected and then ordered for testing to determine if they would actually bind to CCL21. The top three compounds predicted to be the strongest binders based on their binding pose from this initial round were then used in additional rounds of docking in order to determine the conformational space being sampled and bias the results towards these scaffolds with the desired chemotypes that took advantage of the chemical space in the binding site. The docking results were visually analyzed once again, and the top compounds were selected and ordered for testing.

NMR 2D $^1$H-$^{15}$N HSQC perturbation assays were carried out using radiolabeled CCL21 and titrating the sample along with dissolved compounds that were ordered from the docking experiments. Five compounds were found to bind to CCL21, all with affinities in the mM range (Table 4.2). The chemical shifts were mapped onto a model of CCL21 to determine where the
compounds were binding and confirm the interactions were localized to the target binding sites on the chemokine. All the compounds that were found to bind to CCL21 in the NMR titration assays showed similar chemical shift data, and all with residues that were found to be within the predicted sY8 binding pocket. The compound exhibiting the highest affinity for CCL21 was compound BV15-017, which showed chemical shift perturbations with residues Lys16, Ala19, Val21, and Leu63; these four residues are located on the N-loop or α-helix portion of CCL21. Compounds BV15-019 and BV15-020 showed chemical shift perturbations with residues Lys16 and Val21; both residues are found exclusively along the N-loop of CCL21. Compound BV15-012 was found to interact with Lys16 and Val21, as well as with Thr70 which is found on the C-terminal tail of CCL21. More specifically, Thr70 is located at the very end of the α-helix portion of the protein which makes up the far end of the sY8 pocket. Similarly, the compound BV15-015 was found to interact with Thr70, but also Ile17 a residue also located the N-loop of CCL21.

This data suggested that the BV15-012 and BV15-015 compounds when binding to CCL21, spanned the entirety of the predicted sY8 binding pocket, but also showed the weakest estimated binding affinities for the protein (Table 4.2).

4.4 Discussion

4.4.1 Significance of the C-terminal Tail for CCL21

The literature has shown that the C-terminus of CCL21 is critical to the chemokine’s ability to interact with the glycosaminoglycans (GAGs) found in the extracellular matrix, as truncation of this portion of the protein has resulted in a reduction for its affinity towards GAGs. It has also been shown that the cleavage of the C-terminus occurs \textit{in vivo} via protease catalysis in order to establish a soluble CCL21 chemokine gradient, as opposed to a stationary one. This suggests that a C-terminal truncation to the chemokine is biologically
relevant. Further evidence for the biological relevancy of this truncation has been seen in the research that suggests the C-terminal portion of CCL21 can perform an auto-inhibitory function for the chemokine.\textsuperscript{197} It has been shown that post-translational modification of the CCR7 receptor via the addition of polysialic acid (PSA) in dendritic cells greatly enhances the binding affinity of the receptor for the C-terminal portion of CCL21 and thus increases CCR7 activation and the resultant chemotactic response in the dendritic cells.\textsuperscript{198-200} This evidence suggests that the GAGs that are imparting this inhibitory action when bound to CCL21, while PSA is simply serving to interrupt this interaction between the chemokine and the extracellular matrix.\textsuperscript{195} However, the increase in binding affinity seen for the PSA-altered CCR7 receptor has also been reported to occur in the absence of GAGs, suggesting that PSA acts to interrupt the auto-inhibitory state of full length CCL21.\textsuperscript{199} Furthermore, PSA cannot be acting strictly as a co-activator of CCR7 for CCL21, because C-terminal truncated CCL21 can still activate the CCR7 receptor regardless of whether the PSA modification on the receptor is present or not.\textsuperscript{199} Altogether, the findings from these studies suggest that the C-terminal domain of CCL21 acts to engage an auto-inhibitory mechanism for the chemokine.

Variation in the conformational state of the sY17 recognition site due to the presence of the C-terminal tail can be noted by comparing the CCL21 crystal structure with the previously determined NMR structure of CCL21 (PDB ID: 24LN). In the crystal structure the partially intact C-terminal tail appears to occlude the surface of the protein that has been identified as having a role in recognition of the CCR7 receptor (Figure 4.5). Within the literature there is also an NMR study that determined the residues on CCL21 that interact with the C-terminal tail of the chemokine by mapping the chemical shift perturbations observed when comparing a truncated form of the protein that lacked the C-terminal tail (CCL21 1-79) and a non-truncated construct of
the protein (CCL21 1-111). Additionally, it can be seen in the crystal structure that the C-terminal tail is oriented towards the N-loop of the protein. This orientation suggests that the C-terminal tail may be interacting with the N-loop and neighboring residues; thus, altering their conformational states and influencing CCL21’s ability to interact with the CCR7 receptor (Figure 4.5). The argument could be made that the variation in the N-loop’s conformational state is the result of a crystal packing artifact, but when the crystal structure of CCL21 is compared to the previously solved crystal structure of the related chemokine CCL18 the probability of this being the case diminishes. CCL18 does not have an elongated C-terminus and thus closely resembles the truncated form of CCL21 seen in our crystal structure. Unlike many of the chemokines in the CC subfamily, both CCL18 and CCL21 function as monomers in addition to their display of high similarity in their secondary and tertiary structure (Figure 4.1). However, the CCL18 structure reveals completely different interactions at the crystal packing interface compared to those seen in the truncated CCL21 structure. It is unlikely that these two highly similar proteins would adopt such similar structural conformations as the result of profoundly different crystal packing interfaces. In addition to this, when comparing the structure of our truncated CCL21 to the structure of our truncated CCL19, which also lacks an elongated C-terminal tail, there is high similarity in the conformation of the N-loops (Figure 4.1). These structural comparisons lead to the conclusion that the conformation of the C-terminal tail in our crystal structure is likely to have biological relevance to the function of CCL21. It is worth mentioning that our data does not exclude the possibility that the C-terminal tail of CCL21 is altering the function of the chemokine by having direct interactions with CCR7 and thus additional studies are needed to explore this potential mechanism of action.
4.4.2 Significance of the N-terminal Tail for CCL19

The prevailing thought on what the primary functional significance of the flexible N-terminal tail of chemokines is that they are solely responsible for stabilizing the active form of the chemokine receptors through weak interactions with the trans-membrane domain of the receptor.\textsuperscript{201-204} This notion has been supported by many studies carried out looking at the effects of truncating the N-terminus of different chemokines, which all showed that such alterations affected receptor activation but little to no effect on the binding affinity of these proteins for their cognate receptors.\textsuperscript{177,203,205-209} However, uniquely the eight N-terminal residues of CCL19 were shown to not only effect receptor activation but also have a role in the high affinity binding towards the receptor.\textsuperscript{210} Based on the results of these previous truncation studies it would be reasonable to think the N-terminal residues were forming direct interactions with the receptor in order to contribute to the binding affinity of the CCL19 for CCR7, as these truncations directly resulted in significantly lower binding affinities.\textsuperscript{210} However, later research using NMR perturbation assays that titrated receptor fragments with radiolabeled CCL19 did not indicate that these residues played any part in directly interacting with the CCR7 receptor.\textsuperscript{137} These NMR perturbation assays indicated that the N-loop, β-3 strand, and α-helix of CCL19 were the segments of the protein that interacted with the receptor directly.\textsuperscript{137}

When superimposing the crystal structure of CCL19 onto the previously solved NMR structure of the protein the greatest variation in the conformational state of the protein, due to the presence of the N-terminal tail, can be seen in the N-loop region of the protein (Figure 4.6). In the NMR structure the N-terminal tail can be seen bending back and orienting towards the N-loop and interacting with this segment of the protein. Thus, these interactions appear to alter the binding affinity for the receptor by inducing a conformational state in the N-loop region that is
compatible with receptor binding, rather than the N-terminal residues interacting directly with the receptor themselves. The truncated version of the CCL19 protein (residues 7-70) seen in the crystal structure corresponds to the truncated CCL19 seen in the GTP-γS binding and chemotaxis assays that displayed partial agonist activity, thus it seems that this conformational state must be responsible for this observed reduction in binding affinity and chemotactic response. This is a logical conclusion, especially when considering the results from the previous CCL19 truncation binding assays and NMR perturbation studies, and represents yet another potential biological mechanism for increasing the precision of chemotactic responses. Again, it could be argued that this N-loop conformation could be the result of a crystal packing artifact; however, given the previous comparison of the CCL18, CCL19, and CCL21 crystal structures and the three different unique crystal packing interfaces of these proteins, it is unlikely that this conformation would be the result protein crystal packing (Figure 4.1).

4.4.3 Targeting CCL21 with Structure-Based Drug Design

Chemokines and their receptors have been implicated in a host of disease states and thus have become of great interest to researchers for their potential as therapeutic targets. Drug discovery efforts against this signaling axis have remained predominantly focused on targeting the chemokine receptors due to difficulty in targeting protein-protein interfaces with small molecules. Even so, due to the particular mechanism of chemokine-receptor recognition using sulfotyrosines for affinity enhancement, it has been shown that these small proteins can in fact serve as good targets for small molecule inhibitor discovery, as previous efforts targeting CXCL12 have shown. Given the identification of the this conserved mechanism across chemokines and the impact it has on the affinity of the protein for the receptor it was shown to be feasible to go forward with a structure-based drug design efforts against CCL21. Initial efforts in
using the CCL21 crystal structure to carry out molecular docking studies served as a proof of this concept by allowing for identification of five novel ligands that displayed the ability to bind to the target protein with reasonable affinity for initial hit compounds. These compounds were confirmed to be binding at the sY8 recognition target site through the use of 2D NMR $^{15}$N-$^1$H HSQC assays. Further improvements to binding affinity of these initial hit compounds is necessary, but this will require additional structural information through the capturing of complex crystal structures with these compounds. Determining these structures will allow for the rational addition of enhanced functional groups to the current compound scaffolds. Future studies should aim to improve the interaction with the receptor, targeting the sY17 recognition site despite its lower affinity contributions, as identification of hit compounds here could be used to link with compounds identified to bind at the sY8 site. This would offer the chance to create larger compounds with ultra-high affinity and specificity for CCL21.

4.5 Conclusions

In solving the crystal structure of truncated CCL21 and CCL19 alternative N-loop conformations were identified. Both appear to play a crucial role in receptor recognition, but neither of these alternative conformations were observable in the previously published NMR structures.\textsuperscript{136,137} The CCL19 structure helps resolve the question left by the previously carried out CCR7 binding affinity, chemotaxis, and NMR perturbation studies and also suggests a novel mechanism by which the N-terminal tail of CCL19 serves to regulate the signaling of the protein. The CCL21 structure also further elucidated potential roles the C-terminal tail of CCL21 may play in the biological functioning of the protein, along with an explanation of the potential mechanism of action by which it carries out its auto-inhibitory function. Additionally, the identification of two discreet sulfotyrosine recognition sites on CCL21 provided important
structural information that can be utilized in the targeting of the chemokine for small molecule inhibitor discovery efforts. This was proven by successful identification of the five hit compounds that showed reasonable affinity for initial hits; these compounds represent the only currently reported CCL21-specific small molecule inhibitors and serve as an ideal starting point for further structure-based drug discovery efforts. Furthermore, the results support the hypothesis that it is possible to target a multitude of the chemokine signaling proteins with small molecules by identifying these conserved sulfotyrosine recognition sites and hopefully will lead to additional efforts in designing new drugs for the range of diseases this family of proteins has been implicated in.

4.6 Experimental Procedures

4.6.1 Purification of CCL21

The C-terminal truncated sequence of CCL21 (residues 1-79) was inserted into a pQE30 vector (Qiagen) along with a SUMO tag (His<sub>6</sub>-SMT3) and then transformed into competent BL21 (pREP4) E. coli cells. The cells were then cultured in LB media (50 ml) and incubated overnight at 37°C. 10 ml of the culture is then diluted into 1L culture of LB media containing 25 µg/ml of kanamycin and 50 µg/ml of ampicillin and incubated at a temperature of 37°C until an OD<sub>600</sub> of 1.0 is reached. Expression of the vector was induced with 500 µM IPTG and the culture is incubated for another 4 hours. The cell culture is then pelleted down using the centrifuge at 5,000 x g for a total of 40 minutes. The protein is
expressed in inclusion bodies and thus needs to be resuspended in 20 ml of a denaturing buffer composed of 20 mM Tris pH 8.0, 6 M Guanidine HCl, 300 mM NaCl, and 10 mM imidazole and incubated in a hot water bath of 37˚C for 1 hour, while being occasionally shaken. The sample is then centrifuged again at 40,000 x g for 40 minutes and the supernatant is collected and filtered using a 0.2 µm filter in order to prepare it for high-performance liquid chromatography (HPLC). The sample is loaded to the HisTrap Affinity column at a rate of 1 ml/minute and eluted in a single step using a buffer containing 100 mM Sodium Acetate pH 4.5, 6 M Guanidine HCl, 300 mM NaCl. The sample is then dropwise diluted in 200 ml of refolding buffer containing 100 mM Tris pH 8.0, 10 mM Cysteine (reduced), and 0.5 mM Cystine (Oxidized). The sample is then allowed refold overnight at 4˚C. The sample is then filtered, concentrated, and diluted into a cleavage buffer made of 20 mM Tris pH 8.0, 150 mM NaCl, and 10% Glycerol. The SUMO tag is then cleaved off through proteolysis using 2.4 mg of ULP1 and incubated overnight at 30˚C. ULP1 was selected as the protease due to the protease having high specificity and fidelity for the given SUMO tag. The sample is then filtered in order to remove any precipitate formed during the cleavage process. Then the sample is again concentrated and buffer exchanged to 20 mM Tris pH 8.0, 300 mM NaCl and 40 mM Imidazole in order to prepare for additional HPLC runs. The sample is then run through the HisTrap affinity column again at a rate of 1 ml/minute in order to remove the SUMO tag and protease. The sample is collected and then further concentrated and loaded onto a gel filtration column (HiLoad 16/60 Superdex 75 column) and eluted using a buffer containing 20 mM Tris pH 7.0 and 300 mM NaCl. The sodium chloride is critical to getting a protein yield, due to the ability of CCL21 to bind glycosaminoglycans even in its truncated form. The sodium chloride prevents CCL21 from interacting with the generally inert matrix of the gel filtration column. After collecting the sample from the gel filtration column it is
tested for purity using an SDS-PAGE gel. This resulted in a 10.4 mg/ml yield of CCL21 with a purity that was greater than 95%.

4.6.2 Purification of CCL19

The C-terminal truncated sequence of CCL21 (residues 7-70) was inserted into a pQE30 vector (Qiagen) along with a SUMO tag (His$_6$-SMT3) and then transformed into competent BL21 (pREP4) E. coli cells. The cells were then cultured in LB media (50 ml) and incubated overnight. 10 ml of the culture is then diluted into 1L culture of LB media containing 25 µg/ml of kanamycin and 50 µg/ml of ampicillin and incubated at a temperature of 37˚C until an OD$_{600}$ of 1.0 is reached. Expression of the vector is induced with 500 µM IPTG and the culture is incubated for another 4 hours. The cell culture is then pelleted down using the centrifuge at 5,000 x g for a total of 10 minutes and resuspended in 20 ml of 20 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole, 0.1% β-mercaptoethanol, and Pierce protease inhibitor tablet (Thermo Scientific). The resuspended cells are then lysed using a sonicator set to an amplitude of 6 with cycles of 10 seconds active and 15 seconds resting for a total of 15 minutes. The sample is then collected using additional centrifugation at 40,000 x g for a total of 40 minutes. The protein is expressed in inclusion bodies and thus needs to be resuspended in 20 ml of a denaturing buffer composed of 20 mM Tris pH 8.0, 6 M Guanidine HCl, 300 mM NaCl, and 10 mM imidazole and incubated in a hot water bath of 37˚C for 1 hour, while being occasionally shaken. The sample is then centrifuged again at 40,000 x g for 40 minutes and the supernatant is collected and filtered using a 0.2 µm filter in order to prepare it for high-performance liquid chromatography (HPLC). The sample is loaded to the HisTrap Affinity column at a rate of 1 ml/minute and eluted in a single step using a buffer containing 100 mM Sodium Acetate pH 4.5, 6 M Guanidine HCl, 300 mM NaCl. The sample is then dropwise diluted in 200 ml of refolding buffer containing 100 mM
Tris pH 8.0, 10 mM Cysteine (reduced), and 0.5 mM Cystine (Oxidized). The sample is then allowed refold overnight at 4°C. The sample is then filtered, concentrated, and diluted into a cleavage buffer made of 20 mM Tris pH 8.0, 150 mM NaCl, and 10% Glycerol. The SUMO tag is then cleaved off through proteolysis using 2.4 mg of ULP1 and incubated overnight at 30°C. ULP1 was selected as the protease due to the protease having high specificity and fidelity for the given SUMO tag. The sample is then filtered in order to remove any precipitate formed during the cleavage process. Then the sample is again concentrated and buffer exchanged to 20 mM Tris pH 8.0, 300 mM NaCl and 40 mM Imidazole in order to prepare for additional HPLC runs. The sample is then run through the HisTrap affinity column again at a rate of 1 ml/minute in order to remove the SUMO tag and protease. The sample is collected and then further concentrated and loaded onto a gel filtration column (HiLoad 16/60 Superdex 75 column) and eluted using a buffer containing 20 mM Tris pH 8.0, 150 mM NaCl, and 10% Glycerol. After collecting the sample from the gel filtration column it is tested for purity using an SDS-PAGE gel. This resulted in a 13.3 mg/ml yield of CCL21 with a purity that was greater than 95%.

4.6.3 Crystallization Screening of CCL21

Taking the now purified CCL21 protein screening began for crystallization conditions using Qiagen JCSG I-IV and AmSO₄ suites, along with Hampton Crystal Screens I and II and placing them on the Crystal Phoenix robot, which was used to dispense 0.2 µl of protein solution at 10 mg/ml, into a sitting drop screening tray containing 0.2 µl of crystallization buffer solution in each well. These trays were stored at 20°C in our crystallization incubator. Crystals appeared in the well containing a buffer condition of 3.0 M ammonium sulfate and 1% MPD. This condition was then scaled up using hanging drop crystallization trays and optimized by altering the crystallization buffer condition to 2.7 M ammonium sulfate and 0.5% MPD. Additionally, the
drop ratio was altered to 2 µl of protein for every 1 µl of crystallization buffer solution. In order to prepare the crystals for x-ray diffraction they are then soaked in a cryo-protectant solution containing 2.7 M ammonium sulfate, 0.5% MPD, and 25% glycerol.

4.6.4 Crystallization Screening of CCL19

Taking the now purified CCL19 protein screening began for crystallization conditions using Qiagen MPD and AmSO₄ suites, Jena Bioscience JCSG Plus and PACT screens, and the Rigaku Berkley and Top96 screens and placing them on the Crystal Phoenix robot, which was used to dispense 0.2 µl of protein solution at 13 mg/ml, into a sitting drop screening tray containing 0.2 µl of crystallization buffer solution in each well. These trays were stored at 20°C in our crystallization incubator. Crystals appeared in the well containing a buffer condition of 20% PEG3350 and 0.2 M sodium nitrate. This condition was then scaled up using hanging drop crystallization trays and optimized by altering the crystallization buffer condition to 20% PEG3350 and 0.2 M sodium nitrate with a final pH of 5.5. Additionally, the drop ratio was altered to 2 µl of protein for every 2 µl of crystallization buffer solution. In order to prepare the crystals for x-ray diffraction they are then soaked in a cryo-protectant solution containing 27.5% PEG3350, 0.2 M sodium nitrate, and 17.5% Glycerol.

4.6.5 X-Ray Data Collection and Processing for CCL21

X-ray diffraction data for CCL21 was collected at Argonne National Laboratory’s Advanced Photon Source (APS) using the SER-CAT 22-BM beamline. The data was processed using HKL2000¹³⁴ and the CCP4 suite¹⁰⁵ in order to solve the model structure and carry out refinements. Phaser Cell Content Analysis¹³⁵ was utilized for determining the number of monomer copies contained within the asymmetric unit of the crystal, of which it predicted six.
Balbes\textsuperscript{214} taking the sequence of the previously solved NMR CCL21 structure (PDB ID: 2L4N) was able to provide an initial structure through the use of molecular replacement. The resulting solution only contained four copies within the asymmetric unit ($R_{\text{work}}=45\%$ and $R_{\text{free}}=48\%$). The model created by Balbes was input into Refmac5\textsuperscript{107} for refinement. Taking the Refmac5 solution Parrot\textsuperscript{215} was then used in order to optimize the density. In order to improve the data further Buccaneer\textsuperscript{216} was used, giving it the inputs of the improved density from Parrot and the refined data from Refmac5. Then using CCP4\’s Chainsaw utility the protein structure was pruned to remove all the flexible side chains found on the surface of the protein. In order to search for the additional copies that were missing from the initial output of Balbes the pruned structure from Chainsaw along with the output from Buccaneer were then input into Molrep.\textsuperscript{104} Molrep was able to find another additional copy to put that into asymmetric unit and this was run in Refmac5 for further refinement. We repeated to use Molrep in order to find the final sixth copy that was predicted by Phaser Cell Content Analysis. After successfully finding the final copy Refmac5 was again run on the output. Using our six monomer model further refinements were carried out in Coot\textsuperscript{106} in order to fill in the missing residues and rebuild the sections of the protein visually. Final refinements were carried out using Refmac5 and Phenix\textsuperscript{217} ($R_{\text{work}}=20.5\%$ and $R_{\text{free}}=24.8\%$).

4.6.6 X-Ray Data Collection and Processing of CCL19

X-ray diffraction data for CCL19 was collected at Argonne National Laboratory\’s Advanced Photon Source (APS) using the SER-CAT 22-BM beamline. The data was processed using iMOSFLM\textsuperscript{103} and Scala from the CCP4 software suite\textsuperscript{105} in order to integrate and scale the data. The output density was then input into Balbes\textsuperscript{215}, along with the protein sequence, in order to generate a model structure for the protein. The resulting solution contained sixteen copies within the asymmetric unit ($R_{\text{work}}=42.9\%$ and $R_{\text{free}}=47.1\%$). Taking the Balbes solution Parrot\textsuperscript{215}
was then used in order to optimize the density. Additional copies were added to the model and fit into the density manually through visual refinement of the data using Coot. The output density and model from Coot was then input into Refmac5 for additional refinement. Using the eighteen monomer model further refinements were carried out in Coot in order to fill in the missing residues and rebuild the sections of the protein visually. Final refinements were carried out using Refmac5 ($R_{\text{work}}=32\%$ and $R_{\text{free}}=24\%$).

4.6.7 NMR 2D $^1\text{H}-^1\text{H}$ HSQC Spectroscopy

All NMR spectroscopic data was collected at the NMR facility within the Medical College of Wisconsin on a Bruker Avance 600 mHz spectrometer equipped with a $^1\text{H}/^13\text{C}/^15\text{N}$ cryoprobe. The identification of the sulfotyrosine binding sites located on CCL21 and determination of the effect sulfonation has on binding affinity was carried out by using 100 µM U-$^{15}\text{N}$ CCL21 dissolved in 25 mM MES (deuterated) pH 5.94, 0.02% NaN$_3$, and 10% D$_2$O in the NMR spectrometer. This was then titrated with slowly increasing concentrations of the fragment CCR7 peptides and monitored using the 2D NMR $^{15}\text{N}-^1\text{H}$ HSQC spectra, as previously described. The following molar ratios of U-$^{15}\text{N}$ CCL21 to N-terminal CCR7 fragment peptides were used: 1:0, 1:0.25, 1:0.5, 1:0.75, 1:1, 1:3, 1:7, and 1:10. The amide $^1\text{H}-^1\text{H}$ chemical shift perturbations ($\Delta\delta$) were calculated using the formula: \[\sqrt{(5\Delta\delta_{\text{H}})^2+(\Delta\delta_{\text{N}})^2}\], where $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{N}}$ are the sum of all the changes in backbone amide $^1\text{H}$ and $^{15}\text{N}$ chemical shifts in ppm, respectively. The concentration dependent changes in chemical shift data once titrated with the assorted CCR7 N-terminal fragment peptides was implemented in the calculation of the dissociation constant ($K_d$) values by using a nonlinear fitting equation which factors in ligand depletion, as previously described.
Note To Reader 4.1

Members of the Volkman lab at the Medical College of Wisconsin performed the NMR described in this chapter.

Note To Reader 4.2

Text and Figures were reprinted (adapted) with permission from Emmanuel W. Smith,5 Eric M. Lewandowski,5 Natasha A. Moussouras,3 Kyle G. Kroec,5 Brian F. Volkman,3 Christopher T. Veldkamp,1 and Yu Chen,5 Biochemistry 2016, Copyright of American Chemical Society 2016

Note To Reader 4.3

Text and Figures were reprinted (adapted) with permission from Andrew J. Phillips,1 Deni Taleski,2 Chad A. Koplinski,3 Anthony E. Getschman,3 Natasha A. Moussouras,4 Amanda M. Richard,1 Francis C. Peterson,3 Michael B. Dwinell,4 Brian F. Volkman,3 Richard J. Payne,2 and Christopher T. Veldkamp,1 International Journal of Molecular Sciences 2017, Copyright by the authors 2017

5 Department of Molecular Medicine, University of South Florida, Tampa, Florida 33612, United States
Table 4.1 X-ray Data Collection and Refinement Statistics.

<table>
<thead>
<tr>
<th></th>
<th>SEKI</th>
<th>CCL19</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structure (PDB ID)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Group</td>
<td>$P_2_1$</td>
<td>$P_4_3$</td>
</tr>
<tr>
<td>Cell Dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$, $b$, $c$ (Å)</td>
<td>65.75</td>
<td>87.31</td>
</tr>
<tr>
<td></td>
<td>58.24</td>
<td>87.31</td>
</tr>
<tr>
<td></td>
<td>66.05</td>
<td>143.49</td>
</tr>
<tr>
<td>$\alpha$, $\beta$, $\gamma$ (°)</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>119.94</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.00-1.90</td>
<td>50.00-2.55</td>
</tr>
<tr>
<td>No. Reflections</td>
<td>34110</td>
<td>33235</td>
</tr>
<tr>
<td>$R_{merge}$ (%)</td>
<td>11.5</td>
<td>9.7</td>
</tr>
<tr>
<td>$I / \sigma(I)$</td>
<td>16.5 (2.2)*</td>
<td>8.7 (3.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Redundancy</td>
<td>7.0 (6.2)</td>
<td>3.8 (3.8)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>29.12-1.90</td>
<td>32.20-2.55</td>
</tr>
<tr>
<td>$R_{work}/R_{free}$ (%)</td>
<td>20.5/24.8</td>
<td>32.2/24.4</td>
</tr>
<tr>
<td>No. Heavy Atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3427</td>
<td>2867</td>
</tr>
<tr>
<td>Ligand/Ion</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>Water</td>
<td>188</td>
<td>143</td>
</tr>
<tr>
<td><strong>B-Factors (Å$^2$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>23.75</td>
<td>32.17</td>
</tr>
<tr>
<td>Ligand/Ion</td>
<td>16.68</td>
<td>21.34</td>
</tr>
<tr>
<td>Water</td>
<td>25.53</td>
<td>30.89</td>
</tr>
<tr>
<td><strong>Ramachandran Plot</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most Favored Region (%)</td>
<td>94.2</td>
<td>86.8</td>
</tr>
<tr>
<td>Additionally Allowed (%)</td>
<td>5.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Generously Allowed (%)</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

* Values in parentheses represent highest resolution shells
Table 4.2 CCL21 NMR Titration Assay Hit Compounds.

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>Structure</th>
<th>M.W.</th>
<th>Binding to CCL21</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV15-012</td>
<td><img src="image1" alt="Structure" /></td>
<td>235.2</td>
<td>+</td>
</tr>
<tr>
<td>BV15-015</td>
<td><img src="image2" alt="Structure" /></td>
<td>331.3</td>
<td>+</td>
</tr>
<tr>
<td>BV15-017</td>
<td><img src="image3" alt="Structure" /></td>
<td>286.3</td>
<td>+++</td>
</tr>
<tr>
<td>BV15-019</td>
<td><img src="image4" alt="Structure" /></td>
<td>192.2</td>
<td>++</td>
</tr>
<tr>
<td>BV15-020</td>
<td><img src="image5" alt="Structure" /></td>
<td>238.2</td>
<td>++</td>
</tr>
</tbody>
</table>

* + = Low mM Range Affinity, ++ = Mid mM Range Affinity, +++ = High mM Range Affinity
Figure 4.1. Superimposition of Truncated CCL21 with Truncated CCL19. The x-ray crystal structure of truncated CCL21 (purple, PDB ID: 5EKI) shows a highly similar N-loop conformation when compared to the structures of the related chemokines CCL19 (green) and CCL18 (orange, PDB ID: 4MHE).
Chemokine Receptor Binding Mechanism. Chemokines follow a general two-site/two-step mechanism of binding to their cognate receptor. (A) The chemokine and the target receptor’s flexible extracellular N-terminal domain enter within a proximal distance of each other. (B) The first step of the mechanism begins where the chemokine is recognized by the N-terminal domain of the receptor, with a large part of the affinity for this binding coming from the receptors sulfated tyrosine residues. (C) The second step of the mechanism begins when the flexible N-terminal domain of the chemokine then gets internalized within the channel formed by the seven transmembrane domains of the chemokine receptor, where the tail binds. This results in the receptor initiating a downstream signaling cascade on the intracellular side of the membrane.
Figure 4.3. Truncated CCL21 (Residues 1-79) Crystal Structure. (A) The asymmetric unit of the CCL21 crystal which contains 6 monomeric copies of the protein. (B) Each monomer from the asymmetric unit superimposed onto chain B (monomer with the lowest average B-factor), showing that the variation between each copy of the monomer within the asymmetric unit is negligible. (C) Monomer B displayed from two sides and all the important secondary structure components labeled.
Figure 4.4. CCL21 Sulfotyrosine Recognition Site NMR Analysis. (A) Chemical shift perturbations in response to titration with the CCR7 fragment peptides plotted on graphs and mapped onto the CCL21 structure, with the statistically significant shifts shown in orange. The transparent orange surfaces on the CCL21 structures outline the binding pockets of the CCR7 fragments. (B) Estimated binding affinities ($K_d$) and free energy contributions ($\Delta G$) for the binding of the corresponding CCR7 peptide fragments to CCL21 based on the NMR titration data.
Figure 4.5. Superimposition of CCL21 Crystal Structure with CCL21 NMR Structure. The x-ray crystal structure of truncated CCL21 (green) displays additional structural information about residues 71-77 when compared to the previously solved NMR structure ensemble of CCL21 (purple, PDB ID: 2L4N). This new information about the conformation of the C-terminal tail seen in the crystal structure shows it blocking a portion of the sulfotyrosine recognition site identified by the NMR titration experiments and points towards the N-loop.
Figure 4.6. Superimposition of CCL19 Crystal Structure with CCL19 NMR Structure. The x-ray crystal structure of truncated CCL19 (green) shows conformational variation in the N-loop when compared to the previously solved NMR structure ensemble of CCL19 due to the presence of the N-terminal tail (purple, PDB ID: 2MP1).
Chapter 5:

Summary

The focus of all three projects was using structure-based drug design in order to develop or improve small molecule inhibitors for the purpose of alleviating the various disease states that rely on the biochemical pathways being targeted. The first project sought to design novel dual activity inhibitors that bound to both the LpxA and LpxD proteins of *Psuedomonas aeruginosa*; where these enzymes are involved in the synthesis of Lipid A and are crucial to cell survival and viability of the bacterium. Through the use of molecular docking, complex crystal structure determination, and SPR assays, compounds that bind to both enzymes were able to be identified and serve as proof of concept for the central hypothesis of this project. Additionally, the crystal structures determined for LpxD displayed a novel structural mechanism for allosteric ligand inhibition that may lead to a greater understanding of the structure-function relationship of the enzyme and potentially more potent inhibitors. Future research for this project should seek to improve the affinity and specificity of the current lead compounds, as well as find additional lead compounds with unique scaffolds and chemotypes when compared to the previously identified small molecule inhibitors. It is also worth looking further into the novel mechanism of allosteric ligand inhibition and identifying how and why this effect is observed. This project would greatly benefit by gathering biochemical data to support the biophysical data observed in the SPR assay, as this would potentially offer further support of the current findings and possibly help lead to improved inhibitors with better binding affinities.
The second project focused on determining why temocillin is ineffective in treating *P. aeruginosa* infections and elucidating the mechanism by which temocillin evades known mechanisms of antibacterial resistance in other infectious strains. By implementing fluorescence polarization assays, thermal shift assays, and determining complex crystal structures of temocillin and the parent compound ticarcillin with PBP3, the significant role the 6-α-methoxy group plays in destabilizing the temocillin-PBP3 complex structure was able to be determined to be the reason behind the inability of temocillin to treat *P. aeruginosa* infections. Furthermore, the blocking of Lys70 from being used as a catalytic base was determined to be the likely mechanism by which temocillin displays its ability to counter β-lactamase activity, as seen in the complex crystal structure between CTX-M-14 and temocillin. Future research for this project should look into the replacement of the 6-α-methoxy group on temocillin with other functional groups that would allow the compound to still counter β-lactamase activity while not destabilizing the PBP3-product complex, which would potentially allow the improved inhibitor to exhibit antibiotic properties against *P. aeruginosa* infections.

The third project aimed to determine crystal structures for the CCR7 binding chemokines, CCL19 and CCL21, and to utilize the sulfotyrosine recognition sites on the signaling proteins as hotspots for protein-protein interaction inhibitor design. By utilizing NMR perturbation assays and the crystal structure determined for CCL21, molecular docking experiments were carried out to identify lead compounds that targeted one of the sulfotyrosine recognition sites identified on the CCL21 protein. These compounds were confirmed to bind to the target through the use of additional NMR perturbation studies. The crystal structures determined for both CCL19 and CCL21 support the existence of two unique regulatory mechanisms for altering the signaling patterns exhibited by these proteins when binding to the CCR7 receptor. Future studies for this
project should seek to determine complex crystal structures of the lead compounds found through molecular docking studies to determine the exact interactions involved in the binding of the small molecules to CCL21. By understanding what specific interactions are involved in compound binding improvements to the affinities and specificities of the hit compounds can be made through alteration and addition of functional groups. Additional research should be carried out to identify the specific sulfotyrosine recognition sites found on CCL19 and then utilize this structural data for carrying out additional molecular docking studies that would identify novel inhibitors that target the protein-protein interface found on CCL19. Further research should also be carried out to determine if the C-terminal tail of CCL21 and the N-terminal tail of CCL19 exhibit regulatory functionality over the binding of the chemokines to their cognate receptor through structural analyses.
References

1 CDC. Antibiotic Resistance Threats in the United States. 114 (Center For Disease Control and Prevention [CDC], 2013).
36 Rupp, B. *Biomolecular crystallography : principles, practice, and application to structural biology.* (Garland Science, 2010).


57 J., U. & K., S. 1 online resource (Society of Pharmaceutical Sciences and Research, Phnakhula (HR), India, 2010).


Appendix 1:

Permissions

4/3/2019

Re: Using A Figure From Your Paper

mterrak@uliege.be

Thu 3/28/2019 3:47 AM

To: KroecK, Kyle <kkroecK@health.usf.edu>

Dear Kyle,

Yes you can use the figure

Mohammed Terrak

----- Mail original ----- 
De: "KroecK, Kyle" <kkroecK@health.usf.edu>
À: "mterrak" <mterrak@ulg.ac.be>
Objet: Using A Figure From Your Paper

Dear Dr. Terrak,

I am working on my PhD dissertation and would like permission to use Figure 1 from your paper entitled "Glycosyltransferases and Transpeptidases/Penicillin-Binding Proteins: Valuable Targets for New Antibacterials" in the dissertation. Your publisher said I should contact you for permission. Let me know if this is acceptable to you. Thank you!

Kyle KroecK
Graduate Research Assistant
Morsani College of Medicine
Department of Molecular Medicine,
University of South Florida
12901 Bruce B. Downs Blvd.
Tampa FL 33612
Email: kkroecK@health.usf.edu
Title: Crystallographic Structure of Truncated CCL21 and the Putative Sulfotyrosine-Binding Site
Author: Emmanuel W. Smith, Eric M. Lewandowski, Natasha A. Moussouras, et al
Publication: Biochemistry
Publisher: American Chemical Society
Date: Oct 1, 2016
Copyright © 2016, American Chemical Society

Quick Price Estimate
Permission for this particular request is granted for print and electronic formats, and translations, at no charge. Figures and tables may be modified. Appropriate credit should be given. Please print this page for your records and provide a copy to your publisher. Requests for up to 4 figures require only this record. Five or more figures will generate a printout of additional terms and conditions. Appropriate credit should read: "Reprinted with permission from {COMPLETE REFERENCE CITATION}. Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.

I would like to... [reuse in a Thesis/Dissertation] [Author (original work)] [Full article] [Print and Electronic] [No] [ переводе]
Select your currency [USD - $]
Quick Price [Click Quick Price]

To request permission for a type of use not listed, please contact the publisher directly.

Copyright © 2019 Copyright Clearance Center, Inc All Rights Reserved. Privacy statement, Terms and Conditions Comments? We would like to hear from you. E-mail us at customerservice@copyright.com
Re: Permissions for Adapting "CCR7 Sulfotyrosine Enhances CCL21 Binding" for my Dissertation

Veldkamp, Christopher T
Wed 4/3/2019 4:51 PM

Kyle:

As the corresponding author, you have my permission to use and modify whatever you need from the article titled "CCR7 sulfotyrosine enhances CCL21 binding" for your dissertation. If you need something more formal than this email, please draft a letter including what you require for me to work off of and I'll return it to you. Good luck with your dissertation and it's defense. I am sure you'll do great.

Chris

Christopher Veldkamp, Ph.D.
Associate Professor
University of Wisconsin-Whitewater
800 West Main Street
Whitewater, WI 53190
veldkamp@uww.edu
1-262-472-5267

On Apr 3, 2019, at 3:16 PM, Kroech, Kyle <kkroech@health.usf.edu> wrote:

Hi Dr. Veldkamp,

It's Kyle from Dr. Yu Chen's lab at USF. I am currently in the process of writing my dissertation and would like to include a bunch of the NMR data you and Brian published on CCR7 and CCL21 in the article "CCR7 Sulfotyrosine Enhances CCL21 Binding". I contacted the publisher but they relayed me to contact the corresponding author of the paper to obtain permissions since the journal you published in grants the authors the copyright. I just wanted to check to make sure I have your permission and so I can have in formal email to add to my ETD submission of my dissertation so everything clears properly. Thanks!

Kyle Kroech
Graduate Research Assistant
Morsani College of Medicine
Department of Molecular Medicine,