2019

# YCZR, a New Case of PLP-Dependent MOCR/GABR Type Transcription Regulator in Klebsiella Pneumonia

Yuanzhang Zheng

LOYOLA UNIVERSITY CHICAGO

# YCZR, A NEW CASE OF PLP-DEPENDENT MOCR/GABR TYPE TRANSCRIPTION REGULATOR IN *KLEBSIELLA PNEUMONIAE*

AN DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

PROGRAM IN CHEMISTY AND BIOCHEMISTRY

BY

YUANZHANG ZHENG

CHICAGO, IL

MAY 2019

Revised by Yuanzhang Zheng, 2019

ACKNOWLEDGMENTS

There is a list of people I need to say thank you, who play important roles not only in this paper but also through my Ph.D. program.

First, I would like to express my great appreciation to my advisor Dr. Dali Liu for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. He introduced and gave me the opportunity to work in X-ray Crystallography. His guidance helped me with research and writing of this thesis continually. Without his advice and persistent help, this dissertation would not have been possible. He also taught me how to be a man and a logical thinker.

Besides my advisor, I would like to thank the rest of thesis committee Members: Dr. Miguel A. Ballicora, Dr. Domenic Castignetti, and Dr. Wei-Tsung Lee, not only for their insightful comments and encouragement but also for the hard questions which incented me to widen my research from various perspectives. Their assistance and valuable advices on experimental design and data interpretation helped me a lot during my Ph.D. program. I appreciate their patience to allow me to learn everything as a beginner. Without their precious support it would not be possible to conduct this research.

I would also like to express my sincere gratitude to Dr. Quyen Q. Hoang from Indiana University School of Medicine and Dr. Jianbin Zheng from Northwestern University Chemistry Department. Some of the data in this thesis were collected in collaboration with them. Their valuable and constructive suggestions during the planning and development of this research work

have helped this project immensely.

I appreciate for their patient guidance, enthusiastic encouragement and useful critiques of this research work. Additionally, I would like to thank Argonne National Laboratory Structural Biology Center (SBC). This research used resources of the Advanced Photon Source on my X-ray crystallography projects.

My thanks are also extended to my fellow lab mates Dr. Romila Mascarenhas, Daniel Catlin, Dr. Rui Wu, Arseniy Butrin, Hiralben Patel, Rania Hussien, Mashael Alghamdi, and undergraduate students Nicolas Marotta, Harjot Uppal, Toaha Hussain and Audrey O'Neill for their valuable technical support on this project.

I would like to thank Loyola University Chicago and the Department of Chemistry and Biochemistry. I appreciate the Loyola University of Chicago for permission to include copyrighted photographs as part of my thesis/dissertation and provide financial support through the Ph.D. program.

I am so grateful to the researching and teaching assistantship scholarship of Loyola University Chicago for providing financial support. And also thanks for Chinese Scholarship Council (CSC) to for making it possible for me to study at Loyola Unversity Chicago.

I take this opportunity to record our sincere thanks to all the faculty members of the Department of Chemistry and Biochemistry for their help and encouragement. I also want to thank the department office staff, Carol Grimm, Stacey Lind, and Denise Hall for their help with paperwork matters; and Matthew Sara from the Stockroom for helping with laboratory/chemicals matters.

Last but not least I want to thank my family and friends, and particularly to my master supervisor Dr. Wang for helping with my graduate school application and for the constant

encouragement and guidance. I would like to thank my parents Fuhong Zheng and Zhenqing Zhang, and my lovely wife Ruiying Zhang for their unconditional love, support, and prayers, which have enabled me to complete this dissertation.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF SCHEMES xiii

LIST OF ABBREVIATIONS xiv

ABSTRACT xix

CHAPTER ONE: YCZR, A NEW CASE OF PLP-DEPENDENT MOCR/GABR TYPE TRANSCRIPTION REGULATOR IN *KLEBSIELLA PNEUMONIAE* 1
- Research Aims 1
- Gene Transcription and Transcription Factor 2
- GntR family and the subfamily classification 3
- MocR/GabR 8
- PLP Pyridoxal phosphate 11
- Domain of unknown function 12
- MocR/YczE 13

CHAPTER TWO: CRYSTAL STRUCTURE OF *KLEBSIELLA PNEUMONIAE* YCZR, THE TRANSCRIPTIONAL REGULATOR OF THE DOMAIN OF UNKNOWN FUNCTION 161 (DUF161) 15
- Research Significance 15
- YczR is a Homodimer 16
- PLP is Bound in the C-terminal AT-Fold Domain 19
- Structure of Apo-YczR 24

CHAPTER THREE: YCZR REGULATION DOMAIN CATALYZES A REVERSIBLE HALF-TRANSAMINASE REACTION INSTEAD OF BEING A NON-ENZYMATIC REGULATOR LIKE GABR 31
- UV Absorption 31
- Specificity test of YczR with amino acids or small molecular 34
- Fluorescence emission spectra 39
- An End-point Assay to Monitor the YczR Catalyzed Half Transaminase Reaction 41

CHAPTER FOUR: YCZR BINDS DNA SEQUENCE BETWEEN YCZR GENE AND YCZE GENE IN A SIMILAR FASHION TO GABR 44
- YczR Binds DNA with High Affinity 44
- YczR likely binds to the regulation region with a 2:1 stoichiometry 49
- YczR binds tighter to the site closer to yczE gene 49

Materials and Methods 52
Materials 52
Methods 52
DNA Cloning 52
Protein Purification 53
Crystallization 54
X-ray Diffraction Data Collection 55
Structure Determination, Model Building, and Refinement 56
UV-Absorption Spectroscopy 56
End-point Assay to Monitor the YczR Catalyzed Half Transaminase Reaction 57
Fluorescence emission spectroscopy 57
Size Exclusion Chromatography with Multi-Angle Light Scattering (SEC-MALS) 57
Fluorescence Polarization Assay 58

CHAPTER FIVE: CONCLUSIONS AND FUTURE INVISTIGATIONS 61
Future Work 65

CHAPTER SIX: THE AQUEOUS DECAY OF A VANADIUM COMPLEX SEREDIPUOSLY REVEALS MISSING KNOWLEDGE ON THE REDOX STATES OF PTP1B 66
Research aims 66
Results and Discussion 70
Structure of PTP1B and $VO(acac)_2$ Complex 71
Kinetic results of PTP1B with $VO(acac)_2$ Complex 81
Materials and Methods 84
Materials 84
Methods 84
DNA Cloning 84
Protein expression and purification 84
Synthesis Specific compounds as potential PTP1B inhibitors 85
Crystallization 86
Diffraction data collection, Structure Determination, and Refinement 86
PTP1B inhibition assay 87

CHAPTER SEVEN: CONCLUSIONS 89

REFERENCE LIST 93

VITA 102

LIST OF TABLES

Table 1. Crystallographic data collection and refinement statistic for YczR and Apo-YczR 9

Table 2. Crystallographic data collection and refinement statistics for PTP1B 10

LIST OF FIGURES


| | | |
|---|---|---|
| Figure 1. | Classification of *Mycobacterium smegmatis* GntR regulators including representatives of all subfamily regulators from different bacterial genomes with 1000 bootstrap replicates | 5 |
| Figure 2. | Quaternary structure of *Thermotoga maritima* GntR/ FadR bacterial transcription regulator TM0439 | 6 |
| Figure 3. | Quaternary structure of *Bacillus subtilis* GntR/ HutC bacterial transcription regulator YvoA | 7 |
| Figure 4. | Quaternary structure of *Corynebacterium glutamicum* GntR/ YtrA bacterial transcription regulator CGL2947 | 8 |
| Figure 5. | Quaternary structure of *Bacillus subtilis* GntR/ GabR bacterial transcription regulator GabR | 9 |
| Figure 6. | Schematic of GabR binding sites at the gabTD and gabR promoters | 10 |
| Figure 7. | The gel filtration purification profile of YczR | 16 |
| Figure 8. | Quaternary structure of *K. pneumoniae* YczR regulator domain | 17 |
| Figure 9. | Quaternary structure of *K. pneumoniae* YczR regulator domain | 18 |
| Figure 10. | Quaternary structure of YczR AT-fold domain monomer | 19 |

Figure 11. Structure of YczR regulator domainPLP-binding pocket 20

Figure 12. PLP binding pockets 22

Figure 13. Comparison of the structure of *K. pneumoniae* YczR with *B. subtilis* GabR 23

Figure 14. Quaternary structure of apo-YczR AT-fold domain 24

Figure 15. Comparison of the YczR-PLP and apo-YczR AT-domain 26

Figure 16. Mass spectrometry data of YczR-PLP monomer 28

Figure 17. Molecular weight determinations of YczR in solution 29

Figure 18. The overall comparison of *K. pneumoniae* YczR and *B. subtilis* GabR SAXS models 30

Figure 19. Time-resolved UV absorption spectra of YczR 32

Figure 20. UV absorption spectra of YczR-L-alanine with pyruvate 33

Figure 21. Specificity test of YczR with amino acids or small molecular 34

Figure 22. Specificity test of YczR with amino acids or small molecular 35

Figure 23. Specificity test of YczR with amino acids or small molecular 36

Figure 24. Specificity test of YczR with amino acids or small molecular 37

Figure 25. Specificity test of YczR with amino acids or small molecular 38

Figure 26. Fluorescence emmission spectra of YczR 40

Figure 27. An end-point assay to monitor the YczR catalyzed half transaminase reactio 43

Figure 28. Sequence alignment of the winged-helix domain sequences of *B. subtilis* GabR gabR-gabT DNA binding region and *K. pneumoniae* YczR yceE DNA binding region 45

Figure 29. Fluorescence polarization assays to measure DNA binding by YczR 47

Figure 30. Specificity test of YczR fluorescence polarization assays 48

Figure 31. Fluorescence polarization assays to measure DNA_1 binding by YczR 50

Figure 32. Fluorescence polarization assays to measure DNA_2 binding by YczR 51

Figure 33. Synthesis of $VO(acac)_2$ complex 69

Figure 34. Overlapping X-ray crystal structures of PTP as ribbon diagram 1B with $VO(acac)_2$ as a ribbon diagram 72

Figure 35. Electron density of the vanadate modified catalytic cysteine 73

Figure 36. Electrostatic interactions between the $VO(acac)_2$ inhibitor and the PTP1B active site 76

Figure 37. Superimposed images of sulphenyl-amide PTP1B and $VO(acac)_2$ inhibition of PTP1B show structural changes in the catalytic Cys-215 active site 78

Figure 38. Superimposition of the oxidation state structures of the PTP1B catalytic cysteine 79

Figure 39. Overlay of the WPD-loop open vanadate-bound and a WPD-loop closed conformations in the active sites of the PTP1B enzymes 80

Figure 40. Activity versus substrate concentration for PTP1B against $pNPP^{2-}$ in the 82

presence of varying $VO(acac)_2$ complex concentrations

Figure 41. Activity versus time for PTP1B 83

LIST OF SCHEMES

Scheme 1. Crystallographic data collection and refinement statistic for YczR and Apo-YczR 41

Scheme 2. The general mechanism of the PTP-catalyzed reaction 70

## LIST OF ABBREVIATIONS

| | |
|---|---|
| UTI | Urinary Tract Infections |
| CDC | Centers for Disease Control and Prevention |
| HAI | Healthcare-associated Infections |
| CRE | Carbapenem-resistant Enterobacteriaceae |
| CR-Kp | Carbapenem-resistant *K. pneumoniae* |
| KPC | *Klebsiella pneumoniae* carbapenemases |
| DoLR | Drug of Last Resort |
| TALE | Transcription Activator-like Effector |
| HTH | Helix-Turn-Helix |
| EB/o | Effector Binding and / or oligomerization domain |
| *B. subtilis* | *Bacillus Subtilis* |
| *E. coli* | *Escherichia coli* |
| *C. glutamicum* | *Corynebacterium glutamicum* |
| DNA | Deoxyribonucleic Acid |
| GntR | Repressor of the Gluconate *Operon* |
| GABA | γ-amino butyric acid |
| HEPES | N-2-Hydroxyethylpiperazine-N′-2-ethanesulfonic Acid |
| DTT | Dithiothreitol |

| | |
|---|---|
| GABA-AT | GABA Aminotransferase |
| SSA | Succinic Semialdehyde |
| EC | Enzyme Commission |
| DUF | Domains of Unknown Function |
| CDD | Conserved Domain Database |
| FPLC | Fast Protein Liquid Chromatography |
| PEG | Polyethylene Glycol |
| PLP | Pyridoxal 5'-Phosphate |
| PMP | Pyridoxamine 5'-Phosphate |
| SEC MALS | Size Exclusion Chromatography with Multi-Angle Light Scattering |
| SDS PAGE | Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis |
| TAMRA | Tetramethylrhodamine |
| SAXS | Small-Angle X-ray Scattering |
| TEM | Transmission Electron Microscopy |
| wHTH | winged Helix-Turn-Helix |
| TFs | Transcription Factors |
| Pfam | Protein Families Database |
| DBD | DNA Binding Domain |
| PDB | Protein Data Bank |
| NAD | Nicotinamide Adenine Dinucleotide |

| | |
|---|---|
| NADH | Nicotinamide adenine dinucleotide |
| AT | Aminotransferase |
| Asp-AT | Aspartate Aminotransferase |
| NMR | Nuclear Magnetic Resonance |
| BLAST | Basic Local Alignment Search Tool |
| FP | Fluorescence Polarization |
| KD | Dissociation Constant |
| UV | Ultra Violet |
| LB | Lysogeny Broth |
| IPTG | Isopropyl-β-D-Thiogalactopyranoside |
| NADP | Nicotinamide Adenine Dinucleotide |
| NADPH | Nicotinamide Adenine Dinucleotide Dehydrogenase |
| SAXS | Small Angle X-ray Scattering |
| ADP | Adenosine Diphosphate |
| ADP-Glc | Adenosine Diphosphate Glucose |
| CRC | Comprehensive Cancer Center |
| BSA | Bovine Serum Albumin |
| NCBI | National Center for Biotechnology Information database |
| PTP | Protein Tyrosine Phosphatase |
| PTP1B | Protein Tyrosine Phosphatase 1B |
| $VO(acac)_2$ | Bis(acetylacetonato)oxovanadium(IV) |
| CVO | Covalent modification Vanadium Oxide |

| | |
|---|---|
| Å | Angstrom |
| Ki | Inhibition constant |
| $pNPP^{2-}$ | *para*-nitrophenylphosphate |
| Km | Michaelis-Menten constant |
| pNP | *para*-nitrophenylphosphate |
| Kcat | Turnover number |
| $pNP^{-}$ | *para*-nitrophenol |
| n | Hill equation constant |
| OD | Optical density |
| PCR | Polymerase Chain Reaction |
| Ala/A | Alanine |
| Arg/R | Arginine |
| Asn/N | Asparagine |
| Asp/D | Aspartic Acid |
| Cys/C | Cysteine |
| Glu/E | Glutamic Acid |
| Gln/Q | Glutamine |
| Gly/G | Glycine |
| His/H | Histidine |
| Ile/I | Isoleucine |
| Leu/L | Leucine |
| Lys/K | Lysine |

| | |
|---|---|
| Met/M | Methionine |
| Phe/F | Phenylalanine |
| Pro/P | Proline |
| Thr/T | Threonine |
| Trp/W | Tryptophan |
| Tyr/Y | Tyrosine |
| Val/V | Valine |
| Ser/S | Serine |

ABSTRACT

An increasing number of genes encoding PLP-dependent transcription regulators and MocR/GabR type regulators have been identified in various bacterial genomes. However, only a handful of them, including MocR, PdxR, and GabR have been studied experimentally. They control different aspects of the bacterial metabolism. Only GabR has a reported crystallographic structure. MocR/GabR regulators possess a chimeric structure consisting of a winged helix-turn-helix (wHTH) DNA binding domain and an aminotransferase-like regulatory domain, which can bind PLP as an effector in transcription regulation. Such a chimeric construct presents a compelling case in molecular evolution. The regulation domains of all MocR/GabR type regulators loss their catalytic capacity during evolution and function as means of effector recognition and transcription regulation. A MocR/GabR homolog in *Klebsiella pneumoniae* has recently been studied; this homolog is currently named Duf161R (YczR) since it is putatively believed to control a gene encoding a membrane protein annotated as “domain of unknown function 161”.

The three-dimensional crystal structure of the regulatory domain of YczR have been determined to a resolution of 1.79 Å. Our crystallographic studies hav revealed the structure of a truncated regulation domain with PLP bound, and our spectroscopic studies have collected evidence to support at least partial transaminase-like catalytic activity of the regulatory domain. Together with DNA binding studies, we provide data on a new case of MocR regulation and its currently unconfirmed biological function, which is likely augmenting pathogenesis via

facilitating taurine trafficking in *K. pneumoniae*.

Together with DNA binding studies, we provide data on a new case of MocR regulation and its currently unconfirmed biological function, which is likely augmenting pathogenesis via facilitating taurine trafficking in *K. pneumoniae*.

During my Ph.D. program, I also worked on several other protein projects. I have included chapter two for the PTP1B project. Protein tyrosine phosphatase 1B (PTP1B) is an enzyme shown to play an essential role in insulin regulation. PTP1B is a critical negative regulator of insulin and leptin signaling pathways by removing phosphate groups ($PO_4^{3-}$) from insulin receptor and other post-receptor substrates. Previous studies have identified transition metal compounds that exhibit insulin-mimetic effects. A plausible explanation is that vanadium-containing compounds and zinc-containing compounds inhibit PTP1B activity, which allows the required phosphorylation reaction to proceed normally. One specifically modified vanadium-containing compound has been synthesized. This research has determined the three-dimensional crystal structure of PTP1B and $VO(acac)_2$ complex to a resolution of 2.2 Å. Furthermore, the kinetic data suggest a mixed inhibition because of the aqueous study of the vanadium complex.

# CHAPTER ONE

# YCZR, A NEW CASE OF PLP-DEPENDENT MOCR/GABR TYPE TRANSCRIPTION REGULATOR IN *KLEBSIELLA PNEUMONIAE*

## Research Aims: YczR may be Potential Anti-Virulence Target for *Klebsiella pneumoniae*

*Klebsiella pneumoniae* is a Gram-negative, facultatively anaerobic bacterium[1-3]. It can reside in the human flora of the mouth, skin, and intestines[2, 4, 5]. Once aspirated to the lungs, specifically to the alveoli, it can cause hemoptysis and destructive changes to human and animal lungs[3, 6]. It also can lead to a range of different illnesses, such as wound infections, urinary tract infections (UTI), bloodstream infections, meningitis, surgical site infections, and pneumonia[7-9]. *K. pneumoniae* is a multi-drug resistant bacterium and consumes taurine as part of the pathogenesis infecting humans.

*Klebsiella* infections commonly occur among patients who take long courses of certain antibiotics. Patients who have intravenous catheters or are on a ventilator are are also at a higher risk of *Klebsiella* infections[10-14].

According to the Centers for Disease Control and Prevention (CDC) Healthcare-associated Infections (HAI) reports, doctors use several powerful antibiotics, such as polymyxins, fosfomycin, and aminoglycosides in conjunction to treat *Klebsiella* infections. However, *Klebsiella* is increasingly becoming resistant to antibiotics[4]. The highly antibiotic-resistant "superbug" bacteria belong to the carbapenem-resistant *Enterobacteriaceae* (CRE) family[5, 15].

Carbapenem-resistant *K. pneumoniae* (CR-Kp) is the most common type of CRE. According to CDC guidelines, CR-Kp is responsible for approximately 7,900 infections and 520 deaths each year[16, 17]. Some strains of *K. pneumoniae* can produce *Klebsiella pneumoniae* carbapenemases (KPC) to render the antibiotics ineffective. KPC-producing bacteria are resistant to most classes of known antibiotics, including carbapenems, which are considered as a drug of last resort (DoLR) against antibiotic-resistant Gram-negative infections[6, 16]. Drug resistance and diminished effect of medication on bacteria and other infectious diseases are becoming one of the most significant health issues of the 21st century. To combat this problem, novel antimicrobial targets need to be identified[2].

## Gene Transcription and Transcription Factors

Gene transcription is fundamentally important for metabolism and normal physiological activities at the first step of gene expression. During transcription, RNA polymerase uses a particular segment of genomic DNA as a template to synthesize a single-stranded complementary mRNA primary transcript[18]. Therefore, transcription is a highly regulated process.

Bacteria adapt to environmental changes by carrying out highly sophisticated responses via transcription regulation. Transcription factors (TFs) control transcription regulation by binding to specific DNA sequences and leading to either activation or repression of certain genes. TFs recognition of specific DNA binding sites is essential for mediating gene expression and repression in different cellular contexts.

TFs employ a variety of mechanisms to recognize target nucleotide sequences such as hydrophobic contacts and hydrogen bonds between DNA bases and amino acid side chains. They can bind to the specific promoter or regulatory regions of DNA by a variety of unique domains or motifs such as zinc fingers, leucine zippers, helix-turn-helix (HTH) domains, sheet DNA-binding proteins, and transcription activator-like effector (TALE) domains. TFs can also recognize the local three-dimensional (3D) shape of the target DNA sequence at the protein-binding site, including sequence dependent narrowing of the DNA minor groove.

TF's play several critical biological roles in bacteria, such as basal transcription regulation, cell cycle control, intercellular signal response, pathogenesis, environmental response, and differential enhancement of transcription[19, 20]. Therefore, TFs are potential antimicrobial targets.

Based on biophysical, molecular, computational, and X-ray crystallography studies over the past decade, scientists have located numerous TFs binding sites, by deducing from aligning a set of DNA sequence that is experimentally known[21]. However, understanding the TFs binding mechanisms has proven to be a challenge, especially concerning mechanisms details. The primary difficulty is in bridging the gap between genomic-scale overview and molecular description at molecular level.

## GntR family and the subfamily classification

The GntR family is a widespread and large group of bacterial TFs. It was named GntR after Haydon and Guest had identified and described the first GntR family member the *Bacillus subtilis* gluconate operon repressor in 1991[22]. The database of protein families (Pfam) currently stores approximately 8,561 regulatory proteins of the GntR family (Pfam family: GntR,

PF00392), which spread out among 764 bacteria taxa[23]. In the protein data bank (PDB) the metabolite-response GntR family has approximately 106 protein structures been solved.

GntR family TFs often bind to DNA sequences adjacent to the genes which they control. There are many cases in which TFs bind as effectors. In some instances, GntR TFs are also not located close to the gene it regulates. One of the well-studied examples is FadR, the multifunctional regulator of fatty acid metabolism in *E. coli*[24].

The GntR family TFs can regulate various biological processes including metabolism of fatty acid, amino acids, utilization of sugars and other carbon sources[22, 25]. Among all bacterial transcription regulators, the GntR superfamily proteins are structurally featured by N-terminal wHTH DNA-binding domains along with a heterogeneous C-terminal regulatory, effector-binding/oligomerization (Eb/O) domain, which provides a basis for the constituent subfamilies classification.

While the DNA binding domains in this superfamily share a significant level of similarity, as all exhibit the wHTH domain with the canonical HTH motif followed by a beta-turn-beta wing hairpin, the C-terminal regulatory ligand binding domains vary significantly. The characteristics of the C-terminal domains provide a basis for the current classification of the GntR subfamilies: HutC, MocR, YtrA, AraR, DevA, PlmA[26-28] (Figure 1) and, the largest subfamily, FadR[25, 27, 29].

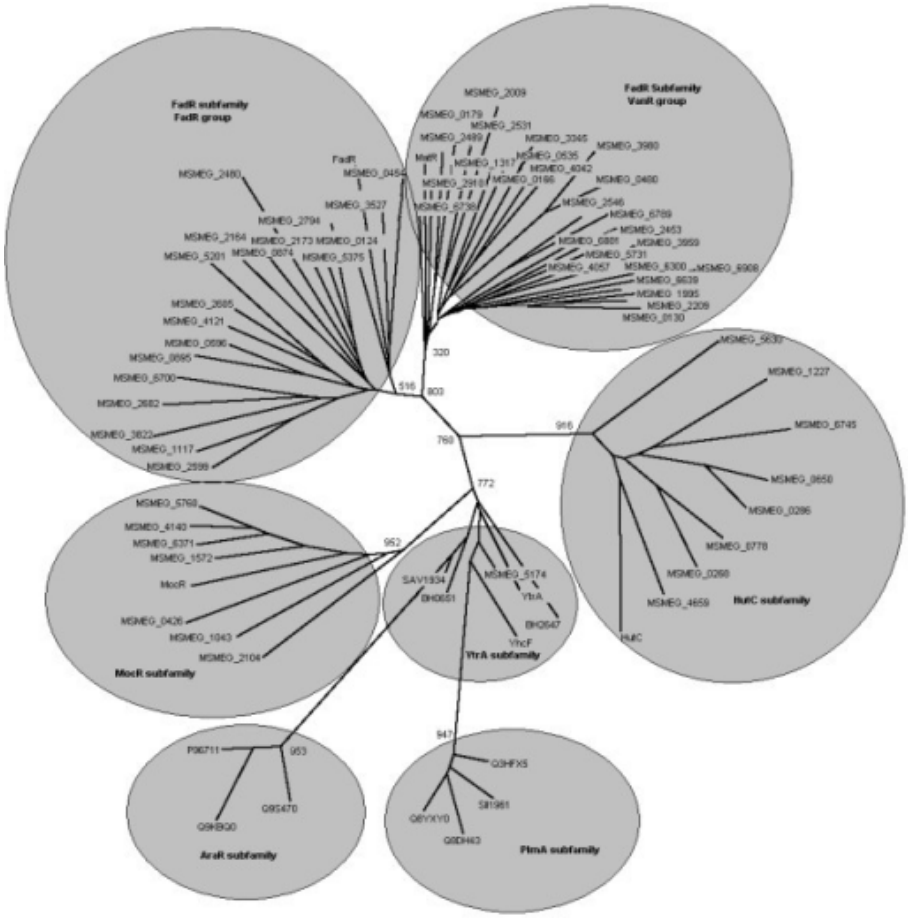



Figure 1. Classification of *Mycobacterium smegmatis* GntR regulators including representatives of all subfamily regulators from different bacterial genomes with 1000 bootstrap replicates. All the GntR regulators are clustered into six subfamilies. FadR subfamily is branched again into two groups (VanR and FadR)[26].

FadR subfamily is the largest GntR subfamily of all GntR-like regulators. The FadR family primarily functions in regulating the expression of enzymes responsible for oxidizing substrates related to amino acids, or substrates that emerge from the central metabolism. FadR is named after the best-characterized GntR repressor, the acyl-CoA responsive regulator that regulates fatty acid metabolism[30, 31]. The crystal structure for FadR of *Thermotoga maritima* transcription regulator TM0439 is shown in Figure 2 (PDB ID: 3FMS)[32].

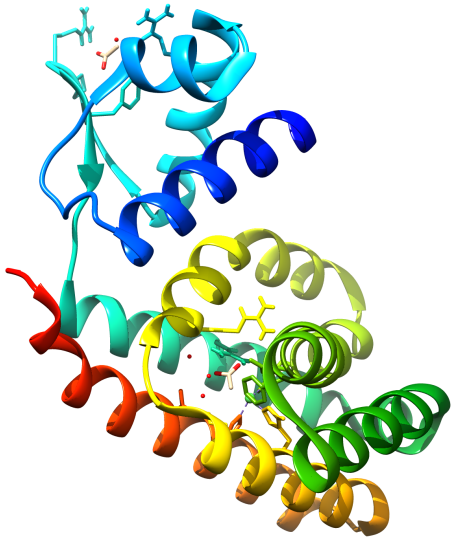


Figure 2. Quaternary structure of *Thermotoga maritima* GntR/ FadR bacterial transcription regulator TM0439 (PDB: 3FMS)[32].

The FadR C-terminal domain characteristically contains 150-160 amino acids, which form a six to seven α-helical bundle that binds carboxylic acids and other small organic ligands. A unique feature of this family is a kink in helix four which is involved in dimerization.

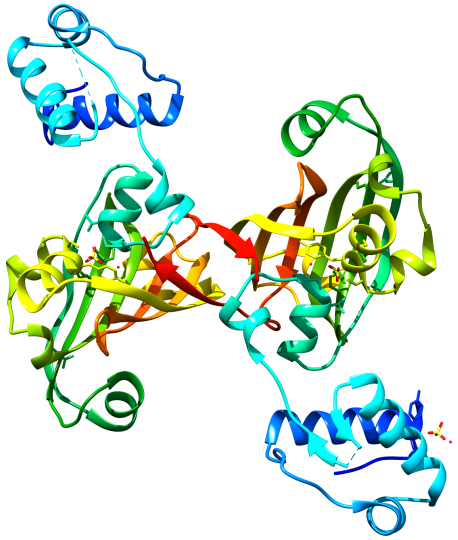


Figure 3. Quaternary structure of *Bacillus subtilis* GntR/ HutC bacterial transcription regulator YvoA(PDB: 2WV0)[33].

HutC is the second major GntR subfamily accounts of all GntR-family transcription regulators[31]. The HutC C-terminal domain is approximately 170 amino acids in length and consists of α-helices and β-sheets [25]. Although the HutC C-terminal domain lacks enzymatic activity, it shares a similar fold with chorismate from *E.coli*; it can bind effectors such as histidine, fatty acids, and sugars. This subfamily derives its name from HutC regulator in *Pseudomonas putida* which is involved in regulating histidine utilizing genes[31]. Capable of binding to a wide range of effector ligands, HutC regulators are involved in a variety of processes such as conjugative plasmid transfer, sensing of nutritional status in the environment, and antibiotic production[30, 31]. The X-ray structure of the GntR/HutC bacterial transcription regulator YvoA is shown (Figure 3; PDB ID: 2WV0)[33].

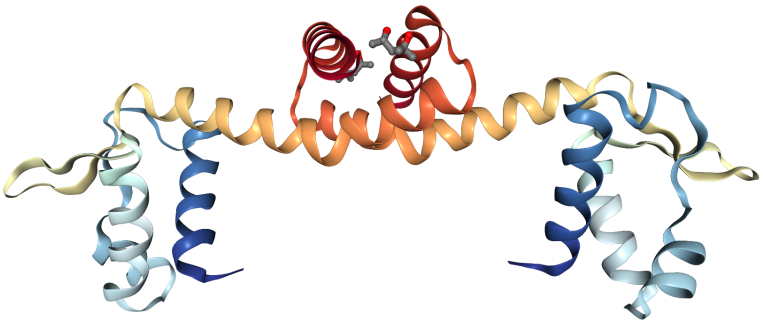


Figure 4. Quaternary structure of *Corynebacterium glutamicum* GntR/ YtrA bacterial transcription regulator CGL2947 (PDB: 2DU9).

Most members of the YtrA subfamily form operons with ATP-biding cassette transport systems[25]. This subfamily contains a characteristically short 50 amino acid two helices C-terminal domain. The short domain suggests that effector ligand binding is not likely without dimerization. Dimerization is commonplace as many GntR-like palindromic repeats have been

identified upstream of regulated operons[30]. Furthermore, the abundance of hydrophobic, aromatic, negatively charged, and positively charged residues at the end of the domain suggests that dimerization most likely occurs through side-chain-side-chain interactions and salt bridges[25]. This hypothesis was verified via the crystal structure of CGL2947 from *Corynebacterium glutamicum* (Figure 4), which showed a homodimer assembly[34].

**MocR/GabR**

MocR/GabR, the last of the major families, is characterized by a large C-terminal/oligomerization domain with an average of length of 350 amino acids (Figure 5). This domain is homologous to type I aminotransferases, which catalyze the formation of α-keto acids from amino acids using PLP as a cofactor[25]. GabR from *B. subtilis*, the representative member of this family, regulates γ-aminobutyric acid (GABA) metabolism. GABA is an important carbon and nitrogen source. In many bacteria, *B. subtilis* GABA could act as the sole nitrogen source[35]. GabR regulates the gabTD pathway in a PLP and GABA dependent manner (Figure 6). In the presence of PLP, GabR will repress its transcription, upon binding to GABA, GabR upregulates the gabTD pathway promoting GABA metabolism and glutamate biosynthesis. The gabTD pathway consists of two genes, which allow for the use of intracellular GABA to generate succinate and glutamate[35].

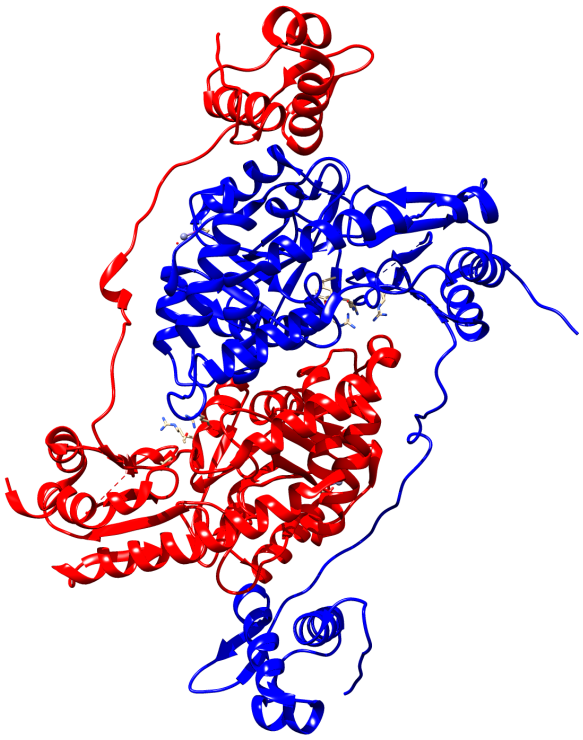


Figure 5. Quaternary structure of *Bacillus subtilis* GntR/ GabR bacterial transcription regulator GabR (PDB: 4N0B)[36].

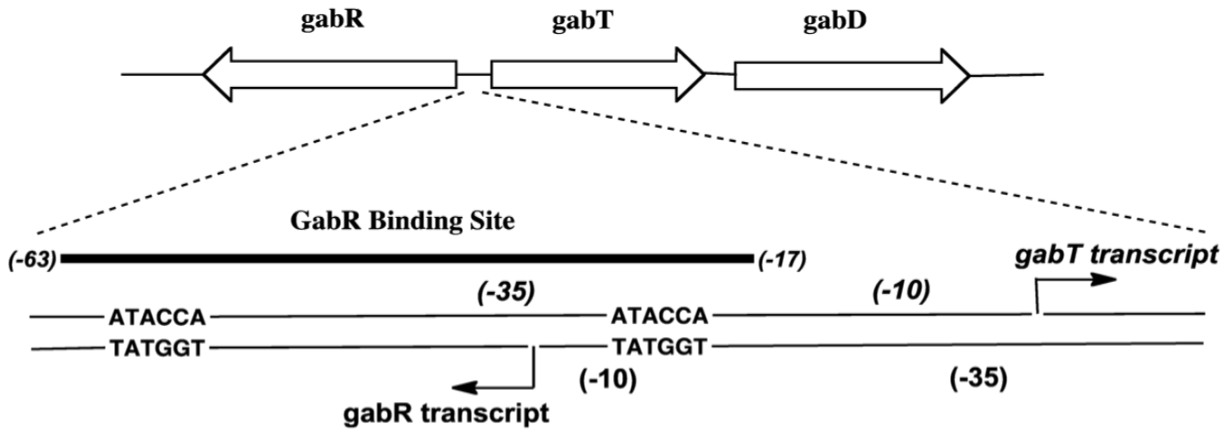



Figure 6. Schematic of GabR binding sites at the gabTD and gabR promoters. A schematic of the gabR and gabTD gene location is shown here. The arrows indicate the transcription directions and starting sites for the gabTD and gabR genes. Within the gabR-gabT regulatory region, the

two direct repeats (ATACCA) are indicated, and the 47-bp GabR binding site (-63 to -17) is shown in detail.

The two genes, gabT and gabD, encode for the enzymes GABA aminotransferase (GABA-AT) and succinic semialdehyde dehydrogenase (SSADH), respectively. GABA-AT catalyzes the ping-pong transamination using GABA as an amino group donor, PLP as a coenzyme, and α-ketoglutarate as an amino group acceptor. In the first half reaction, GABA donates its amino group to PLP forming pyridoxamine 5'-phosphate (PMP) and releases succinic semialdehyde. In the second half reaction, PMP donates its amino group to α-ketoglutarate forming glutamate and regenerating PLP. Downstream, the GabD succinate semialdehyde dehydrogenase will convert succinic semialdehyde to succinate[37].

## PLP Pyridoxal phosphate

Pyridoxal 5'-Phosphate (PLP) is the active biological form of vitamin B6, which exists in the form of six interconvertible vitamers: pyridoxine, pyridoxamine, pyridoxal and their 5'-phosphorylated forms[38-40]. PLP acts as a coenzyme in a variety of enzymatic reactions in all organisms, such as deamination, decarboxylation, racemization reactions of amino acids and in vast majority of transamination reactions[41].

The Enzyme Commission (EC) has identified more than 145 PLP-dependent activities which corresponding to about 4% of all classified activities[42]. Almost 1.5% of genes generally encode PLP-dependent enzymes in the free-living prokaryotic genomes[42]. The most common PLP-dependent enzymes are phosphorylases, aminotransferases, decarboxylases, glycogen amino acid racemases, and enzymes catalyzing β- or γ-elimination or replacement. The utility of PLP lies in its ability to stabilize carbanionic reaction intermediates by acting as an electrophilic catalyst after covalently binding to substrates. PLP is also involved in stress responses, especially to oxidative stress, which can quench reactive oxygen species[43]. The dysregulation of PLP metabolism can cause various physiological deficiencies, such as reduced glucan production, lowered stress tolerance, growth inhibition and biofilm formation[43].

Some PLP-dependent enzymes generate enamine/imine intermediates that can inactivate other enzymes by covalent binding to their active site, whereas the RidA-family proteins quench these reactive intermediates to prevent metabolic damage[44].

The PLP aldehyde group can form a Schiff base linkage (internal aldimine) with the ε-amino group of the specific lysine group of the aminotransferase enzyme. In the

transaldimination, the amino acid substrate α-amino group displaces the ε-amino group of the active-site lysine. And then, the external aldimine can lose a proton, an amino acid side chain or carbon dioxide to form a quinoid intermediate, which can act as a nucleophile in several reaction pathways[38, 45, 46].

**Domain of unknown function**

Domains of unknown function (DUF) are a large set of uncharacterized protein families that are found in the Pfam database that has no characterized function[47, 48]. More than 20% of all protein domains in the protein family's database were annotated as DUFs in 2013[49]. The number of DUFs is substantially increased with the rapidly accumulating genome sequencing data[47]. These genes have been collectively given in the Protein family's database the prefix DUF followed by a number. Some DUFs are highly conserved, indicating an important role in biology. However, the biological function of many DUFs often remains unknown because such proteins are not essential[48, 49]. Structural genomics program has attempted to understand the function of DUFs through structure determination. The structure of more than 250 DUF families proteins have been solved[49]. These protein families have been collected together in the Pfam database using the prefix DUF followed by a number. There are over 3,000 DUF families in the Pfam database representing over 20% of known families[49].

DUF161 is an uncharacterized hypothetical membrane protein, belonging to YitT family. This entry describes proteins with a trans-membrane domain. The functions of this family are unclear. The PDB entries over 50,000 protein structures, but less than 1% of them are membrane

proteins. Membrane proteins have proven to be challenging to study owing to their flexibility, partially hydrophobic surfaces, and lack of stability[50].

**MocR/YczE**

Generally, YczE and YitT family protein have been annotated as "membrane proteins containing the DUF161 domain". YczR, in *K. pneumoniae*, is the transcriptional regulator of DUF161. The crystal structure of YczR would provide information to speculate on the function of DUF161.

The TF of yczE genes, named YczR, predicted to constitute a MocR subfamily. Bacterial YczE family proteins are predicted to be unknown function membrane proteins which prosessing five trans-membrane helices. Based on the Conserved Domain Database (CDD), code COG2364 belongs to the YczE family. Corresponding to the Pfam database (code: PF02588), InterPro databank (code IPR003740), and CDD (code COG1284), YczR and YitT family shared the presence of DUF161 domains. Based on the annotations report of the databanks, YczE and YitT family do not have apparent differences[51]. Possibly, they may be distinguished by the absence of the C-terminal DUF2179 domain in YczE, corresponding to the Protein Data Bank (PDB) structure PDB: 3HLU, present in the members of the YitT family[51, 52].

YczR, in *K. pneumoniae*, is the transcriptional activator of YczE, which belongs to the transcriptional regulator MocR family. It is supposed to regulate the expression of the yczE genes. MocR regulators are a subfamily belonging to the class of GntR regulators characterized by the presence of a short N-terminal wHTH DNA-binding domain and a long C-terminal

regulatory domain. The C-terminal domains belong to the superfamily of the PLP enzymes of the fold type I[53].

Two members of the MocR/GntR family have been characterized in detail, GabR, and PdxR. GabR positively regulates the expression of the gabTD operon, which is responsible for the utilization of gamma-Aminobutyric acid[36, 54]. PdxR regulates the expression of the pdxST genes[55].

Obtaining the structure of YczR will help us in speculating the function of YczE. In our current research, we have solved the *K. pneumoniae* YczR regulator domain crystal structure and apo-YczR crystal structure. Based on the structure, we speculate that YczR might have a half alanine-glyoxylate transaminase activity. Based on this speculation, we designed the UV-absorption spectroscopic measurements, end-point enzymatic assays, and fluorescence spectroscopy. Alanine-glyoxylate aminotransferase catalyzes the transfer of the L-alanine to glyoxylate, forming glycine and pyruvate. In the first half-transamination reaction, L-alanine reacts with the pyridoxal form of the enzyme to yield the pyruvate and the pyridoxamine form of the enzyme[56-60]. We also used size exclusion chromatography with multi-angle light scattering (SEC-MALS), small-angle x-ray scattering (SAXS), mass spectrometry, and transmission electron microscopy (TEM) methods to gain an in-depth understanding of the characteristics of YczR and valuable clues for the function of YczE. We used fluorescence polarization assay to determine the YczR binding DNA sequence. Fluorescence polarization assays are suitable for the analysis of binding between the small fluorescence-labeled DNA fragment and its significantly larger binding protein[61].

# CHAPTER TWO

# CRYSTAL STRUCTURE OF *KLEBSIELLA PNEUMONIAE* YCZR, THE TRANSCRIPTIONAL REGULATOR OF THE DOMAIN OF UNKNOWN FUNCTION 161 (DUF161)

## Research Significance

YczR is a member of the understudied MocR/GabR subfamily of the GntR family of transcription regulators. A Typical GntR/MocR subfamily protein contains a large C-terminal putative type I aminotransferase domain and a small N terminal wHTH DNA binding domain [25, 35].

In the present study, we report two *K. pneumoniae* YczR regulator domain crystal structures: a 1.79 Å structure of YczR with PLP bound and the 2.0 Å apo structure of YczR regulator without PLP. We also used size exclusion chromatography with multi-angle light scattering (SEC-MALS), small-angle x-ray scattering (SAXS), and mass spectrometry methods to gain an in-depth understanding of the biological assembly of YczR. To our best knowledge, these crystal structures of YczR were described for the first time.

## YczR is a Homodimer

The YczR protein of *K. pneumoniae* has a predicted size of 475 amino acids and a theoretical molecular mass of 52.25 kDa. YczR was eluted as one major peak on a Superdex 200 gel-filtration column during protein purification, with an apparent molecular weight of a dimer (108 kDa), indicating a dimer of YczR in solution (Figure 7).

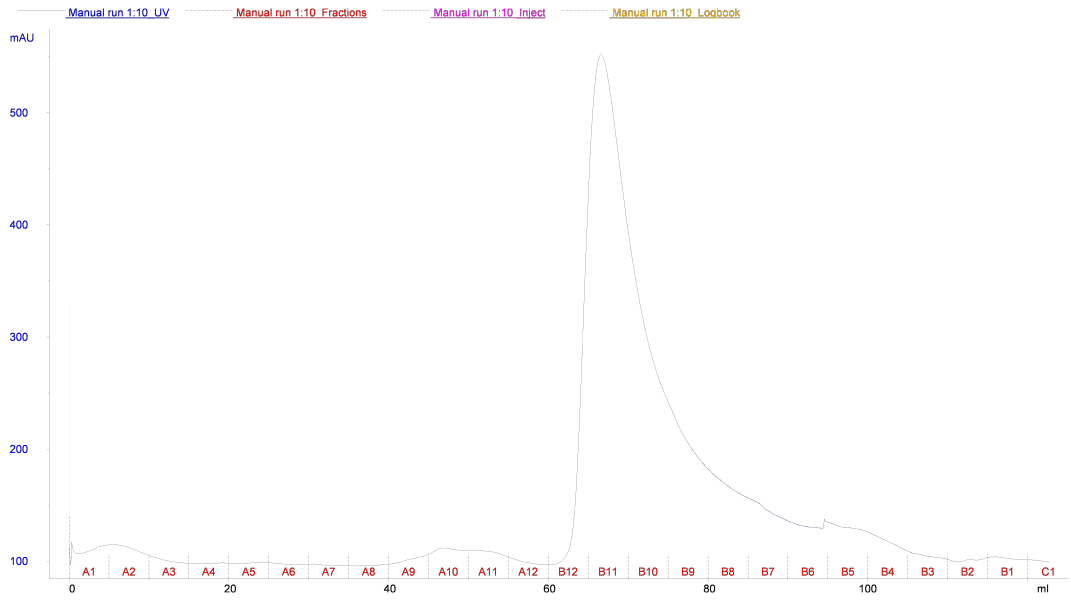



Figure 1. The gel filtration purification profile of YczR. The YczR major peak (fractions B9-B12) on Superdex 200 16/60 corresponds to an apparent molecular weight of a dimer.

YczR crystallizes as a homodimer with the two AT-fold domains organized in a head-to-head domain arrangement (Figure 8 & 9). As seen in all type I aminotransferases[62, 63], this interface of YzcR contains two PLP-binding pockets bound by one molecule of PLP in each pocket. The large dimer interface between the two AT-domains further suggests that any such domain movements by the wHTH domain do not require restructuring of the AT-AT interface.

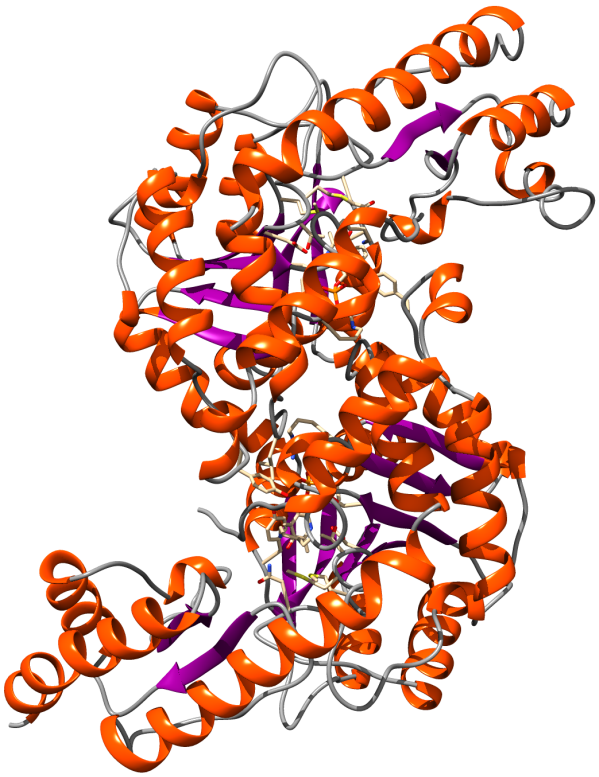


Figure 2. Quaternary structure of *K. pneumoniae* YczR regulator domain. YczR should be a head-to-tail domain-swap homodimer. The AT-fold domains are shown in red and blue. An orthogonal view of the structure is shown in Figure 9.

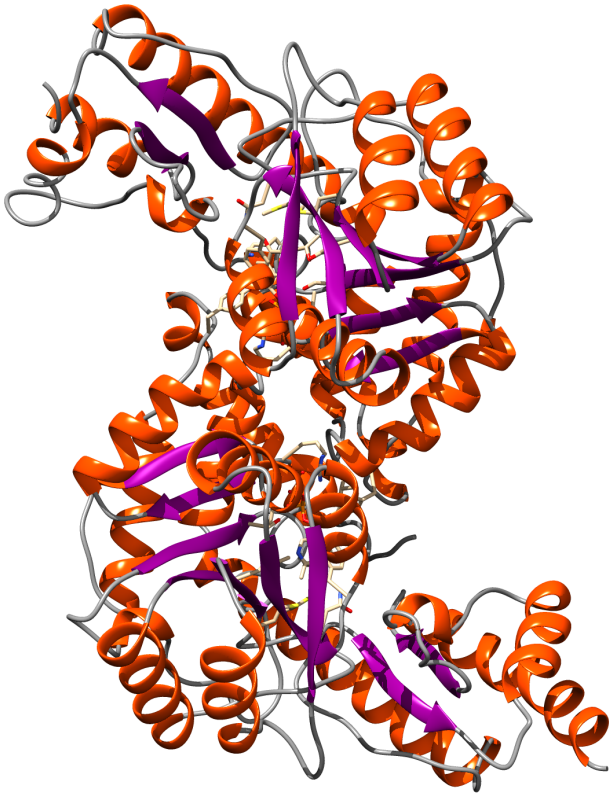


Figure 3. Quaternary structure of *K. pneumoniae* YczR regulator domain. The structure is rotated 90° counterclockwise along the y axis relative to Figure 8.

The structure of YczR was solved by molecular replacement and refined to 1.79 Å resolutions. The final model contains 385 residues, yielding a crystallographic R-value of 19.11% and an R-free value of 22.84% with near ideal chemical geometry. Two molecules in an asymmetric unit formed a homodimer. Two active sites at the interface of the subunits are shown with PLP in element-colored sticks.

**PLP is Bound in the C-terminal AT-Fold Domain**

Based on the amino sequence, the YczR should have a short N-terminal wHTH DNA-bing domain and a long C-terminal PLP-binding putative aminotransferase domain. However, the N-terminal wHTH DNA-binding domains have not been found in YczR crystal structure. We speculated that we lost the DNA-binding domain when we purified the protein or during crystallization. The regulation domain structure still provides insights into YczR funcations.

PLP is bound to the C-terminal regulator Domain (Figure 10). The PLP binding pocket also corresponds to the cofactor binding sites in the related enzymes and therefore defines the active site of YczR.

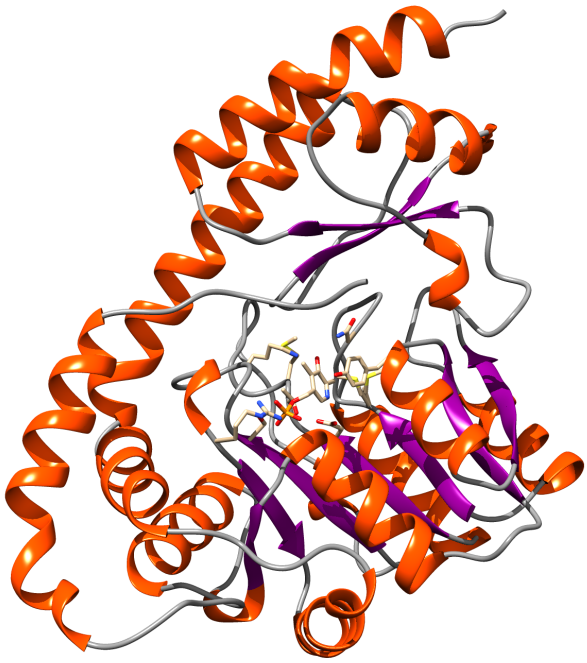


Figure 4. Quaternary structure of YczR AT-fold domain monomer. PLP is bound in the AT-fold domain of YczR. Side chains of the key residues that interact with PLP are shown in stick representation. The secondary structure elements are shown in cartoon form.

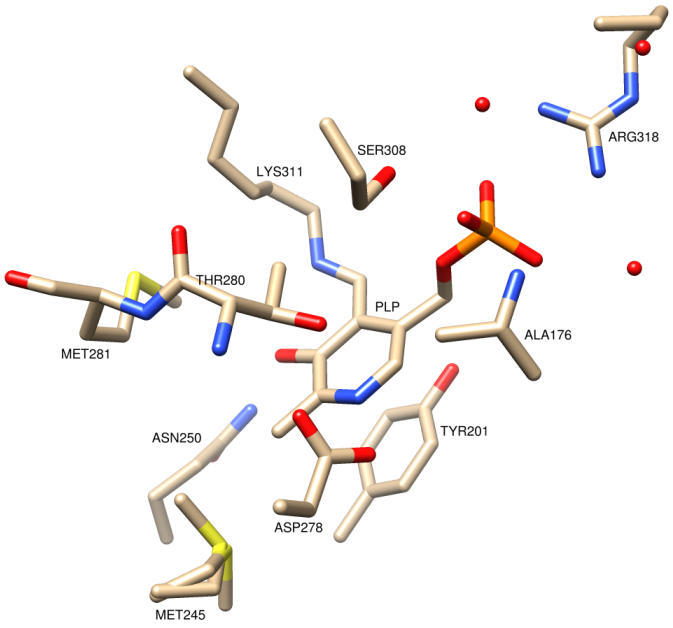



Figure 5. Structure of YczR regulator domainPLP-binding pocket. The residues with 4 and 5 Å from K311-PLP were labeled. The PLP and surrounding side chains are shown as sticks. Carbon atoms are shown in gray, oxygen atoms in red, and nitrogen atoms in blue.

One molecule of PLP forms an internal aldimine with the ω-amino group of K311 per YczR monomer. As with other types I aminotransferases[62, 63], the YczR AT-fold domain has two subdomains, the small subdomain that typically interacts with the α- carboxylate group of the substrate and the large subdomain containing the invariant lysine (K311) that forms the Schiff base with PLP. The interaction between the ASP278 and pyridine nitrogen of PLP probably could protonate the pyridine nitrogen. Although the effectors were unknown, those structure features remain.

In amino transferase reaction, the principal function of PLP is to stabilize the negative charge generated at external aldimine between PLP and the amino acid substrate or the Cα atom

of the transition state by delocalizing the negative charge through the π-electron system of the cofactor[36].

The YczR AT-fold domain residues that interact with PLP are shown (Figure. 11): The C4A atom of PLP is covalently attached to the LYS311 through the formation of an internal Schiff base. The PLP-pyridine ring stacked with residues Tyr-201 by formed pi-pi hydrophobic interactions. The C2A atom of PLP exhibits hydrophobic interaction with a Met-245 side chain. The phosphate group of PLP forms hydrogen bonds and salt bridges with Ser-308, Ala-176, and Arg-318. The side chains of Thr-280, Asp-278 and Asn-250 are hydrogen bonded to N-1 and O-3 of the PLP-pyridoxal, respectively. The phosphate moiety of the PLP formed hydrogen bonds with the peptide side chains of residues Arg-318 and Ala-176. This conformation allows hydrogen bond with the Lys-311 and could also interact with the amino group from the pyridoxamine intermediate during the enzymatic reaction.

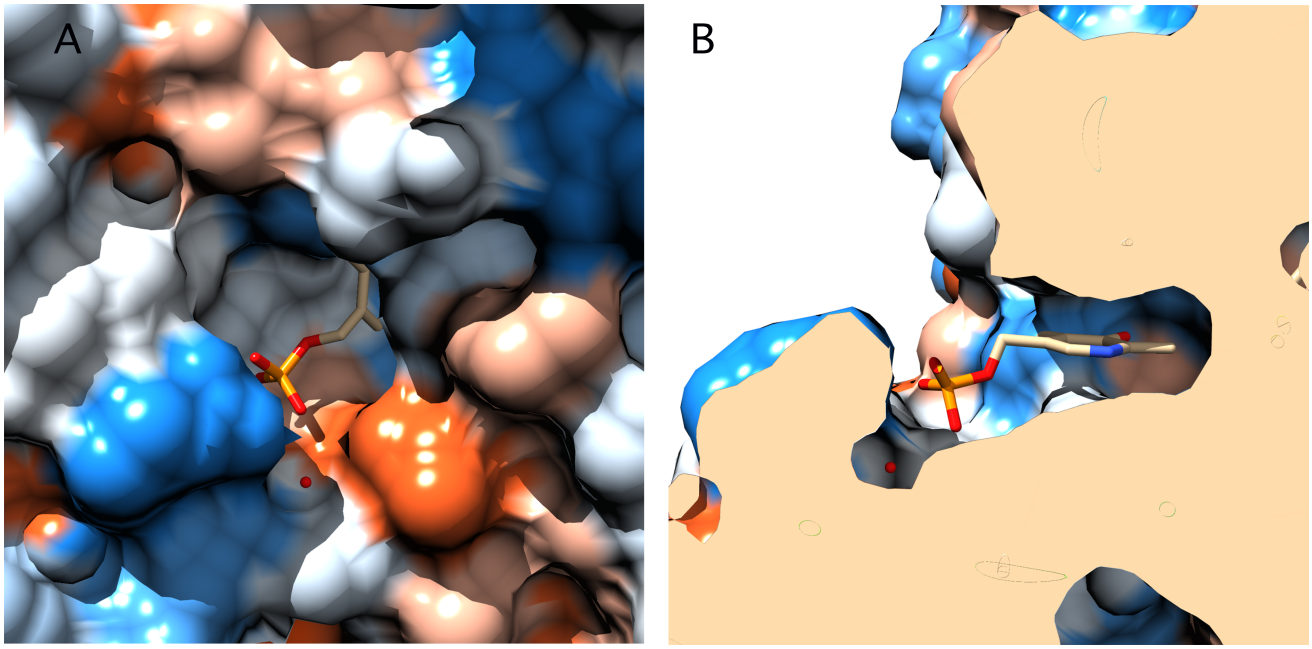



Figure 6. PLP binding pockets. (A) A surface-coated YczR(blue) is shown with the PLP product as sticks (carbons colored gray, oxygen colored red and nitrogen colored blue). The PLP is buried below the surface. (B) A rotated, cut-away view of panel A, showing a defined pocket for binding the PLP, while the opened ring is pointing toward the solvent.

Compared with the structure of GabR, PLP binding pocket in the YczR AT-fold domain has bigger size of the opening (Figure 12). PLP is buried below the surface of the protein. Once the internal aldimine between PLP and the K311 broken, the molecular PLP or PMP would be released from binding pocket easily.

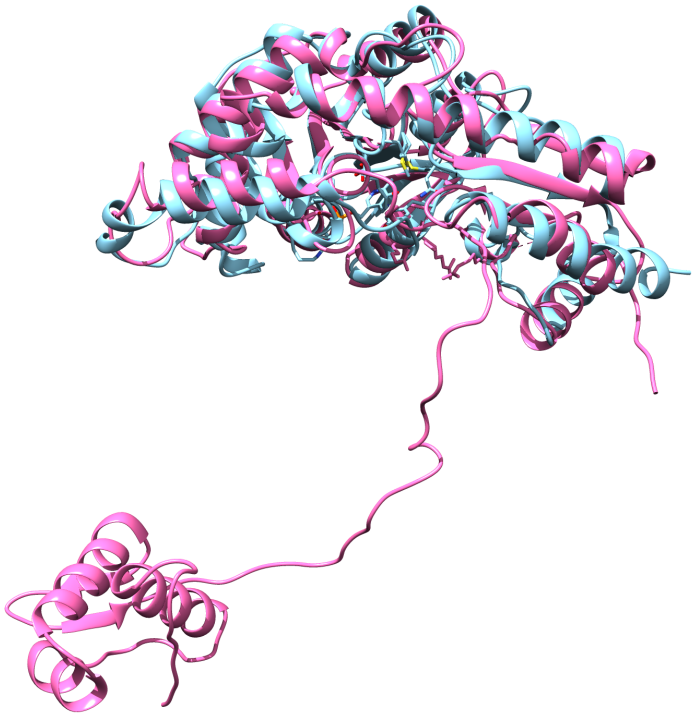


Figure 7. Comparison of the structure of *K. pneumoniae* YczR with *B. subtilis* GabR (PDB ID: 4N0B). An orthogonal view of the GabR structure is shown in pink. An orthogonal view of the YczR AT-fold domain is shown in blue.

The crystal structures of *K. pneumoniae* YczR and *B. subtilis* GabR (PDB ID: 4NOB) were compared to analyze the difference in the overall structure (Figure 13). The YczR AT domain size was very similar to that *B. subtilis* GabR reported. GabR has a long linker connecting C-terminal AT-fold domain to the N-terminal winged-helix domain. YczR lost the N-terminal DNA binding domain in the crystal structure.

**Structure of Apo-YczR**

We also attempted to get YczR with L-alanine co-crystal. In the resulting crystal structure, PLP is no longer bound in the PLP-binding pocket. The PLP-binding site of the crystal structure of apo-YczR is shown (Figure 14). The purified YczR protein is pale yellow, which suggested that endogenous PLP from *E. coli* bound to YczR after purification. After co-crystallization of YczR and L-alanine, the apo-YczR crystal is clear in color. This result indicated that Schiff base between the invariant lysine (Lys-311) and PLP had been broken.

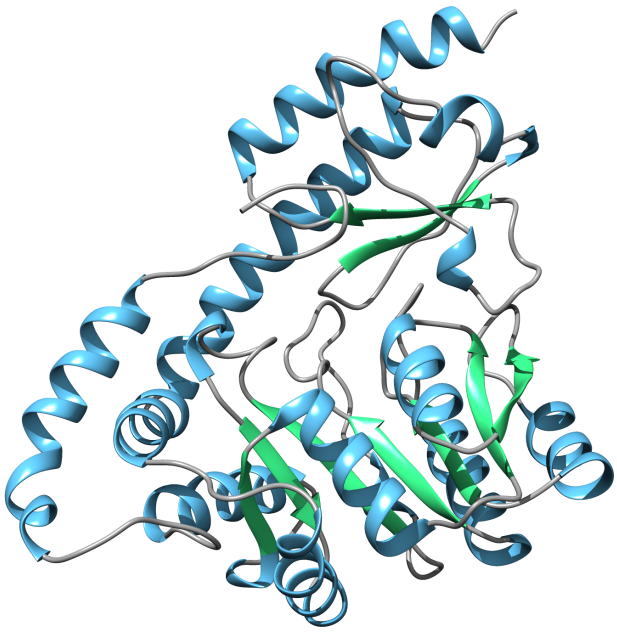


Figure 8. Quaternary structure of apo-YczR AT-fold domain. PLP is no longer bound in the PLP-binding pocket. The secondary structure elements are shown in cartoon form.

One small molecular binds with Ser-308. Apo-YczR crystal crystallization solution contained 8% (V/V) Tacsimate, pH 6.0. Tacsimate is a unique crystallization reagent, which is composed of a mixture of titrated organic acid salts, developed exclusively by Hampton Research. Tacsimate contains malonic acid, ammonium citrate tribasic, succinic acid, DL-malic acid, sodium acetate trihydrate, sodium formate, and ammonium tartrate dibasic. We cannot jump to the conclusion that L-alanine replaced PLP. We still need to get the same structure result from crystallization solution without Tacsimate.

The crystal structures of YczR-PLP and apo-YczR were compared to analyze the difference in the overall structure (Figure 15). The obtained apo-YczR structure with no PLP bound has a similar overall conformation when compared to YczR-PLP. The overall similarity suggests that the apo-YczR AT-fold domain maintains the overall structural integrity even when PLP is no longer bound.

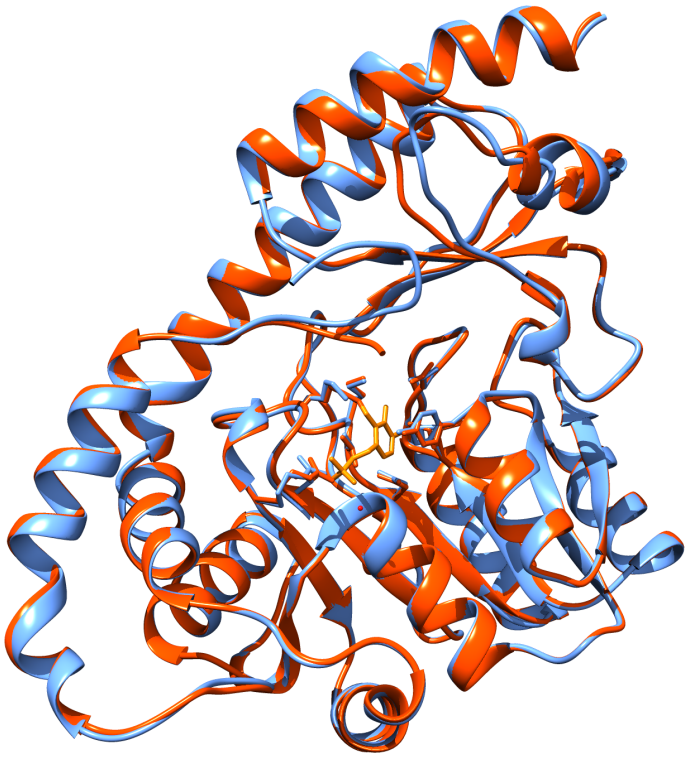


Figure 9. Comparison of the YczR-PLP and apo-YczR AT-domain. The YczR-PLP structure is shown in orange red and in cyan for the apo-YczR structure.

Table 1. Crystallographic data collection and refinement statistics for YczR and Apo-YczR

| | YczR | Apo-YczR |
|---|---|---|
| Data Processing | | |
| Space group | $C222_1$ | $C222_1$ |
| Cell dimension | | |
| α, β, γ (deg) | 123.482, 181.139, 108.567 | 123.24, 180.005, 106.006 |
| a, b, c (Å) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 |
| Resolution (Å) | 1.82 | 2.46 |
| Resolution at I/σ (I) =2 | 1.85 | 2.73 |
| $R_{pim}$ [a](%) | 2.9(38) | 5.5(60) |
| I/σ (I) | 25.5(13.4) | 13.2(10.9) |
| CC ½ [b] | 0.757 | 0.575 |
| Completeness (%) | 99.4(91.7) | 99.7(97.3) |
| Multiplicity | | |
| No. Reflections | 110262 | |
| No. Unique Reflections | | 45635 |
| Refinement | | |
| $R_{work}$ [c]/$R_{free}$ [d](%) | 17.64/21.13 | 20.8/28.96 |
| No. of Atoms | | |
| protein | 8448 | 8001 |
| ligand | 60 | 30 |
| water | 650 | 152 |
| B factors (Å$^2$) | | |
| protein | 30.0 | 48.912 |
| ligand | | |
| RMSD [e] | | |
| bond lengths (Å) | 0.02 | 0.01 |
| bond angles (deg) | 2.07 | 1.67 |
| Ramachandran plot (%) | | |
| most favored | | |
| allowed | | |
| outliers | | |

[a]Precision-indicating merging R
[b]Pearson correlation coefficient of two "half"data sets
[c]Rwork = Σ|Fobs − Fcalc|/ΣFobs
[d]Five percent of the reflection data were selected at random as a test set, and only these data were used to calculate Rfree
[e]Root-mean square deviation
[h]Not applicable

To determine at which step the DNA-binding domain of YczR was lost, mass spectrometry is used to determine the molecular weight of YczR monomer (Figure 16) immediately after FPLC purification. YczR monomer has 474 amino acids, which theoretically molecular weight is 52,112 Da. In the absence of PLP, the Mass Spectrometry plot shows a predominant peak at 52933 Da, which suggesting a full-length of YczR in solution after FPLC purification. We speculate that protein lost N-terminal winged-helix domain during crystallization step.

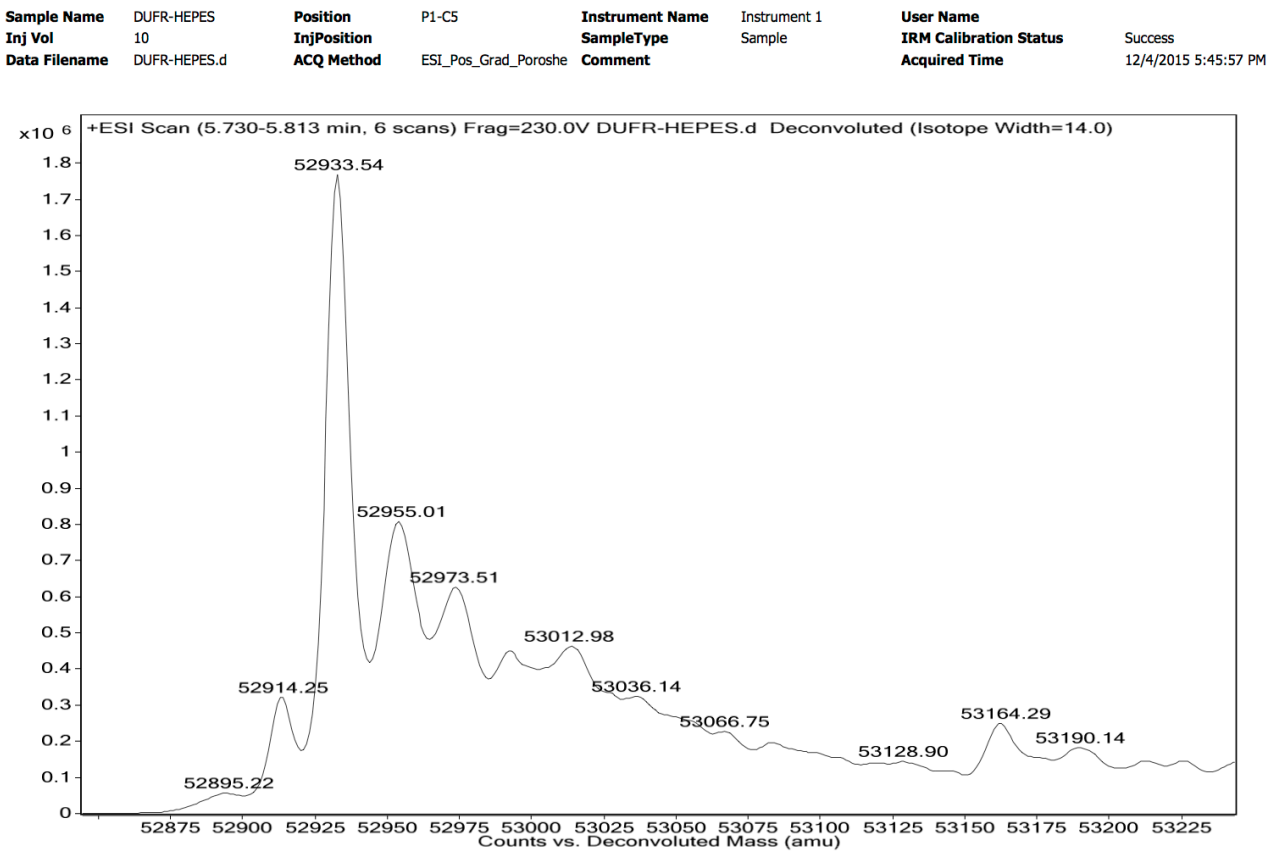



Figure 10. Mass spectrometry data of YczR-PLP monomer.

To determine if YczR is a dimer in solution, SEC-MALS was used to assay YczR oligomers (Figure 17). Purified YczR was eluted from the size-exclusion column as a single peak detected

by Abs280. The molecular weight of 121 ± 5 kDa for this peak is determined by light scattering, which matches the mass calculated for the dimer from the expected sequence (52112 Da × 2 = 104224 Da). YczR came out to be pretty homogeneous with small levels of aggregates that eluted right before the main peak and hence, may lead to an overestimation of the molecular weight of the main peak. Therefore, speculate YczR even at very low protein concentration still forms a dimer that appears a little bigger in solution. SEC-MALS data do not show YczR tetramers or higher oligomers exist in solution; however, further YczR oligomerization may still occur when the protein binds to DNA.

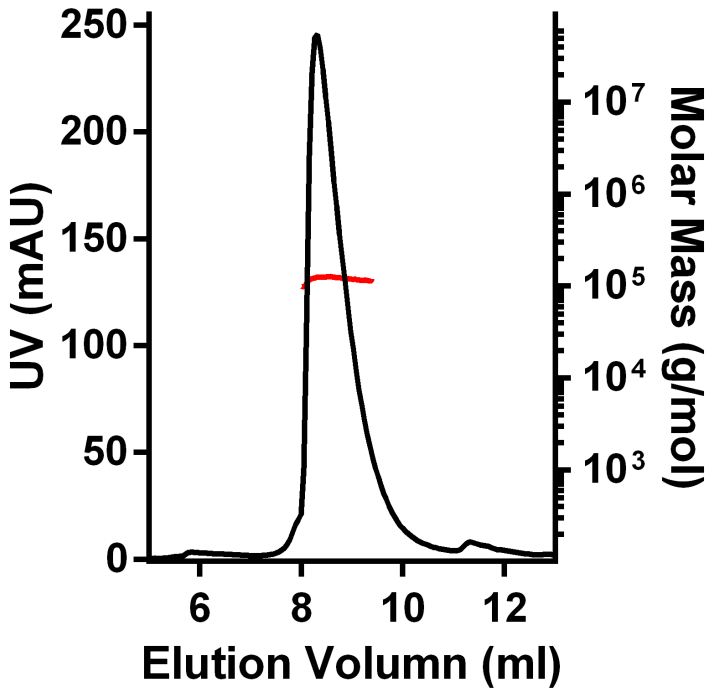



Figure 11. Molecular weight determinations of YczR in solution. SEC-MALS profiles are shown for YczR. The absorbance at 280 nm is shown as a black line, and the calculated molecular weight for protein in the eluent at a particular time is shown in color, as noted. YczR (molecular weight in red) gives a single peak with a molecular weight of 121 ± 5 kDa.

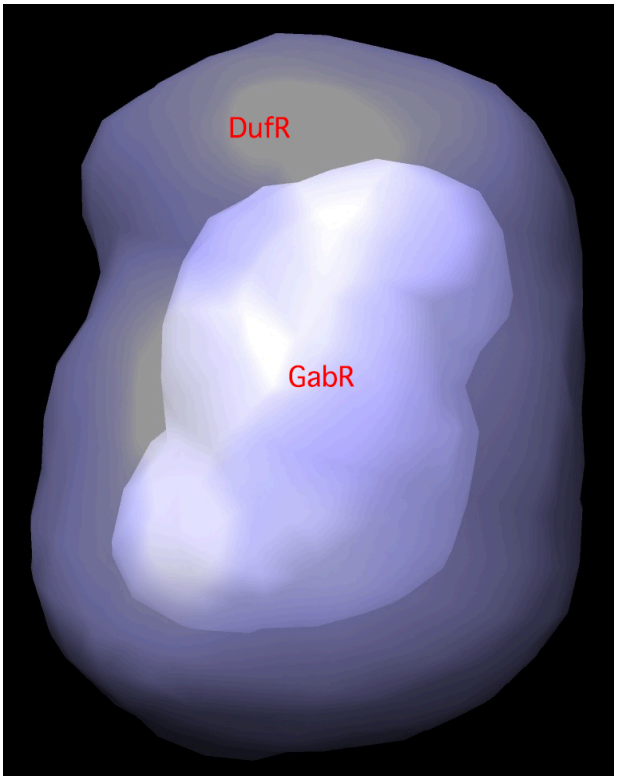



Figure 12. The overall comparison of *K. pneumoniae* YczR and *B. subtilis* GabR (PDB ID: 4N0B) SAXS models.

Both YczR and GabR belong the MocR/GabR subfamily of GntR family of transcription regulator[36]. The SAXS model of *K. pneumoniae* YczR and *B. subtilis* GabR (PDB ID: 4NOB) (Figure 18) were compared to analyze the difference in overall size. Based on the SAXS model comparison result, the volume of YczR dimmer in solution was bigger than GabR. The result suggested that YczR was containing the much extensional DNA-binding domain. The whole structure of YczR should be more extensional.

# CHAPTER THREE

# YCZR REGULATION DOMAIN CATALYZES A REVERSIBLE HALF-TRANSAMINASE REACTION INSTEAD OF BEING A NON-ENZYMATIC REGULATOR LIKE GABR

The crystal structure of YczR showed that PLP was bound to the regulator domain reacted with the lysine (Lys-311) to form a stable internal aldimine, which likely controls the transcription from the *yczE* promoter in *K. pneumoniae.*

Alanine-glyoxylate aminotransferase catalyzes the transfer of the amine of L-alanine to glyoxylate, forming glycine and pyruvate. In the first half-transamination reaction, L-alanine reacts with the pyridoxal form (PLP) of the enzyme to yield the pyruvate and the pyridoxamine form (PMP) of the enzyme[56-60]. Based on comparison of the crystal structure of YczR with alanine-glyoxylate transaminase structure, we speculated that YczR possibly possess half of the alanine-glyoxylate transaminase activity. To test this speculation, we designed the UV-absorption and fluorescence spectroscopic experiments, and end-point enzymatic assays.

## UV Absorption

To test potential YczR substrates, we designed UV-Vis spectroscopy experiments in which 20 amino acids and other possible substrate were added to the protein.

The UV absorption of YczR alone in the reaction buffer (50 mM HEPES, pH 7.5, 300 mM NaCl) has two peaks maximum at 320 nm and 425 nm (Figure 19). The absorption curve stayed unchanged in 20 mins measurement. The 330nm and 420 nm peaks of the Schiff base (aldimine)

formed between an active-site lysine residue of the protein and PLP (Figure. 19, A) are generally ascribed to a ketoenamine tautomer, enolimine tautomers, a substituted aldamine respectively or PMP[64-68]. The peak at 420 nm signifies the formation of internal aldimine between an active-site lysine residue of the protein and PLP.

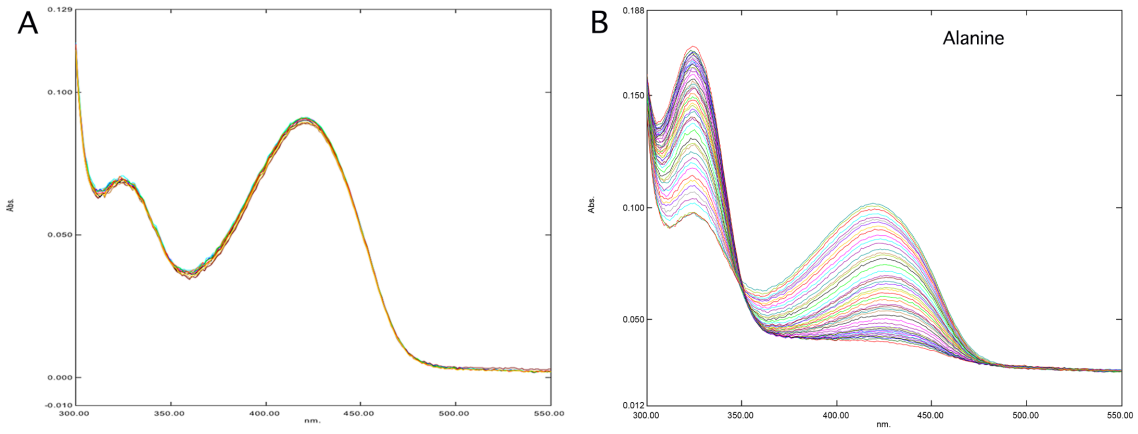



Figure 1. Time-resolved UV absorption spectra of YczR. 6 µM YczR in buffer (50 mM HEPES, pH 7.5, 300 mM NaCl) containing 6 µM PLP was taken at 25°C in the absence (A) or presence (B) of 500 mM L-alanine. After the addition of L-alanine, UV spectra were taken for 192 minutes at 2 mins intervals.

Upon mixing the YczR with 500 mM L-alanine, a biphasic spectral change was observed as shown (Figures 19, B). In a time-dependent manner, a slow decrease in the 425 nm peak with a concomitant increase in the absorbance at 325 nm was observed within 192 minutes.

Decrease of the absorption peak at 425 nm and increased that at 330 nm is consistent with PLP turning into PMP. We speculated that L-alanine reacted with PLP to form one of the intermediates of the half-transamination.

We speculate that the products of PLP and L-alanine reaction were PMP and pyruvate. UV

absorption spectra of YczR-PLP showed an increase in absorption in the range of 300-320 nm in reaction with L-alanine (Figure 19, B), suggesting the formation of PMP and possibly pyruvate, which both has absorpt at 320 nm.

To prove one of the products was pyruvate, we added 100 mM pyruvate into the reaction cuvette after YczR and L-alanine reaction completed (Figure 20, A). Upon addition of pyruvate, the maximum absorption at 320 nm decreased (Figure 20, B). Moreover, maximum absorption at 425 nm increased at the same time (Figure 20, C). The absorption change suggested that pyruvate was a product of PLP and L-alanine reaction. High concentration of pyruvate prompted the product from the first half reaction (PMP) reacted with pyruvate producing L-alanine. The increase in the UV absorption at 425 nm was observed during the addition of pyruvate, which confirms the formation of the Schiff base (aldimine) between PLP and an active-site lysine residue of the YczR. It is likely that that PLP-L-alanine reaction is a reversible reaction. Once the concentration of the product is increased, it will promote PLP formation.

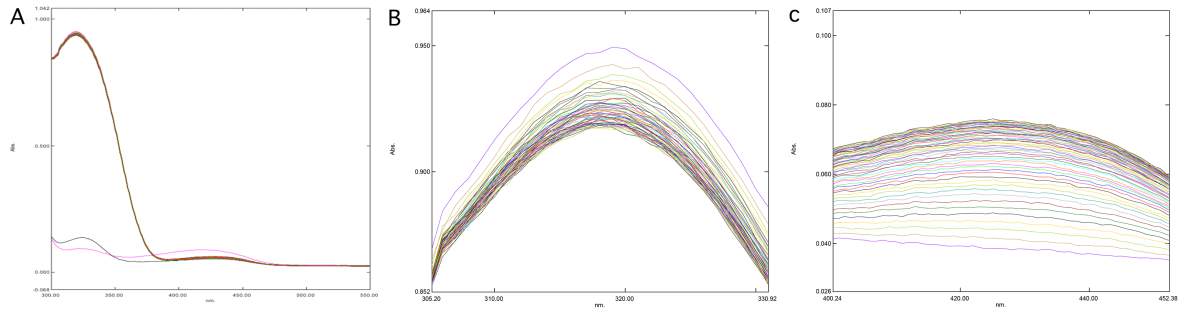



Figure 2. UV absorption spectra of YczR-L-alanine with pyruvate. Reaction of 6 uM YczR and 30 mM L-alanine should be completed in 192 mins. 100 mM pyruvate was added into the cuvette. UV absorption was measured for 2 hours. Figures shown are for UV absorption change in 300 nm-600 nm (A), 306nm-330nm (B) and 400-450 nm (C). Brown line indicates absorption before addition of L-alanine. Pink line indicates that after PLP L-alanine reaction was complete.

**Specificity test of YczR with amino acids or small molecular**

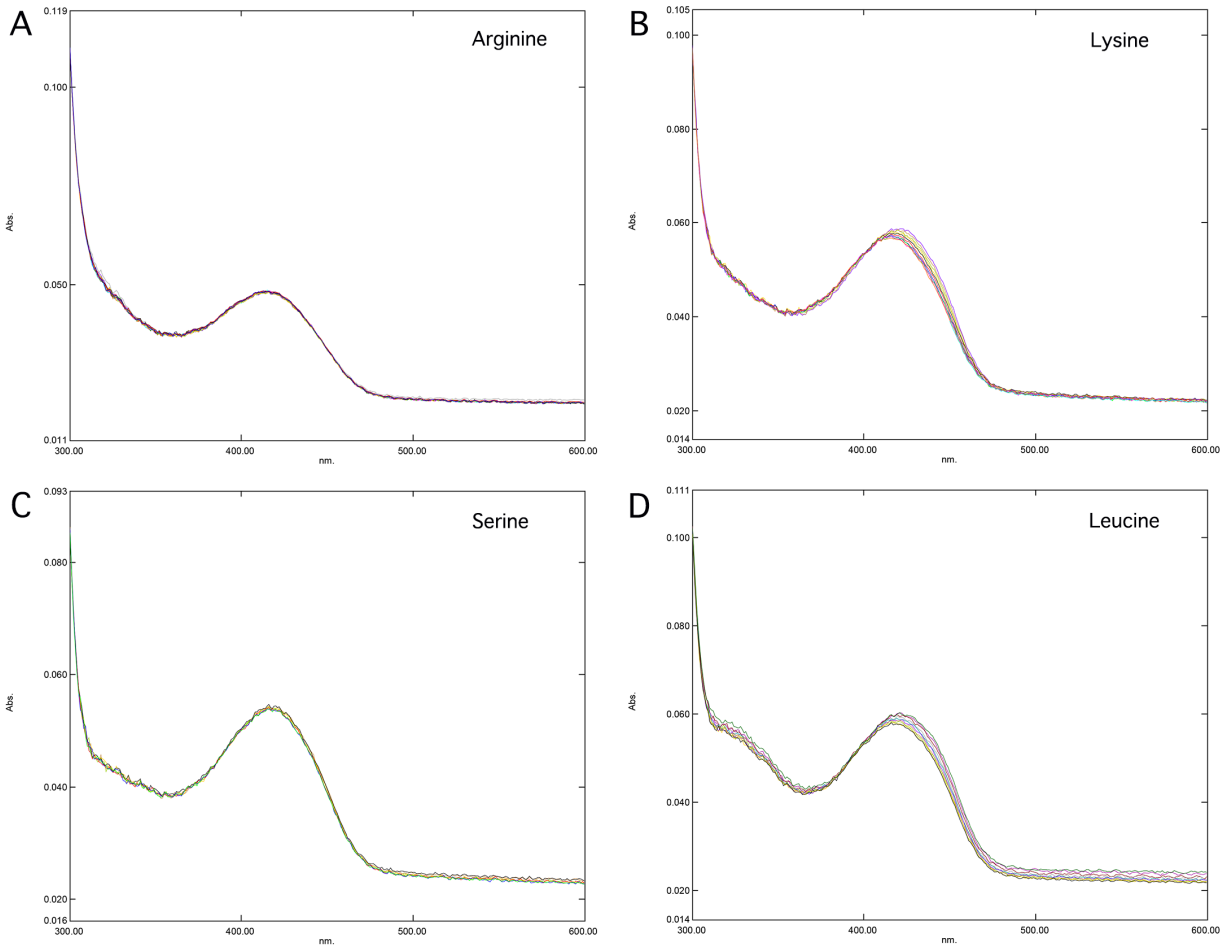



Figure 3. Specificity test of YczR with amino acids or other small molecules. Shown are UV absorption changes of 7 μM YczR in the presence of 300 mM Arginine (A), Lysine (B), Serine (C) and Leucine (D). In each case, the buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl.

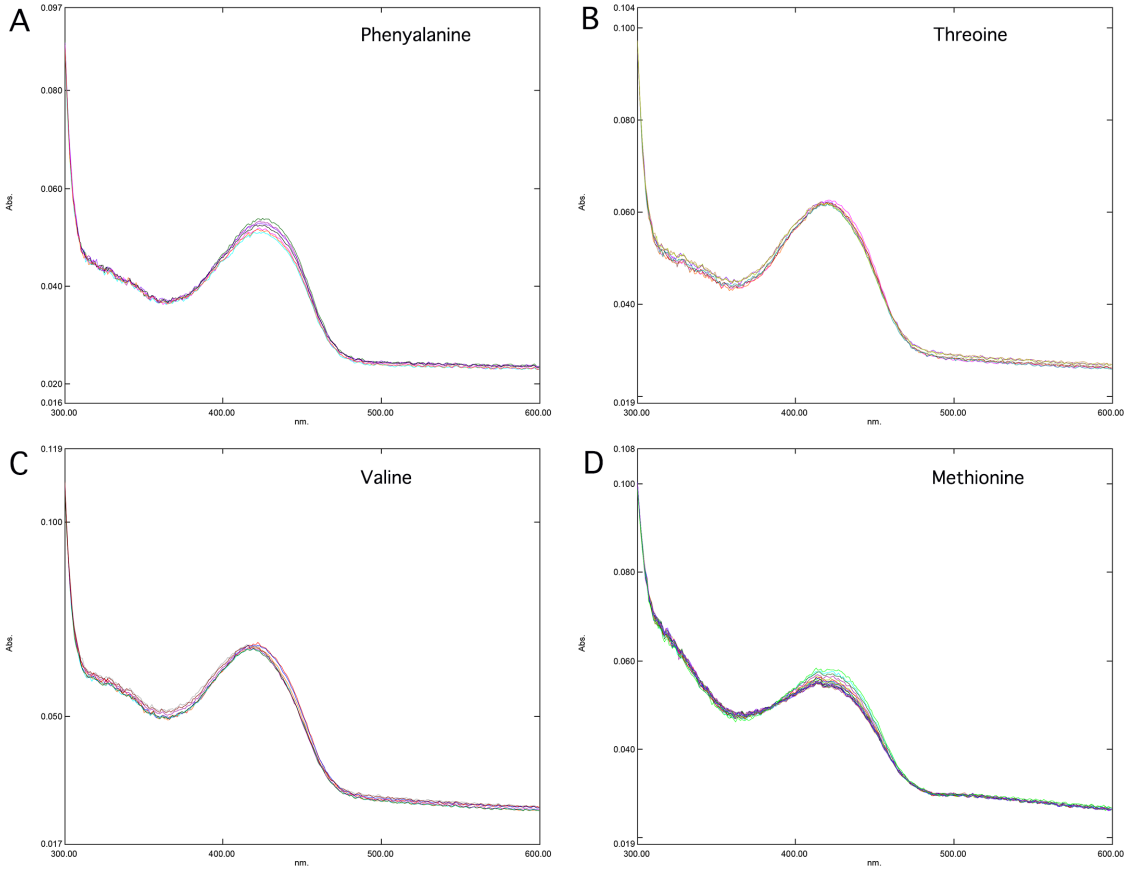



Figure 4. Specificity test of YczR with amino acids or other small molecules. Shown are UV absorption changes of 7 μM YczR in the presence of 300 mM phenyalanine (A), threoine(B), valine (C) and methionine (D). In each case, the buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl.

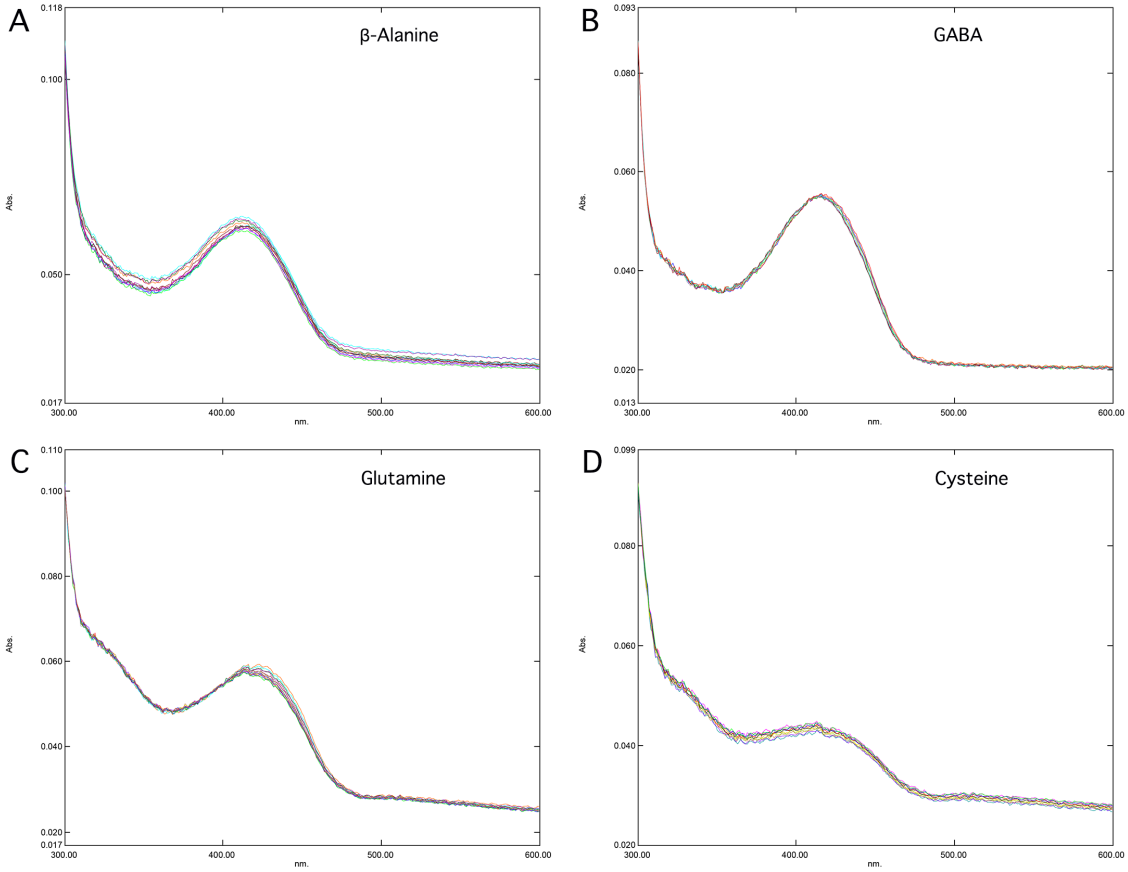



Figure 5. Specificity test of YczR with amino acids or other small molecules. Shown are UV absorption changes of 7 μM YczR in the presence of 300 mM β-Alanine (A), GABA (B), Glutamine (C) and Cysteine (D). In each case, the buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl.

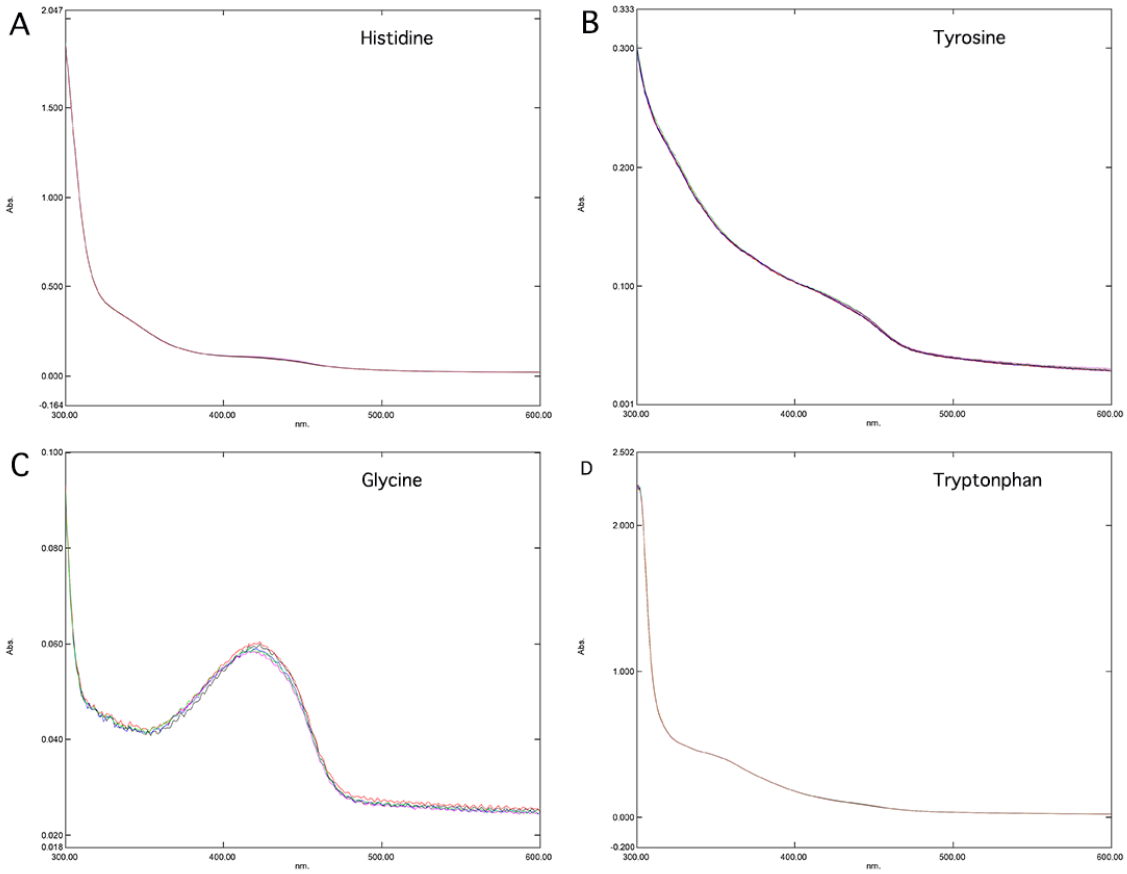



Figure 6. Specificity test of YczR with amino acids or other small molecules. Shown are UV absorption changes of 7 μM YczR in the presence of 300 mM histidine (A), tyrosine (B), glycine (C) and tryptophan (D). In each case, the buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl.

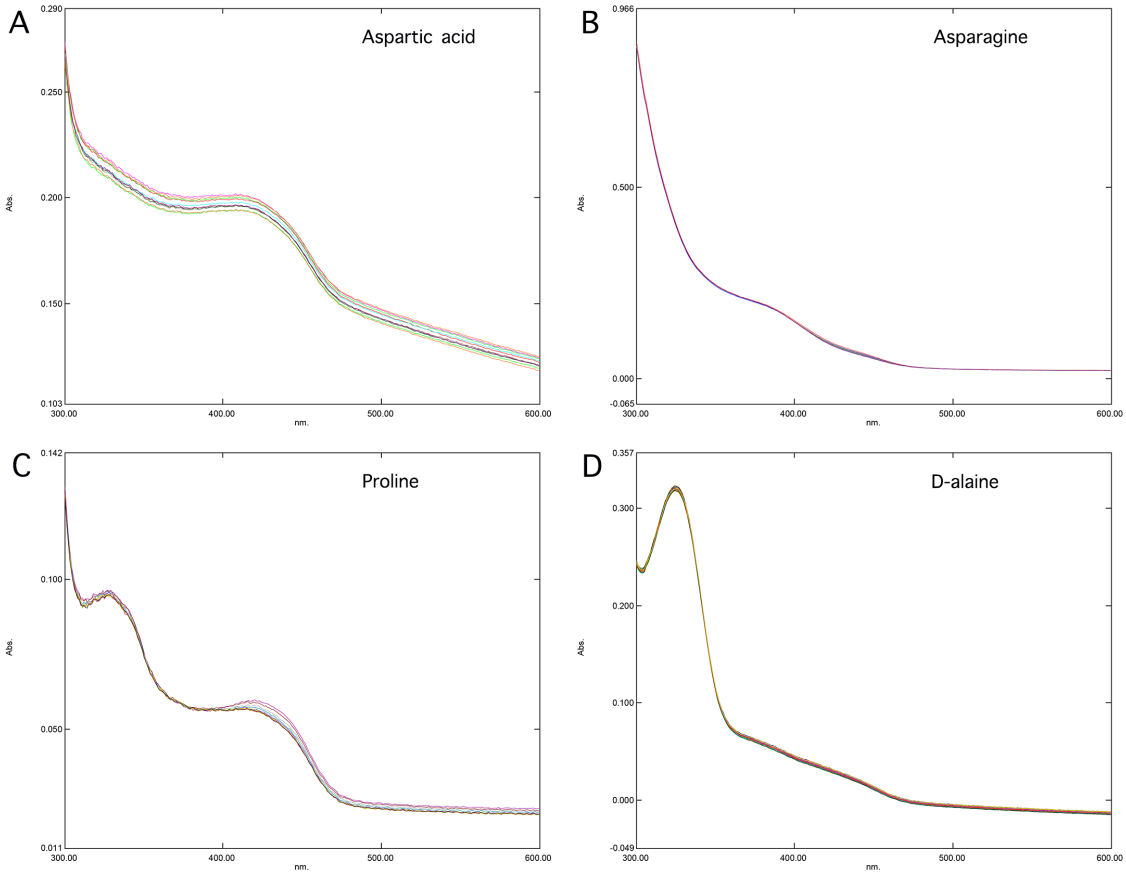



Figure 7. Specificity test of YczR with amino acids or other small molecules. Shown are UV absorption changes of 7 μM YczR in the presence of 300 mM aspartic acid (A), asparagine (B), proline (C) and D-alanine (D). In each case, the buffer was 50 mM HEPES, pH 7.5, 300 mM NaCl.

The UV-visible spectrum of YczR reacted with other amino acids, and possible substrates were recorded (Figure 21-25). The time-dependent UV absorption of 7 μM YczR-PLP with 300 mM B-Alanine, serine, arginine, lysine, tyrosine, histidine, aspartic acid, phenylalanine, leucine, proline, threonine, valine, GABA, tryptophan, asparagine, methionine, and glutamine was measured separately in a cuvette with a 1 cm path length. The UV abortion spectral changes that

occurred with the addition of L-alanine were not observed when other amino acids, such as the D-alanine, glycine, glutamine or asparagine, were added. The results suggest that YczR specifically reacted with L-alanine, which matches the purpose of YczR as it activates transcription of the *yczE* operon. L-alanine or molecules with similar structure may regulate YczR function though reaction with PLP.

**Fluorescence emission spectra**

More information about conformational changes of the Schiff base after addition of L-alanine can be obtained by fluorescence emission spectroscopy.

Time-dependent fluorescence emission spectra results of YczR with L-alanine is shown (Figure 26). YczR (6 μM) and L-alanine (500 mM) were mixed in buffer (50 mM HEPES, pH 7.5, 300 mM NaCl) at 25 °C. After the addition of L-alanine, fluorescence spectra was taken for 20 mins at 4 mins intervals. When L-alanine-reacted YczR was excited at 283 nm, the characteristic emission peaks of the Shiff base between PLP and lysine were observed at approximately 326 nm[69].

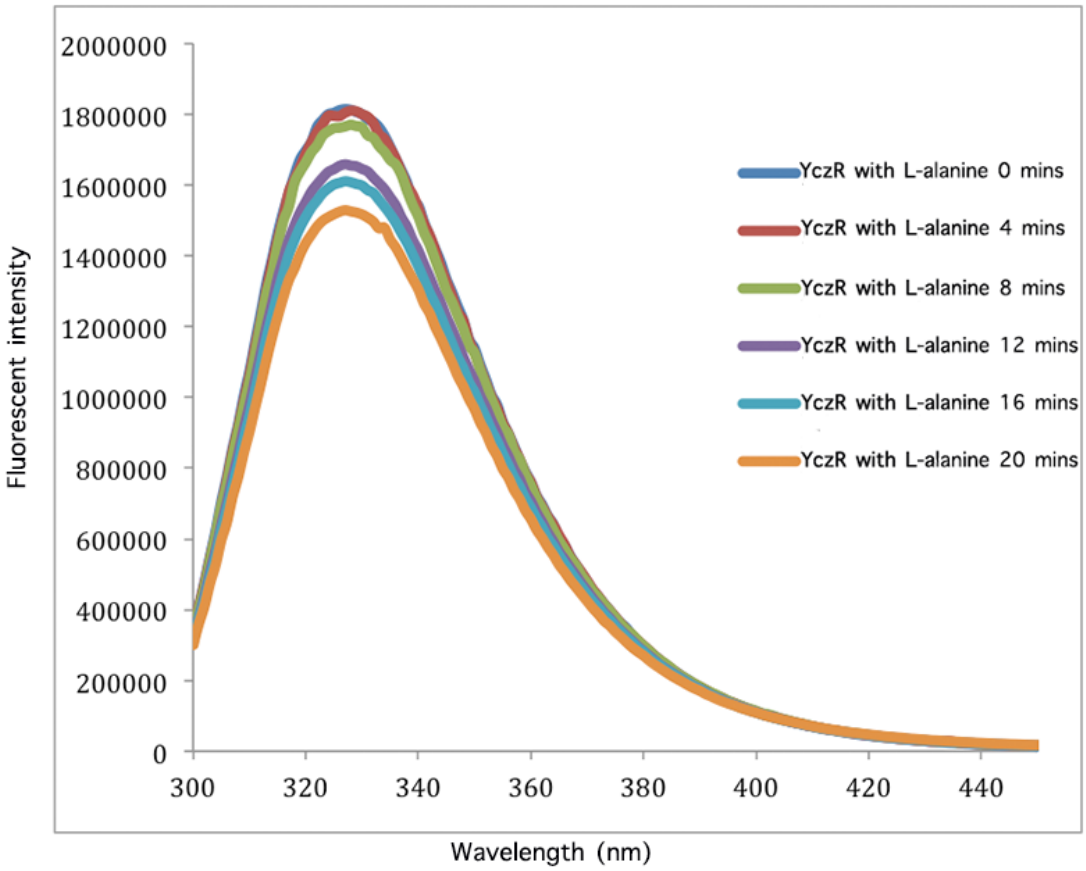



Figure 8. Fluorescence emmission spectra of YczR. Fluorescence spectra of YczR-PLP (1.4 μM) in the buffer (50 mM HEPES, pH 7.5, 300 mM NaCl) was taken at 25°C in the absence of 500 mM L-alanine at an excitation wavelength of 283 nm.

Upon mixing the YczR with L-alanine, a change in fluorescence spectra was observed. The peak in the absorbance at 326 nm decreased slowly. The 326 nm absorption peak of the L-alanine-reacted YczR can be ascribed to the external Schiff base associated with L-alanine. These results suggest that the Shiff base formed between PLP in YczR was broken in the presence of L-alanine.

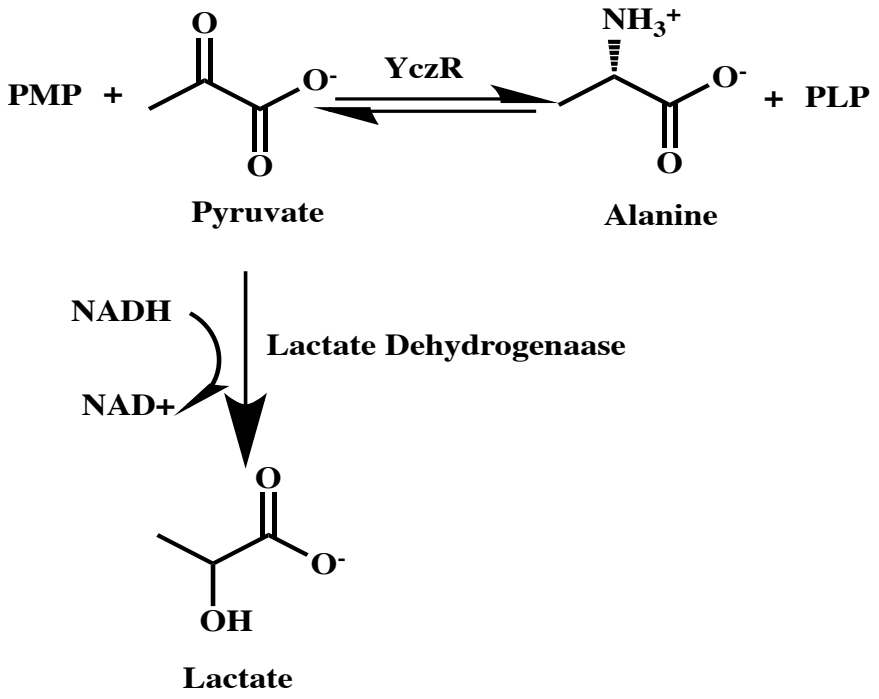
The fluorescence spectral changes that occurred with the addition of L-alanine were not

observed when other amino acids, such as the D-alanine, glycine, glutamine or asparagine, were added.

Based on the results of Fluorescence emission spectra and UV absorption spectra, we speculate that the internal aldimine of YczR can react with L-alanine specifically.

**An End-point Assay to Monitor the YczR Catalyzed Half Transaminase Reaction**

Based on the UV absorption spectra result, we speculate that L-alanine formed pyruvate after reaction with PLP. An End-point assay was designed to confirm this hypothesis (Scheme 1). YczR was incubated with L-alanine excess and coupled with a lactate dehydrogenase to monitor the consumption of the NADH at 340 nm.




Scheme 1. The end-point YczR-catalyzed half-aminotransferase reaction. The first half reaction (L-alanine to pyruvate) of the YczR-catalyzed transamination is shown on the top. The second half-reaction of Lactate dehydrogenase-catalyzed NADH oxidization is shown at the bottom.

To let the L-alanine and PLP reaction to be completed, 1 μM, 3.5 μM and 7 μM YczR had been added in the reaction buffer containing 500 mM L-alanine, 50 mM HEPES, (pH7.5), 300 mM NaCl, which was kept at room temperature for 15mins. The lactate dehydrogenase consumed all pyruvate in each reaction solution to produce lactate. The data above shows measured absorbance change at 340 nm from the concomitant oxidation of the cofactor, NADH to $NAD^+$.

The concentrations of protein should affect the starting absorption points. The higher concentration of YczR in reaction solution had a higher starting absorption point. The absorption of 1 μM, 3.5 μM and 7 μM YczR decreased by 0.01, 0.05 and 0.1, absorption units (Figure 27). Those results suggest that 1 μM PLP in the YczR solution could react with 1 μM L-alanine to produce 1 μM pyruvate. In the YczR-catalyzed first half aminotransferase reaction, PLP and PMP are in dynamic equilibrium when pyruvate is present. All of the PLP was converted to PMP when pyruvate was omitted.

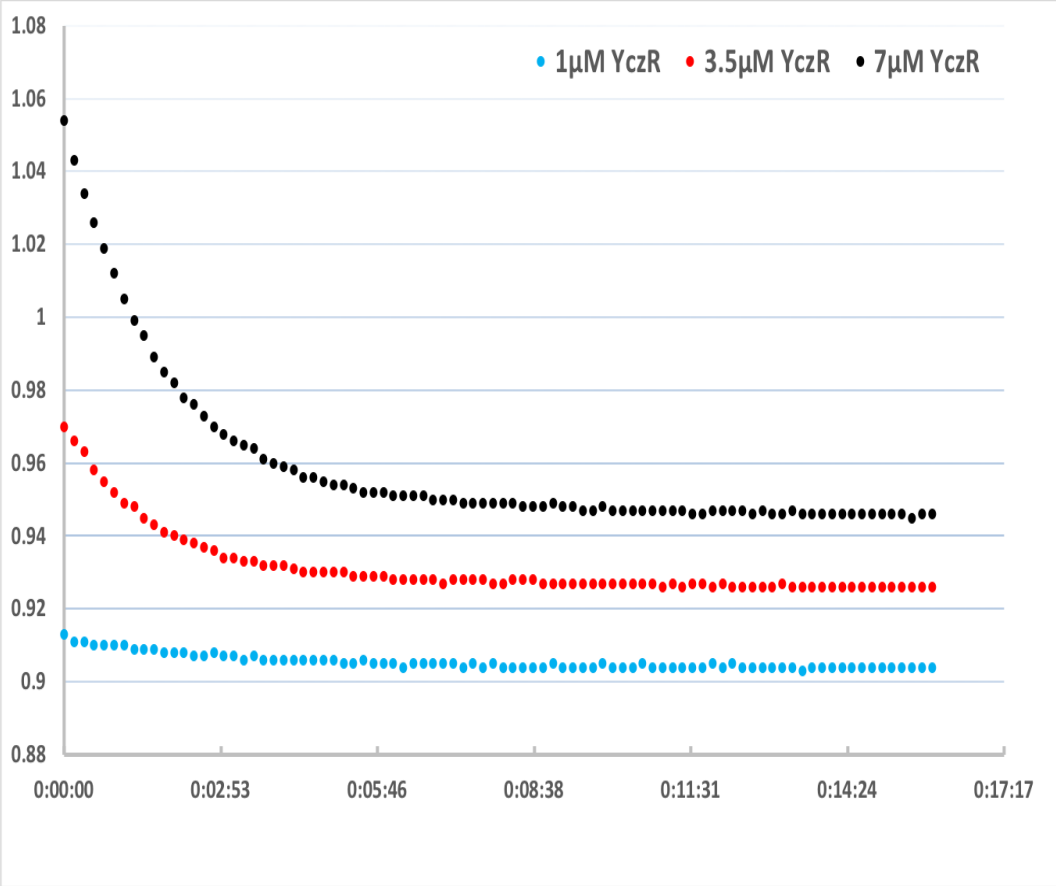



Figure 9. An end-point assay to monitor catalysis of half transaminase reaction by YczR. Absorption changes of NADH in the presence of 1 μM (blue), 3.5 μM (red) and 7 μM (black) YczR. 1 μM, 3.5 μM and 7 μM YczR has been added in the reaction buffer containing 500 mM L-alanine, 50 mM HEPES, (pH7.5), 300 mM NaCl, which was kept at room temperature 12 hours. The lactate dehydrogenase consumed all pyruvate in each reaction solution to produce lactate. Absorbance change at 340 nm from the concomitant reduction of the cofactor, NADH to $NAD^+$ was measured.

# CHAPTER FOUR

# YCZR BINDS DNA SEQUENCE BETWEEN YCZR GENE AND YCZE GENE IN A SIMILAR FASHION TO GABR

## YczR Binds DNA with High Affinity

YczR DNA binding sites and binding affinity was determined via fluorescence polarization (FP) experiments[70]. The *K. pneumoniae* YczR regulatory region contains 88 bp and has two 5bp direct repeats (CCTCA & AATGG).

5'ATTTTTTCCTCAACATCATTGCGAATGCCCTCATCATCAGCAAAAATGGACTG ATAATTAATGGCCAGTTTTGGAAAAGTGGATCGCT3' within this fragment, the two direct repeats of sequences CCTCA and AATGG, appear to be likely for YczR binding. The bindings sites between CCTCA and AATGG repeats are separated by 34 bp, which corresponds to approximately three helical turns of DNA. Therefore, the corresponding nucleotides of the binding sites should lie on the same face of the DNA helix.

GabR protected a 49 bp region of DNA identified previously for GabR-dependent regulation *in vivo*[35, 36]. Sequence alignment results indicate that many residues of the DNA binding regions between GabR and YczR are conserved (Figure 28). The GabR and YczR DNA binding regions are similar in type, direction, and spacing. We predict that some protein-DNA interactions between GabR-DNA and YczR-DNA complexes will be similar. Although a multitude of GntR family DNA bindings sites has been determined, the studies did not indicate a clear consensus sequence, rather sparse alignment was detected over short fragments [54, 55, 71].

Based on the pair of two direct sequence repeats, two 68-bp complementary single-stranded DNA oligonucleotides containing the YczR binding region (5'TTTCCTCAACATCATTGCGAATGCCCTCATCATCAGCAAAAATGGACTGATAATTA ATGGCCAGTTTT-TAMRA-3', where Fl means fluorophore) were purchased from Integrated DNA Technologies (IDT). Fluorescence polarization assay studies were conducted to estimate 1nM as the optimal DNA concentration for further experimentation.

```
CLUSTAL O(1.2.1) multiple sequence alignment

gb|CP015004.1|:441329-441434      TTTGTAGTTTCTCCTTCTGATACCATCAAAAAGTTATAATTGGTACTTTTCATCATACCA
gb|CP013711.1|:73923-74010        -----ATTTTTTCCT-CA----ACATCATTGC------GAATGCCCTCATCATCAGCA--
                                       * *** **** *:    .*****::..      .:: * .** :****** ..

gb|CP015004.1|:441329-441434      AAGAGAAGTCAGAATGATAAGAAAATACCGACAGGGG----GATATTCAC-
gb|CP013711.1|:73923-74010        -AAAATGGACT----GATAATTAATGGCCAGTTTTGGAAAAGTGGATCGCT
                                   *.*.:.*:*:    ***** :**: .**.. :  **    *: .:**.*
```

Figure 1. Sequence alignment of the winged-helix domain sequences of *B. subtilis* GabR gabR-gabT DNA binding regions (top) and *K. pneumoniae* YczR yceE DNA binding regions (bottom). The black asterisks highlight DNA-binding residues of YczR that are identical or similar in GabR.

The binding affinity of YczR for DNA containing the 68-bp YczR binding site was determined by FP (Figure 29). Different concentrations of YczR were incubated with TAMRA-labeled DNA fragments containing the promoter regions of yczE and assayed for the formation of protein-DNA complexes. The result showed a dose-dependent polarization increase, indicative of YczR binding to DNA. The fluorescence polarization data set was fitted into the modified Hill equation to determine the apparent dissociation constant of the YczR-DNA complex, as 26.47 nM.

To rule out nonspecific effects from the TAMRA label and establish binding specificity, a

negative control assay was performed with Bovine Serum Albumin (BSA) and the TAMRA-labeled DNA (Figure 30). A constant concentration of TAMRA-labeled DNA was incubated in several aliquots with increasing concentrations of BSA protein. The fluorescence polarization did not show any significant change with increasing concentration of BSA, which suggests that YczR specifically binds to the TAMRA-labeled DNA.

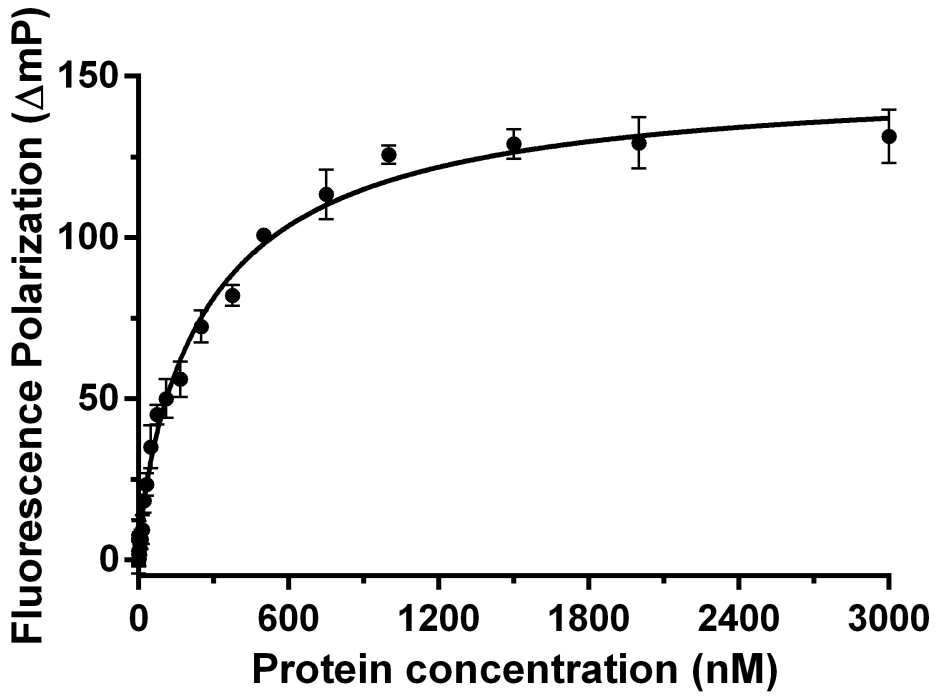



Figure 2. Fluorescence polarization assays to measure DNA binding by YczR. Titration of YczR against 1 nM TAMRA-DNA. The YczR concentrations were varied from 0 to 3,000 nM. The average values and errors were plotted and fitted to the Hill equation. Each data point was measured in triplicate. Error bars are too small to be displayed if the errors are less than 1/1,000th of the measured value.

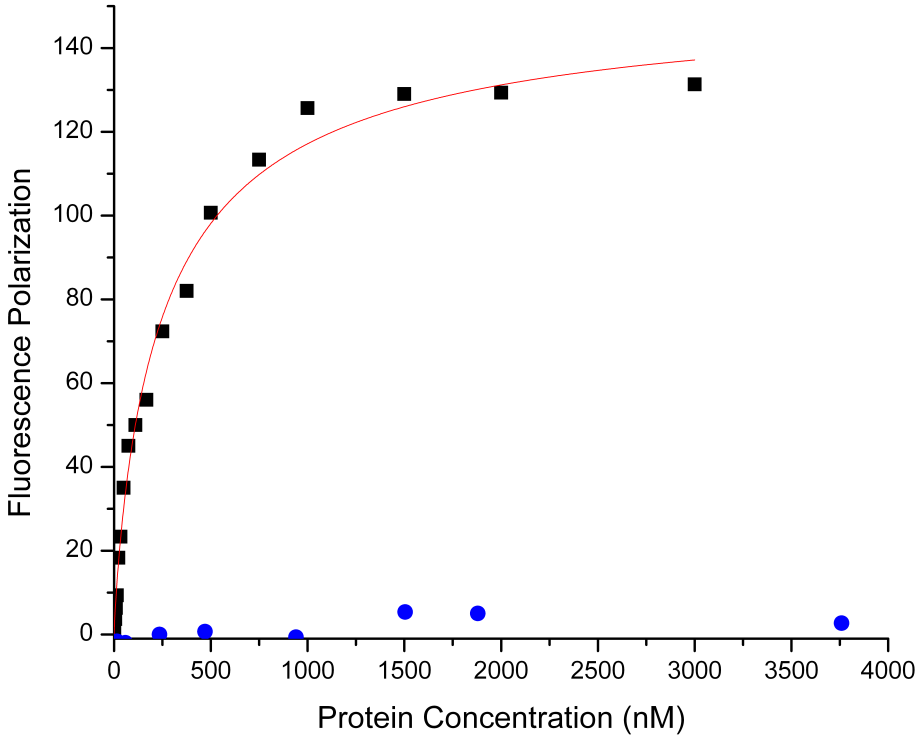



Figure 3. Specificity test of YczR fluorescence polarization assays. In the negative control binding assay, the same increasing concentrations BSA were used against a fixed concentration of TAMRA-labeled DNA to indicate specific binding.

**YczR likely binds to the regulation region with a 2:1 stoichiometry**

Based on the CCTCA and AATGG direct repeats on the DNA sequence, two 34-bp complementary single-stranded DNA oligonucleotides containing the YczR binding region DNA_1 (5'- TTTCCTCAACATCATTGCGAATGCCCTCATCATTC-TAMRA-3') and DNA_2 (5'- AGCAAAAATGGACTGATAATTAATGGCCAGTTTT-TAMRA-3') were designed and purchased from IDT. The two 34 bp DNA fragment located within the YczR protected region is close in size to a minimal fragment allowing YczR binding *in vitro*.

**YczR binds tighter to the site closer to yczE gene**

DNA_1 has one specific binding site for YczR (Figure 31). The dissociation constant (Kd) of the YczR with DNA_1 complex was determined to be 135.1 nM. DNA_2 also has one specific binding site for YczR (Figure 32). The dissociation constant (Kd) of the YczR with DNA_2 complex was determined to be 33.63 nM. We speculate that YczR binds on the DNA_2 binding site first at low concentration and then it binds the DNA_1 binding site at high concentration.

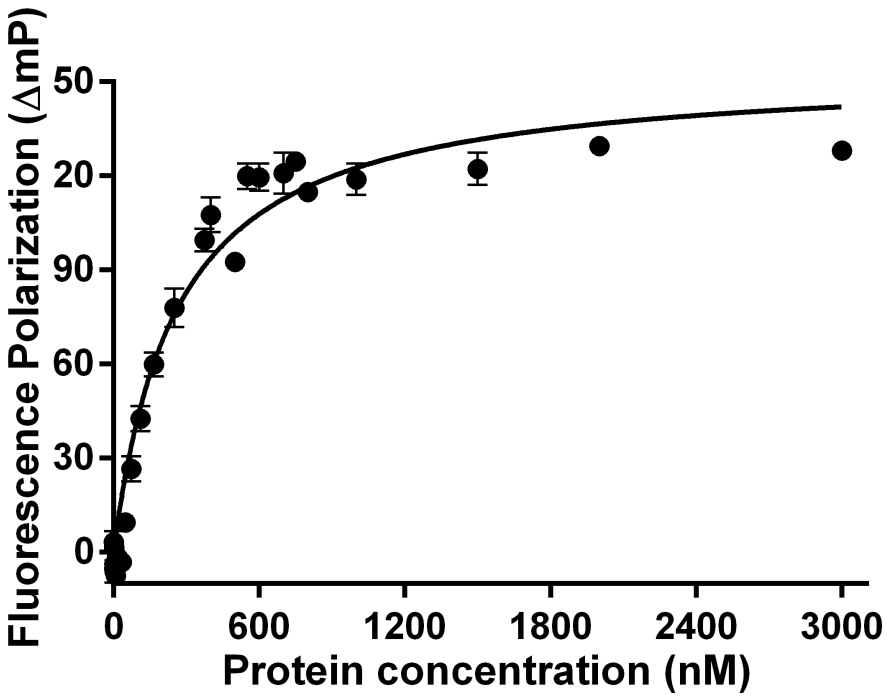



Figure 4. Fluorescence polarization assays to measure DNA_1 binding by YczR. Titration of YczR against 1 nM TAMRA-DNA. The YczR concentrations were varied from 0 to 3,000 nM. The average values and errors were plotted and fitted to the Hill equation. Each data point was measured in triplicate. Error bars are too small to be displayed if the errors are less than 1/1,000th of the measured value.

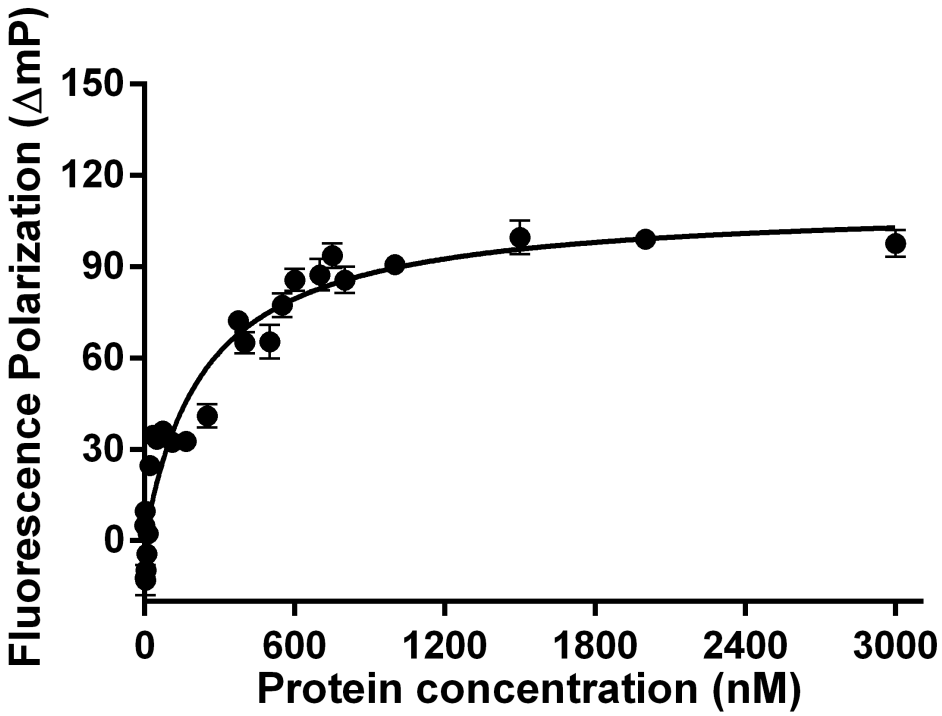



Figure 5. Fluorescence polarization assays to measure DNA_2 binding by YczR. Titration of YczR against 1 nM TAMRA-DNA. The YczR concentrations were varied from 0 to 3,000 nM. The average values and errors were plotted and fitted to the Hill equation. Each data point was measured in triplicate. Error bars are too small to be displayed if the errors are less than 1/1,000th of the measured value.

# Materials and Methods

## Materials

Chemicals for YczR purification, crystallization, assays were purchased from Sigma-Aldrich (St. Louis, MO), Alfa Aesar (Tewksbury, MA) and Fisher Scientific (Pittsburgh, PA). YczR gene was synthesized with codon optimization for *E. coli* expression from Integrated DNA Technologies (Coralville, IA).

*E. coli* BL21 (DE3) cells and Turbo cells were purchased from New England BioLabs (Ipswich, MA). Bacterial growth media Luria-Bertani (LB) broth and antibiotics were obtained from Sigma-Aldrich. Crystallization screen solutions and other crystallization supplies were purchased from Hampton Research (Aliso Viejo, CA) and Emerald Bio (Bedford, MA). All other materials were purchased at the highest quality available. DNA and protein sequences were analyzed using DNA Strider and the BLAST program[72, 73].

## Methods

**DNA Cloning.** To construct the His-tagged version of the YczR protein, the coding region was amplified by PCR from the synthesized YczR gene for K. pneumoniae genomic DNA using the specific oligonucleotides as primers (forward primer designed as GAAGGAGATATACATATGATCGACCAGAGCGATTGGATA and the reverse primer designed as GTGATGGTGGTGATGATGATCCCCTGTAACGGGGATTTT). The PCR thermocycler program used included an initial denaturation for 120 s at 95 °C; 40 cycles of 95 °C for 30 s, 50 °C for 60 s, and 72 °C for 50 s; and the final extension of 72 °C for 5 min. The PCR product was digested with NdeI and NotI and subsequently cloned into the expression vector

pETite (Lucigen).

The YczR DNA fragment was cloned into the NdeI site of pETite by treated with Phusion HF DNA polymerase, following the manufacturer's instructions. The pETite-kan-c-5His vector with *K. pneumoniae* YczR full-length gene was transformed insert into *E. coli* Turbo competent Cell (New England Biolabs).

After pETite-kan-c-5His vector was purified from *E. coli* Turbo cell, the resulting plasmid was then transferred to competent *E. coli* BL21DE3 cells to overexpress the YczR protein. Digestion of the plasmid from *E. coli* Turbo cell was by DNA Restriction enzymes NdeI and NotI. The restricted DNA fragments were analyzed by agarose gel electrophoresis according to standard protocols.

The purified plasmid from BL21DE3 cell was then sent to sequencing. DNA sequencing result from DNA Sequencing & Genotyping Facility (University of Chicago Comprehensive Cancer Center) confirmed the correct orientation with a C-terminal His6-tag.

**Protein Purification.** The YczR protein from E. coli was expressed and purified by using the intein-mediated purification with an affinity histidine tag system. Transformed E. coli BL21DE3 cultures were grown overnight at 37°C and 250 rpm in 500 ml of LB medium containing 50 μg/ml kanamycin.

Induction of YczR expression was carried out at OD600 of 0.8 by adding Isopropyl-β-D-thiogalactopyranoside (IPTG) to the final concentration of 0.5 mM.

Cells were incubated at 25 °C and harvested after 16 hours by centrifugation at 7000 RPM and 4 °C for about 15 min (Avanti J-E Centrifuge, Beckman Coulter). The cell pellet was

resuspended in Ni column buffer containing 5% (vol/vol) glycerol and 1 mM PLP.

Following sonication (Qsonica sonicators, Q500 Sonicator) to break the cellular wall, the cell extract was cleared by centrifuged twice at 16,500 rpm and 4 °C for 20 min. The supernatant fraction was loaded onto a 10 ml affinity chromatography Ni-NTA column (GE Healthcare) containing $Ni^{2+}$ and previously equilibrated with lysis buffer. Elution of the retained proteins was achieved with a linear imidazole gradient (20 column volume, 10 mM-400 mM).

Fractions containing protein eluted from Ni column were concentrated (Centricon Plus centrifugal filter units; Millipore) to 2 mL, and loaded onto the gel filtration column (HiLoad 16/60 Superdex 200 pg. GE Healthcare), which was equilibrated with gel filtration column buffer (300 mM NaCl, 50mM HEPES, pH 7.5).

The YzcR concentration was determined with the Bio-Rad protein assay kit. The protein molecular weight was analyzed using Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE). Purified YczR protein was used freshly after exchanging into suitable buffers for the crystallization and biological assays. The YczR protein was stored with high salt concentrations (300 mM Nacl) and 5 % imidazole to stay soluble without precipitation. The protein was stored at 4 °C for short-term or flash frozen in1000µL aliquots in liquid nitrogen after the addition of 20% (V/V) glycerol and stored at -80°C for long-term use. The color of concentrated, purified YczR was pale yellow because PLP was bound to the protein.

**Crystallization.** The purified YczR was buffer exchanged into crystallization screening buffer (50 mM HEPES, pH 7.5, 300 nM NaCl) and concentrated to 10 mg/ mL using a 10,0000 molecular weight cutoff (MWCO) Amicon-Ultra centrifugal filter device (Millipore). The initial

sparse matrix screen (Crystal Screen Cryo 1 and 2, Crystal Screen 1 and 2, Index 1 and 2, and PEG/Ion 1 and 2 from Hampton Research and Wizard 1&2 and 3&4 from Emerald BioSystems) were set up with the Gryphon crystallization robot (Art Robbins Instruments) to obtain initial crystallization hits.

Crystals appeared in the well solution containing 20% (W/V) PEG8000, 100 mM HEPES, pH 7.5, 200 mM $(NH_4)_2SO_4$, and 10% (V/V) isopropanol after a week during incubation at room temperature. After the initial crystallization screen, Crystals of YczR were obtained via the sitting drop methods incubated at room temperature. Further optimized manually using a 1:1 well solution: YczR stock solution (10 mg/mL) ratio to obtain crystals with the best size and morphology. Crystallization was conducted in 24-well Cryschem sitting drop plates (Hampton Research). Crystals grew to maximum size in 1-2 weeks and were harvested in about 15 days.

For attempted co-crystallization of YczR and L-alanine, 6 µL of YczR (20 mg/mL) was mixed with 1 µL of 10 mM L-alanine (dissolved in protein crystallization buffer) with 4 µL of well solution containing 8% (V/V) Tacsimate, pH 6.0, 20% (W/V) PEG8000. Crystals were seeded on day 3 with crystals obtained from previous trials. Crystals with good size and morphology were picked directly from the sitting drops into a cryoprotecting solution (well solution supplemented with 25% (v/v) glycerol) and then flash-cooled in liquid nitrogen.

**X-ray Diffraction Data Collection.** Individual YczR crystals were flash frozen in liquid nitrogen using 30% PEG10,000 in the crystallization buffer as a cryoprotectant solution. The diffraction data of YczR were collected at Advance Photon Source (Argonne National Laboratory, Argonne, IL) Structural Biology Center (SBC) 19ID and 19BM beamlines. The

YzcR crystal diffraction images were recorded by the 4×4 tiled 300×300 $mm^2$ sensitive area CCD detector from Area Detector Systems Corp. (ADSC). The data collection was at a wavelength of 1.0 Å. The datasets were indexed, integrated and scaled using the HKL2000[74].

**Structure Determination, Model Building, and Refinement.** The collected datasets were all scaled in the space group C2221 to a resolution of 1.79 Å. Auto build by molecular replacement using PHASER[75] in the PHENIX[76] software suit automatically built about 60% of the residues. Manually build the remainder of the model in COOT[77]. Rigid body refinement followed by restrained refinement was carried out in Refmac5 in the CCP4 software suits until no more side chains can be built in automatically[78]. Final refinement statistics of YczR are listed in Table 1. Structural figures were made using UCSF Chimera[79].

**UV-Absorption Spectroscopy.** UV-Vis absorption spectra were recorded with Jasco V-550 spectrophotometer in 1 mL quartz cuvettes in 50 mM Tris-HCl, pH 7.5, and 300 mM NaCl at 25 °C. The enzyme solution was centrifuged 5 mins at 14,000 RPM to reduce light scattering from the small amount of precipitate. The protein solution contained 10 mM potassium phosphate buffer at pH 7.5.

Absorption of YczR-PLP (7 μM) with L-alanine (50 mM) was observed in the UV range 300-600 nm over time. Also measure YczR-PLP (7 μM) at various with 300 mM B-Alanine, L-serine, Arginine, Lysine, Tyrosine, Histidine, Aspartic acid, Phenylalanine, Leucine, Proline, Threonine, Valine, GABA, Tryptophan, Asparagine, Methionine, and Glutamine separately as the same as YczR and L-alanine experiment.

**End-point Assay to Monitor the YczR Catalyzed Half Transaminase Reaction.** Activity was measured using varying concentrations of YczR protein in an assay buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl, 500 mM L-alanine, 320 μM NADH, and 5 units Lactate dehydrogenase (LDH).

In the first half-reaction, add 1 μM, 3 μM and 7 μM YczR to 500 μM L-alanine overnight at room temperature. In the coupled assay, measure the amount of pyruvate produced from each tube. LDH consumes the product of the first half-reaction (pyruvate) to produce lactate. The coupled reaction is measured spectrophotometrically (absorbance change at 340 nm) from the concomitant oxidation of the cofactor, NADH to $NAD^+$.

**Fluorescence emission spectroscopy.** Fluorescence emission spectroscopy is a type of electromagnetic spectroscopy, which analyzes fluorescence from a sample. Experiments were performed using a PTI Quanta-Master fluorimeter (Photon Technologies, Mirmingham, NJ). Time-based assays were conducted with excitation and emission wavelengths set at 283 and 326 nm, respectively. The measurement was initiated after 1.4 μM YczR-PLP and in the presence of 500 mM L-alanine in the buffer containing 50 mM HEPES (pH 7.5), 300 mM NaCl. Fluorescence emission spectra were measured every 4 mins from 0 min to 20 mins at 25 °C. GraphPad Prizm 3.0 was used for analyzing the time-dependent curves.

**Size Exclusion Chromatography with Multi-Angle Light Scattering (SEC-MALS).** The size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) was used to determine the molecular mass of YczR oligomers in solution. SEC-MALS system was set up

with the AKTA FPLC system (GE Healthcare Biosciences) with a silica-based size-exclusion chromatography column (WTC-030S5, Wyatt Technology) as a liquid chromatography unit.

For the experiments, the column was connected upstream of a refractive index detector (Optilab TrEX, Wyatt Technology), and then followed by a multiangle light scattering detector (Dawn Heleos II, Wyatt Technology) used for determining protein concentration and particle size, respectively.

SEC-MALS sample injection requires 100 µl of a solution containing 7 µM YczR protein in 50mM HEPES (pH 7.5), 300 NaCl buffer. The flow rate was set at 0.4 mL/min, and data were collected at 2 s intervals. Data processing and analysis were performed by the ASTRA software (Wyatt Technology).

**Fluorescence Polarization Assay.** The binding affinity of YczR for a DNA containing the 68-bp YczR binding site was determined by fluorescence polarization assays[70]. Synthesis Labeled DNA Sequence with fluorophore from Integrated DNA Technologies (IDT)

Two 68-bp complementary single-stranded DNA oligonucleotides containing the YczR binding region (5'- TTTCCTCAACATCATTGCGAATGCCCTCATCATCAGCAAAAATGG ACTGATAATTAATGGCCAGTTTT-TAMRA3', where Fl means fluorophore) were purchased from Integrated DNA Technologies (IDT).

Based on the four direct repeats on the DNA sequence, two 34-bp complementary single-stranded DNA oligonucleotides containing the YczR binding region (5'- TTTCCTCAACAT

CATTGCGAATGCCCTCATCATC-TAMRA-3') and (5'-AAAACTGGCCATTAATTA TCAGTCCATTTTTGCT-TAMRA-3') were also purchased from IDT.

Both oligonucleotides were synthesized and labeled with the fluorophore, carboxytetramethylrhodamine (TAMRA) attached at the 3' terminus. TAMRA has an emission maximum of 580 nm and an absorbance maximum of 565 nm.

The double-stranded oligonucleotides were prepared by annealing a mixture containing an equimolar concentration of the two single-stranded oligonucleotides at 25 μM in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 50 mM NaCl. The annealing solution was heated at 90°C for 1 minute, and then cooled 5°C per minute until it reached 10°C by using a Mastercycler Personal (Eppendorf).

For the fluorescence polarization assay, YczR was serially diluted in the buffer (50 mM HEPES, pH 7.5, 5% Glycerol, 300 mM NaCl) to acquire the appropriate concentrations against a fixed concentration of TAMRA-labeled DNA (1 nM).

The constant concentration (0.1 nM) of TAMRA-labeled DNA was mixed with the serially diluted YczR protein (final concentration, 0-1,000 nM). The DNA binding reaction contained 50 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 0.1 μg/μL BSA, 100 mM NaCl, and various concentrations of protein typically in 100 μL. The sample mixture was loaded into the individual wells of a 384-well micro-plate. Fluorescence polarization data for the YczR DNA complex were obtained using a BioTek Synergy 2 Plate Reader (BioTek). The fluorescence polarization (FP) signals of the YczR-DNA mixtures were measured with fixed excitation (531 nm) and emission (595 nm) wavelength filters at room temperature. In the negative control binding assay, Add Bovine serum albumin (BSA) at increasing concentrations against a fixed concentration of TAMRA-labeled DNA.

The change plots in mP as a function of YczR concentration were used to determine the equilibrium dissociation constant (Kd) for the interaction between YczR and DNA binding site. The data was fit to a hyperbola using the program Kaleida Graph (Synergy Software). Determine Kd by: (Bmax)([S])/Kd+[S], where Kd and Bmax are derived from the single-ligand binding plot, and [S] is the concentration of the fluorescein-labeled DNA. Use the Kd result to calculate the Ki using the following equation: Ki=EC50/ (1+[s])/kd, where Kd is for YczR binding to DNA sequence.

# CHAPTER FIVE

# CONCLUSIONS AND FUTURE INVISTIGATIONS

Many widespread DUFs are biologically essential in bacteria, and the importance of understanding and prioritizing DUF research has been recognized in the last few years[80]. DNA sequence alignment results have indicated that YczR in *K. pneumoniae* is the transcriptional regulator of YczE (DUF161) a new MocR/GabR subfamily member of the GntR superfamily.

In this study, we have defined two X-ray structures of the YczR regulation domain: a 1.79 Å structure with PLP bound and a 2.0 Å apo structure without PLP.

To the best of our knowledge, these structures are the first YczR X-ray crystal structures solved. The elucidation of the three-dimensional YczR crystal structure provided important information regarding the structure-function relationship and domain arrangements of the MocR/GabR subfamily. Additionally, this information allowed for transcriptional activity analysis of YczR.

GabR is the best genetically and biochemically characterized MocR/GabR subfamily member. The members of the MocR/GabR subfamily contain an N-terminal wHTH DNA-binding domain and a C-terminal aminotransferase-like PLP binding domain. YczR and GabR belong to the same evolutionary subgroup the MocR/GabR subfamily and share similar mechanistic, functional, and structural features.

The N-terminal helix-turn-helix DNA-binding domain was not found in the two YczR crystal structures. However, SEC-MALS and mass spectrometry results suggested that the YczR protein is full length in solution after purification. We speculated that YczR lost the DNA-binding domain during crystallization, likely due to a contaminated protease activity. Although YczR and GabR are similar in molecular weight, the SAXS results indicated that YczR occupies a greater volume. We surmise that the interactions between the DNA binding domain and the regulation domain within YczR are weaker than in GabR. This allows for full extension of the linker region, which contributes to the loss of the DNA binding domain during crystallization.

Similar to GabR, YczR has a co-factor, PLP, which is a common co-enzyme among a large number of enzymes. Also, like GabR[36], the YczR structure shown indicates that the ε-amino group of lysine331 forms a covalent Schiff base linkage (internal aldimine) with the aldehyde group of PLP; which serves as the only effector of the YczR. In contrast, GabR requires GABA for modulation of its activity in addition to PLP[54]. Based on Belitsky's publication, the intracellular level of PLP in bacteria is most likely not subject to significant variations[54]. It is unusual among bacterial transcriptional regulators to require a B6 vitamin for activity modulation.

YczR is a likely DNA-binding transcription regulator that requires the presence of PLP for activation of the yczE promoter. YczR binds DNA sequence between yczR gene and yczE gene in similar fashion comparing to GabR. From the fluorescence polarization results, it appears that an oligomeric form of YczR binds to two binding sites within the yczE gene. The YczR can recognize and specifically bind to sequences within the yczE promoter fragment.

YczR likely binds to the regulation region with a 2:1 stoichiometry. The dissociation constant (Kd) of the YczR-DNA complex was determined to be 26.47 nM. Although the precise stoichiometric binding of YczR for DNA is unknown, the dissociation constants which been determined from the fluorescence polarization assay suggest that YczR binds DNA with a high affinity. Four direct repeats of the sequence CCTCA and AATGG were located in this fragment, which appears to be essential for YczR binding. YczR binds tighter to the site closer to the yczE gene.

YczR regulation domain catalyzes a reversible half-transaminase reaction instead of being a non-enzymatic regulator like GabR. The only documented role of PLP as a cofactor in numerous enzymes is catalytic[81]. As such most of the aminotransferases are involved in the metabolism of amino group-containing compounds [42, 45]. It seems likely that PLP plays the same type cofactor role in YczR as it does in GabR. The aminotransferases PLP-binding domains evolved specifically to catalyze the aminotransferase reaction[46, 82, 83]. The UV-visible spectrum and fluorescence emission spectra results suggested the possibility that the regulatory domain of YczR maybe positive regulation of transcription because it could catalyze the half enzymatic reaction involving PLP.

Previous research efforts have shown that potent mechanism-based in-activators can be rationally designed against these PLP-dependent drugs-targets. Although the mechanistic function of PLP cofactor has been studied in details, its transcriptional regulator role has not been fully elucidated[54]. Our current research results will contribute to understanding the role of PLP in the *K. pneumoniae* YczR transcriptional regulator domain. We propose several possible

functions for the 330 amino acids long PLP-binding aminotransferase-like regulator domain of YczR can be suggested: (1) it has a structural role to maintain oligomerization overall and/or stability; (2) it is an effector-binding domain; (3) it likely performs an enzymatic reaction unrelated to regulatory functions and the DNA-binding of YczR; it may catalyze an enzymatic reaction, which is essential to exert its regulatory functions of YczR.

PLP also may interact covalently with taurine to activate TauR[55]. YczR likely responds to alanine to regulate taurine trafficking in *K. pneumoniae* to facilitate infection in animals and humans.

Taken together, all the results from UV absorption and half alanine-glyoxylate transaminase assay indicate that the YczR is highly specific for catalyzing L-alanine to pyruvate processing, thereby playing a pivotal role in L-alanine detoxification. In this half-reaction, the L-alanine reacts with PLP and transfers its α-amino group to PLP to make pyridoxamine phosphate (PMP). A new Schiff base is generated once the L-alanine interacts with the PLP commonly referred to as the external aldimine[36]. If in the complete transaldimination reaction, then PMP would transfer its nitrogen to the sugar to make an amino sugar. However, the crystal structure of YczR showns that PMP had been released from the regulation domain active site pocket.

L-alanine could have specificity affected to YczR-DNA interaction. It may affect YczR-DNA interactions though PLP. The use of a variety of biochemical techniques has allowed us to gain an in-depth understanding of the characteristics of YczR.

## Future Work

YczR monomer's molecular weight is 52.2kDa. YczR should be full-length protein in solution depend on mass spectrometry result. To research the selectivity and reaction mechanisms of YczR, future studies could involve active site residues mutations in YczR. Full-length YczR is not very stable at room temperature and low salt concentration solution. DNA binding domain of YczR would be moved to make truncated YczR protein.

DNA foot printing is a method to determining the DNA sequence specificity of DNA-binding proteins binds *in vivo* or *in vitro*. YczR is the transcriptional activator controlling the transcription of YczE (DUF161). The localization of YczR DNA binding sites would elucidate the function of YczE.

Many years of the alarming rise of bacterial resistance and antibiotics abuse has created a dangerous quandary for antibiotics research. Advances in the development of the scientific technology, and more tools have been provided for the improvement of the already established antibiotics and the new antibiotic classes discovery. The research of YczR maybe used to design inactivators, which are specific for drug targets to combat the rise of resistant gram-negative pathogens.

# CHAPTER SIX

# THE AQUEOUS DECAY OF A VANADIUM COMPLEX SEREDIPUOSLY REVEALS MISSING KNOWLEDGE ON THE REDOX STATES OF PTP1B

**Research aims: The decay of a vanadium complex and its redox effect on PTP1B, a target for diabetic treatment, *in vitro***

Diabetes is the seventh leading cause of death in the United States in 2016. The Center of Disease Control (CDC) and the American Diabetes Association (ADA) published the Diabetes Report Card in 2014, stating that 29.1 million Americans are diagnosed with diabetes as of 2012[84]. There are two major types of diabetes with , which people are diagnosed, type 1 and type 2.

Type 1 diabetes is genetically inherited and occurs when insulin is not produced in sufficient quantities. In type 2 diabetes, the insulin is produced in standard quantity, but the insulin receptors are desensitized and unable to transmit the signal inward to the cell[85, 86]. Type 2 diabetes is often seen in patients with unhealthy lifestyles and is of interest due to the prevalence of patients with heart disease[87].

Protein-tyrosine phosphatases (PTPs) are a class of enzymes that regulate a variety of fundamental cellular processes[88, 89]. PTP1B is a critical negative regulator of leptin signaling pathways[90, 91]. Protein Tyrosine phosphatase 1B (PTP1B) was the first PTP purified from human tissue[92].

Previous studies have indicated that PTP1B plays an important role in cell signaling relevant to diabetes, obesity, and cancer[92-94]. PTP1B is a critical negative regulator of leptin signaling pathways[90, 91]. PTP1B also is a key negative regulator of insulin signaling pathways, which inhibit the pathway responsible for controlling the uptake of glucose into the cell and lowering blood sugar levels[95]. Inhibition targeting PTP1B could prove to be effective in the treatment of type 2 diabetes and provide an opportunity to improve insulin sensitivity[84, 96]. However, the difficulty lies in identifying a specific, potent and safe PTP1B inhibitor. The two major challenges in developing potential therapeutics targeting PTP1B are lack of selectivity over other similar PTPs as well as lack of cell permeability[97]. Developing PTP1B inhibitors should simultaneously occupy both active site (Site-A) and the adjacent pTyr binding site (Site-B) partially to address the selectivity issue[98].

Vanadium compounds have been known to be insulin-enhancing agents both *in vitro* and *in vivo*[99]. Orthovandate ($VO_4^{3-}$), the vanadate form which most closely mimics $PO_4^{3-}$, can act as a substrate analog to inhibit all types of phosphatases because it shares similar structural properties with phosphorus[100].

Organic vanadium complexes have been proven more effective *in vivo* and *in vitro* studies than inorganic vanadium compounds[101]. It has been demonstrated that the peroxovanadium compounds irreversibly oxidize the thiol group of the conserved catalytic cysteine in the PTP catalytic sites[102]. Various vanadium compounds may have a different but overlapping spectrum of PTP inhibition activities, due to differences in their structure and stability. Vanadium is not well absorbed by mammalian cells, while some organic vanadium complex are better absorbed

because of the favorable membrane permeability[99, 103]. The detailed mechanisms by which different vanadium complexes enhance insulin's activity are still not fully understood[104, 105].

Another specific product generated during this decay and their effects on PTP1B was not looked at carefully. The prevalent explanation is that vanadium compounds would break down in solution and inhibit PTP1B activity via the generated vanadate acting as a transition state analog in the PTP1B reaction[102, 106, 107].

Vanadyl acetylacetonate, $VO(acac)_2$ is an organo-vanadium complex, where (acac) is the conjugate base of acetylacetone[108]. $VO(acac)_2$ has been shown to inhibit tyrosine phosphatases (PTPases), such as PTP1B[107]. Like other charge-neutral acetylacetonate complexes, it is more soluble in polar organic solvents.

This work will focus on the co-crystallization and inhibition of PTP1B using synthesized $VO(acac)_2$ complex (Figure 33). In addition to vanadium complexes having shown potential in the treatment of Type 2 diabetes, other studies indicated shown that some zinc complexes also have similar insulin-enhancing effect in treatment of type 2 diabetes[109].

Besides the drug design effort, the oxidation state of the catalytic cysteine in the PTP1B activity down regulation still presents some unsolved puzzles. Robin Carr & Harren Jhoti found that after soaking PTP1B with 2-phenl-sioxazolidine-3, 5-dione, the sulphur Sγ atom of Cys-215 formed a covalent bond with the backbone nitrogen atom of Ser-216[110]. In our studies, it was observed that while Cys-215 was oxidized; it did not form a covalent bond with the backbone nitrogen atom of Ser-216, representing a possible intermediate state in redox sensing.

In this study, we report the crystal structure of inhibited PTP1B obtained from the co-

crystallization experiment of PTP1B and $VO(acac)_2$. In addition to the inhibited structure via a vanadate species, the catalytic Cys-215 is shown to be partially oxidized and in two alternative conformations. Combining both inhibition studies and the crystal structures present important gap knowledge in the modification/oxidation of the PTP1B via the aqueous decay of $VO(acac)_2$.

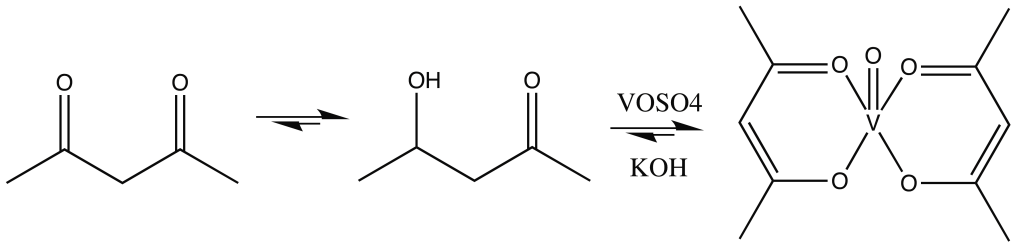
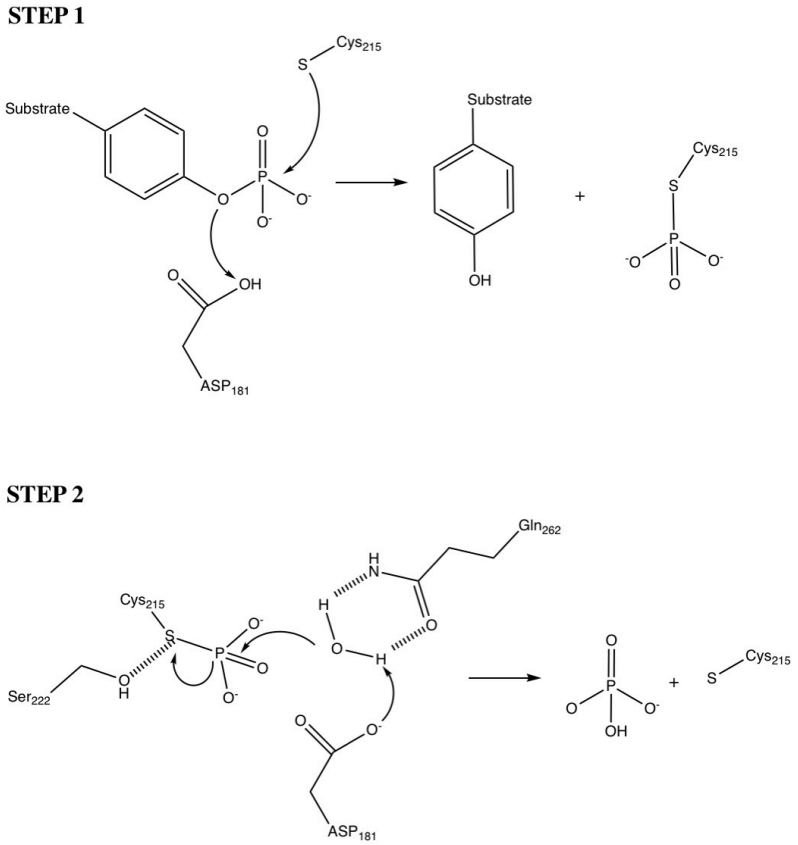



Figure 1. Synthesis of $VO(acac)_2$ complex.

## Results and Discussion

A variety of structural, enzymatic and kinetic studies from several laboratories have made clear that the mechanism of PTPs substrate recognition and catalysis[111-113]. All members of the PTPs are characterized by the presence of a signature motif, which contains the cysteine residue (Cys-215 in PTP1B) that is essential for catalysis[112]. PTP-mediated catalysis proceeds via a two-step ping-pong mechanism (Scheme 2).




Scheme 1. The general mechanism of the PTP-catalyzed reaction. The WPD-loop assumes a catalytically active closed conformation with the general acid in position to protonate the leaving group during formation of the phosphor-enzyme intermediate. In the second step this intermediate is hydrolyzed. After the phosphate product is released the WPD-loop open conformation becomes favored.

In the first step, a nucleophilic sulfur atom of the thiolate side chain of the cysteine attacks the phosphate ester moiety of the substrate, resulting in the formation of a phosphor-enzyme intermediate with the release of the peptidyl-tyrosine. The second step occurs via the phosphoenzyme intermediate attacked of a water molecule, which mediated by Asp-181 and Gln-262, yields the final products inorganic phosphate and the regenerated enzyme.

**Structure of PTP1B and $VO(acac)_2$ Complex**

The X-ray structures for PTP1B in the presence of the competitive inhibitor vanadate was determined by molecular replacement and refined to a 2.2 Å resolution. Overlapping X-ray crystal structures of PTP1B as ribbon diagram with $VO(acac)_2$ as a ribbon diagram is shown (Figure 34).

The PTP1B crystal structure to be solved was that of the 37kDa PTP1B catalytic domain with $VO(acac)_2$ complex. PTP1B is one of the 38 classical PTPs that specifically target phosphorylated tyrosine residues. It is composed of an N-terminal catalytic domain, two proline-rich motifs, and a C-terminal hydrophobic region. The crystals belong to P21 space group with cell dimensions (a= 52.17 Å, b = 71.75 Å, c = 88.01 Å) diffracted to 2.2 Å. The vanadium compound engages the catalytic site of the PTP1B protein.

Figure 2. Overlapping X-ray crystal structures of PTP as ribbon diagram 1B with $VO(acac)_2$ as a ribbon diagram.

The electron density of residues revealed the presence of covalently bound Vanadate in the pocket situated at the interface of the subunits in the crystal structure of PTP1B (Figure. 35). It is of interest to note that the structure of the initial Vanadium complex had degraded before the actual binding of the inhibitor to PTP1B. The pocket also corresponds to the cofactor binding sites in the related enzymes and therefore defines the active site of PTP1B.

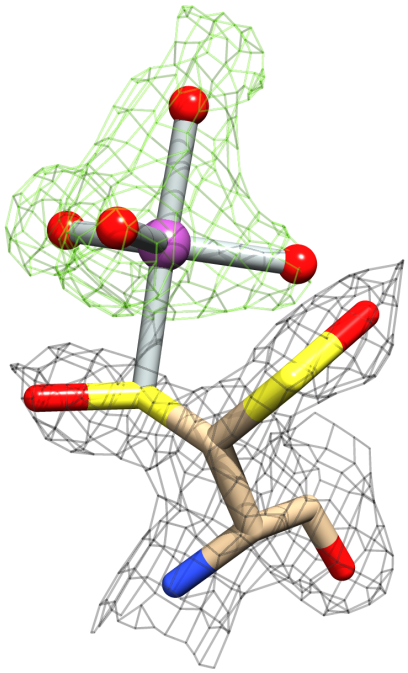


Figure 3. Electron density of the vanadate modified catalytic cysteine. Vanadate is shown with the refined 2Fo-Fc electron density map contoured at 1.0 σ.

Table 2. Crystallographic data collection and refinement statistics for PTP1B

| | PTP1B |
|---|---|
| Space group | P21 |
| Cell dimension | |
| α, β, γ (deg) | 60.743, 61.26, 88.307 |
| a, b, c (Å) | 90.0, 90.041, 90.0 |
| Resolution (Å) | 1.83 |
| Resolution at I/σ (I) =2 | 1.90 |
| $R_{pim}$ [a] (%) | 5.8(43) |
| I/σ (I) | 12.3(8.3) |
| CC ½ [b] | 0.673 |
| Completeness (%) | 91.1(83.4) |
| Multiplicity | |
| No. Reflections | |
| No. Unique Reflections | |
| | |
| $R_{work}$ [c]/$R_{free}$ [d] (%) | 17.7/22.6 |
| No. of Atoms | |
| protein | 13937 |
| ligand | 66 |
| water | 1400 |
| B factors (Å$^2$) | |
| protein | 29.135 |
| ligand | |
| RMSD [e] | |
| bond lengths (Å) | 0.02 |
| bond angles (deg) | 1.96 |
| Ramachandran plot (%) | |
| most favored | |
| allowed | |
| outliers | |

[a]Precision-indicating merging R
[b]Pearson correlation coefficient of two "half"data sets
[c]Rwork = Σ|Fobs − Fcalc|/ΣFobs
[d]Five percent of the reflection data were selected at random as a test set, and only these data were used to calculate Rfree
[e]Root-mean square deviation
[h]Not applicable

The Electron Density in Green is the simulated annealing omit difference map (2Fo-Fc), at 1.0 sigmas, that shows the existence of orthovanadate ($VO_4^{3-}$). The central vanadium atom of orthovanadate is covalently linked to the sulfate of Cys-215. The Electron Density in grey is the simulated annealing composite map (2Fo-Fc) showing sulfenic acid, which is oxidized at the Cys-215 residue and exists in 2 positions at 1.0 sigma.

From the omit map, the thiol group on Cys-215 has been oxidized. The side chain exists in two conformations, with one bound to Vanadate and the other facing out of the active site. Presently, the oxidized Cys-215 remains active. $VO(acac)_2$ forms Electrostatic interactions with active site Cys-215 which may block substrate biding in the active site and/or decrease catalytic activity.

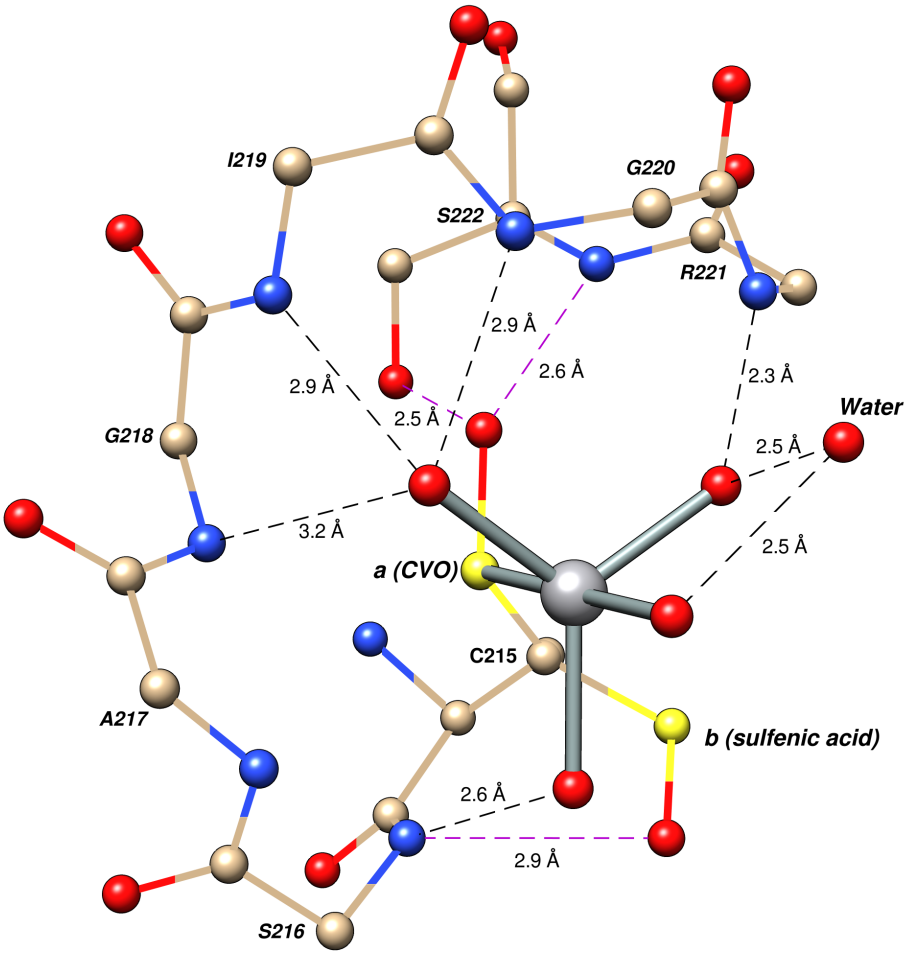



Figure 4. Electrostatic interactions between the $VO(acac)_2$ inhibitor and the PTP1B active site. The hydrogen bonds between inhibitor and PTP1B are indicated by dashed lines. The crystal structure of the PTP1B: $VO(acac)_2$ complex confirmed these noncovalent interactions.

Electrostatic interactions between the $VO(acac)_2$ bound Cys-215 and the PTP1B active site loop are shown (Figure 36). In the PTP1B covalent modification vanadium oxide (CVO) formation, the oxygen atoms in the modified Cys-215 forms two hydrogen bonds with the side chain hydroxyl of Ser-222 and the backbone amide of Arg-221; one of the oxygen atoms of vanadate forms hydrogen bonds with the backbone amides of Gly-220, Ile-219 and Gly-218. The second oxygen atom of vanadate is stabilized by the backbone amide groups of Arg-221 and a

water molecule, which interacts with the third oxygen atom of vanadate. The fourth oxygen atom forms hydrogen bonds with the backbone amide of Ser-216. In the sulfenic acid formation, the oxygen atom in the modified Cys-215 forms one hydrogen bond with the backbone amide of Ser-216.

Previous work by Robin Carr & Harren Jhoti[110] found in the oxidized PTP1B active site, the sulphur atom in the Cys-215 has formed a covalent bond with the backbone amide of Ser-216. A crystallographic soaking experiment of PTP1B with one identified inhibitor (2-phenyl-isoxazolidine-3, 5-dione) formed a sulphenyl-amide derivative of Cys-215. The sulphur Sγ atom of Cys-215 formed a covalent bond between the backbone nitrogen atom of Ser-216[110] (Figure 37, B). The oxidation of catalytic Cys-215 formed sulphenyl-amide formation, which is a protective intermediate in the oxidative inhibition of PTP1B and caused significant changes in the enzyme active site.

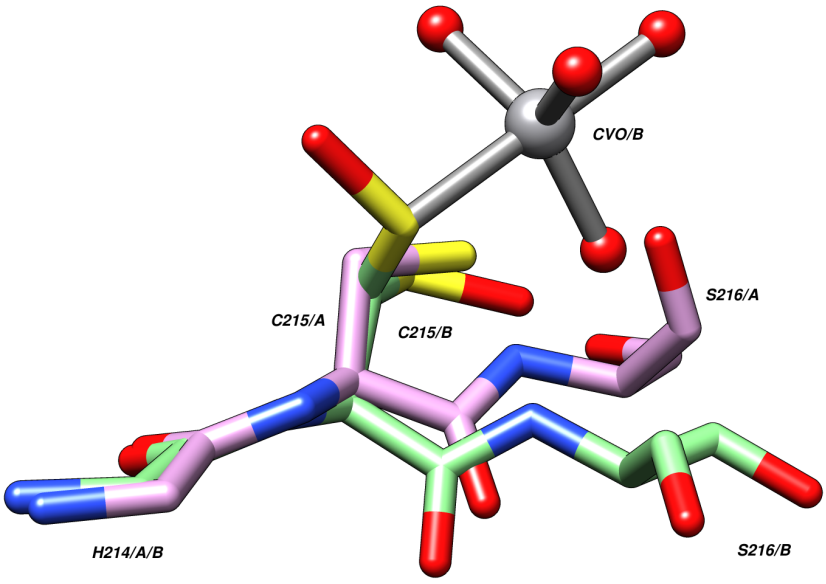



Figure 5. Superimposed images of sulphenyl-amide PTP1B (PDB: 1OES) (purple, A) and $VO(acac)_2$ inhibition of PTP1B (green, B) show structural changes in the catalytic Cys-215 active site. With $VO(acac)_2$ inhibitor bound, the Sy atom of Cys-215 flips away from the backbone nitrogen atom of Ser-216 and can no longer form a sulphenyl-amide bond.

In the previous research, after the oxidation of Cys-215 to sulphenic acid, the nitrogen atom of Ser-216 on the backbone nucleophilic attacked the Sγ atom of Cys-215 and formed a 1.7 Å covalent bond[110]. However, in the vanadate bound Cys-215 crystal structure of PTP1B sulfenyl-amide formation, the distance between the Sγ atom of Cys-215 and the nitrogen atom of Ser-216 was 2.9 Å, which suggested hydrogen bond formed (Figure 37, A). The present structure identifies an intermediate of their proposed mechanism of inactivation. This study provides insight to further understanding of the redox inactivation of such proteins.

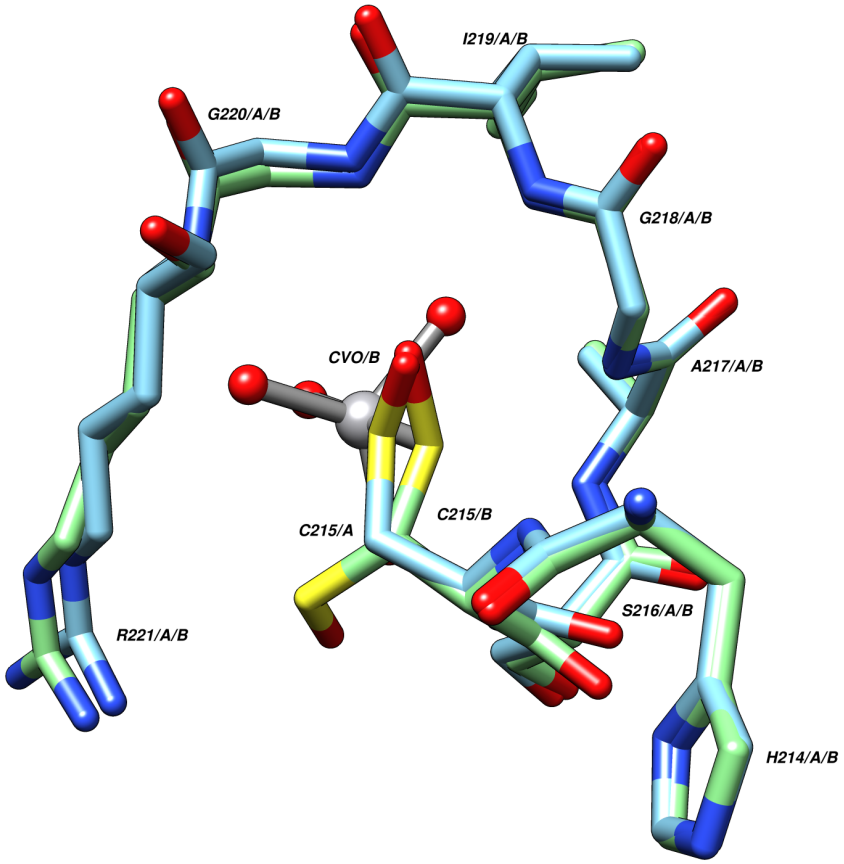



Figure 6. Superimposition of the oxidation state structures of the PTP1B catalytic cysteine. The structure of the sulphenic acid derivative of Cys-215 is shown in blue (A). The structure of the $VO(acac)_2$ derivative is in green (B). The phosphate binding cradle structure is largely conserved with $VO(acac)_2$ bound to the catalytic cystine.

The structures of key residues at the active site of the $VO(acac)_2$ derivative are very similar to the structure of the sulphenyl-amide derivative oxidation states (Figure 38). And the structures of these oxidation states are very similar to the structure of reduced PTP1B. The vanadate bound Cys-215 intermediate is stabilized by the similar interactions observed for the oxygen atoms in the sulphenic acid derivate. However, in the sulfenic acid derivative, the side chain of A Cys-215 is rotated about $180^{o}$.

The movement of the WPD-loop (a conserved protein loop) is important in the PTP1B catalyzed reaction. The WPD-loop has a “closed” and an “open” two distinct conformation. The WPD-loop would fold over the active site and bring the conserved Asp residue up to 8 Å closer to the bound substrate. The WPD-loop has negligible interaction with the P-loop in the “open” conformations[114-117]. The Vanadate also may reduce the mobility of the WPD-loop, which may decrease catalytic activity.

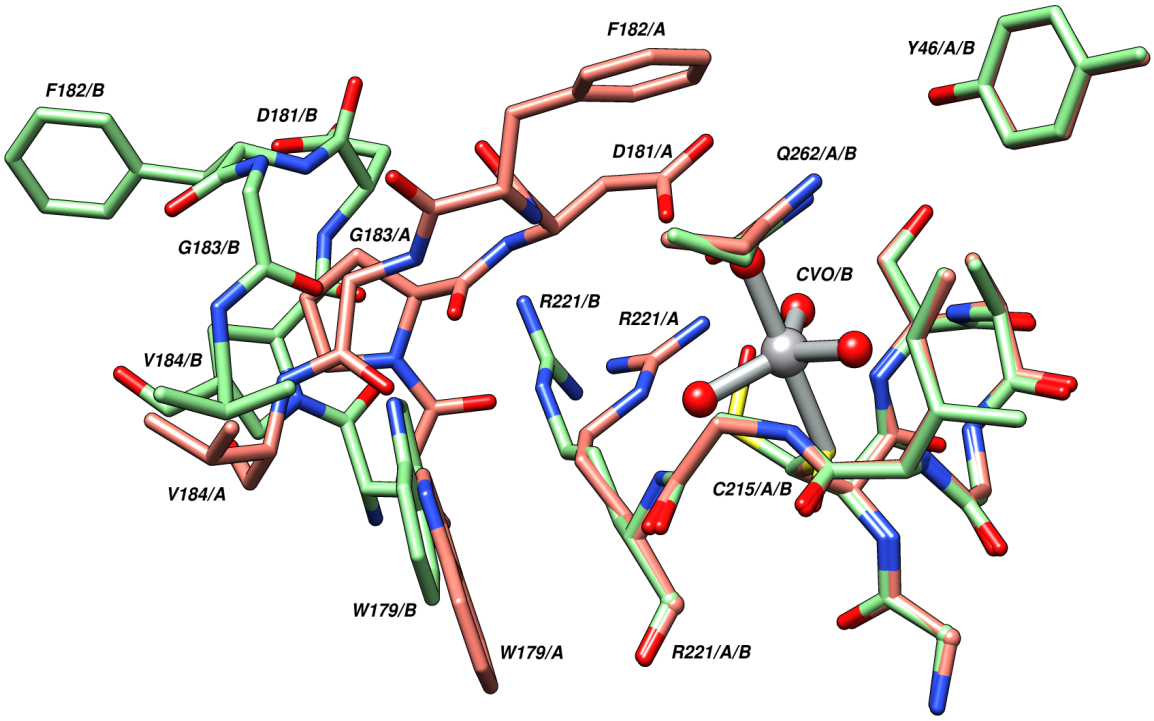



Figure 7. Overlay of the WPD-loop open vanadate-bound (green) and the WPD-loop closed conformations in the active sites of the PTP1B enzymes. The trigonal bipyramidal structure of the vanadate anion is formed by a nucleophilic attack on the catalytic cystine. This also results in a conformational change in the F182 residue.

Different from Johnson's result[112], in the crystal grown in the presence of vanadate, a WPD-loop open vanadate-bound form conformation is observed. The conformational differences between the open and closed conformations in the WPD-loop and the side chain of Arg-221of the PTP1B in both the ligand-bound and ligand-free states had been shown (Figure 39).

In the WPD-loop open vanadate-bound form (Figure 39, B), the bound vanadate exhibits a trigonal bipyramidal conformation with apical positions occupied by the sulfur atom of Cys-215 and an oxygen atom from a water molecule. Compared with the WPD-loop closed form (Figure 39, A), the side chain of Arg-221 is rotated about $70^{o}$ around chi-3 and form a hydrogen bond to the Phe-179 carbonyl group. In both the open and closed forms, the side chain of Phe-179 is embedded into a hydrophobic pocket. Unlike the WPD-loop, the P-loop has no significant motion during catalysis[118]. The P-loop backbone residues align closely in both the ligand-bound and ligand-free states.

**Kinetic results of PTP1B with $VO(acac)_2$ Complex**

The data obtained using $VO(acac)_2$ complex as an inhibitor is shown in Figure 40. The kinetic data suggests a competitive mode of inhibition because the apparent Vmax decreases, while the apparent Km decreases, too. The concentration of vanadium complex varied from 0-0.3333 uM. The initial kinetic parameters obtained without vanadium complex are Vmax= 0.007 U/mM and Km= 2.8±0.1mM. After 16 hours, the PTP1B lost 2/3 activities compared with the new purified protein (Figure 41).

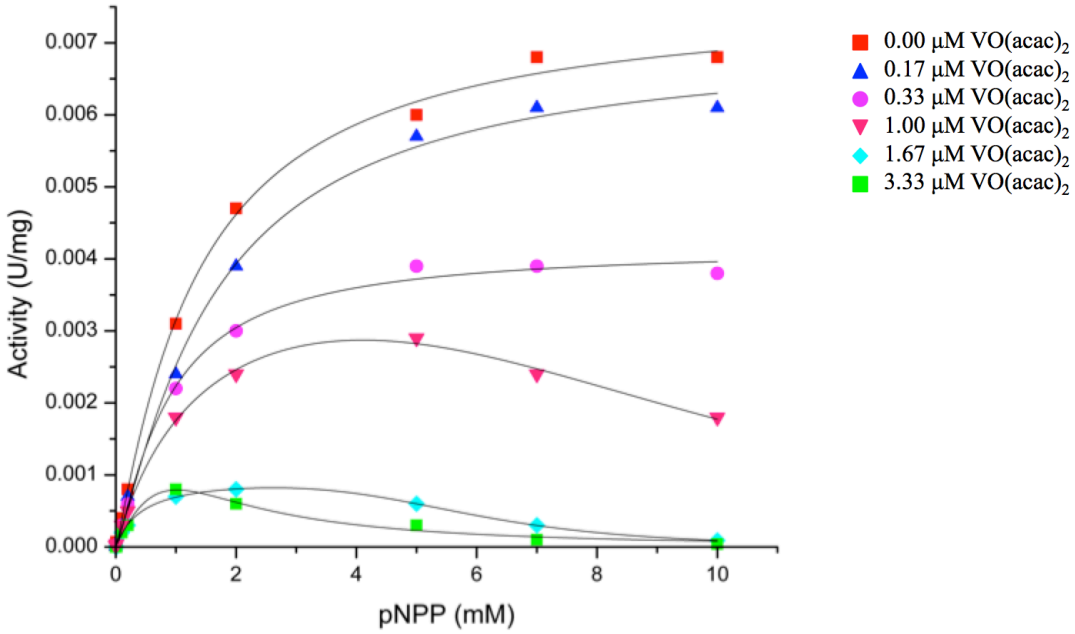



Figure 8. Activity versus substrate concentration for PTP1B against $pNPP2^-$ in the presence of varying $VO(acac)_2$ complex concentrations.

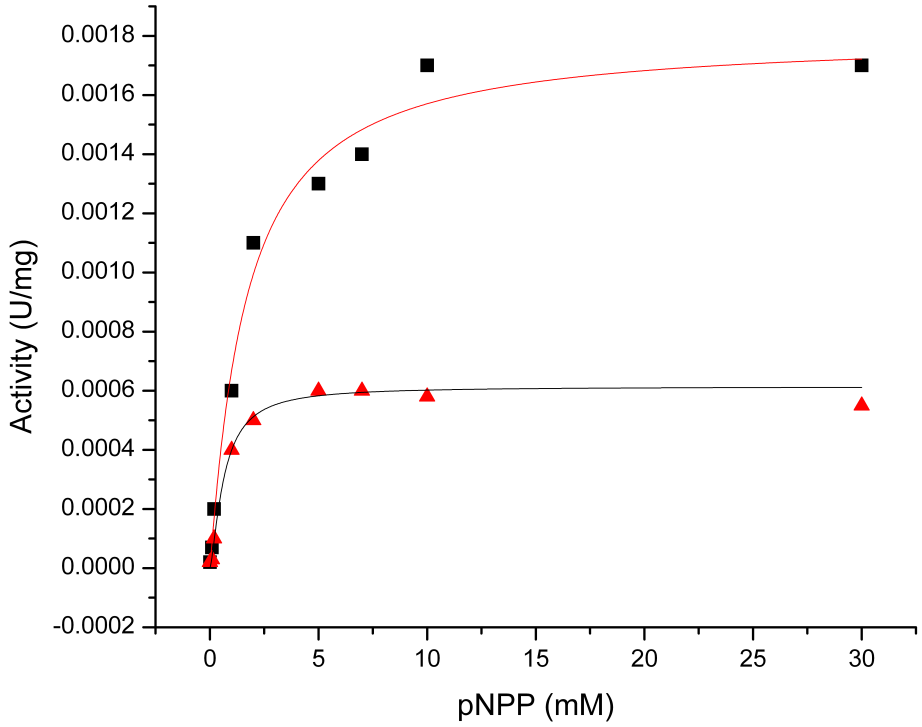



Figure 9. Activity versus time for PTP1B.

# Materials and Methods

## Materials

Chemicals for PTP1B purification, crystallization, assays were purchased from Sigma-Aldrich (St. Louis, MO), Alfa Aesar (Tewksbury, MA) and Fisher Scientific (Pittsburgh, PA). YczR gene was synthesized with codon optimization for *E. coli* expression from Integrated DNA Technologies (Coralville, IA). *E. coli* BL21 (DE3) cells and Turbo cells were purchased from New England BioLabs (Ipswich, MA). Bacterial growth media Luria-Bertani (LB) broth and antibiotics were obtained from Sigma-Aldrich. Crystallization screen solutions and other crystallization supplies were purchased from Hampton Research (Aliso Viejo, CA) and Emerald Bio (Bedford, MA). All other materials were purchased at the highest quality available.

## Methods

**DNA Cloning.** The PTP1B DNA fragment was digested with NdeI and NotI and subsequently cloned into the expression vector pETite. Transform the pETite-kan-c-5His vector with full-length PTP1B insert into *E. coli* Turbo cells for plasmid storage. The PTP1B plasmid was determined that there were no mutations in the entire coding region by DNA sequencing. (DNA sequencing & Genotyping Facility, University of Chicago, Chicago, IL) The resulting plasmid PTP1B was transferred to competent *E. coli* BL21DE3 cells for protein production and purification.

**Protein expression and purification.** The PTP1B protein was purified by using the intein-mediated purification with an affinity His-tag system. *E. coli* BL21DE3 cultures were grown overnight at 37°C while being shaken (250 RPM) in LB medium containing kanamycin (50

ug/mL). Induction of PTP1B expression was carried out at an $OD_{600}$ reached 0.6-0.8 by adding IPTG to a final concentration of 0.5 mM. Cells were harvested after 16 h incubation by centrifugation at 8670 G and 4 °C for 15 min (Avanti J-E Centrifuge, Beckman Coulter).

The cell pellet was resuspended in Ni-NTA column wash buffer (what is Ni column buffer). Using sonication (Qsonica sonicators, Q500 Sonicator) to break the cellular wall, the cell debris was removed by centrifugation at 31000G and 4°C for 20 min.

Protein purification was purified using affinity chromatography (Ni-NTA, GE Healthcare). The Ni-NTA column was equilibrated with a wash buffer (50 mM NaH2PO4, 500 mM NaCl, and 10 mM imidazole, pH 7.5) and the PTP1B fusion protein was eluted from the column using wash buffer with imidazole (100 mM)

The fractions protein was concentrated to 2 mL using a 10 kDa Amicon-Ultra centrifuge filter tube (Centricon Plus centrifugal filter units; Millipore). Further purified by size-exclusion chromatography (HiLoad 16/60 Superdex 200, GE Healthcare), which was equilibrated with gel filtration column buffer (300 mM NaCl, 50mM HEPES, pH 7.5).

The protein concentration was determined with a Bio-Rad protein assay kit. The protein purity assessed by SDS-PAGE. Purified PTP1B protein was concentrated (Centricon Plus centrifugal filter units; Millipore) and flash frozen in 1000μL aliquots in liquid nitrogen after the addition of 20% (V/V) glycerol and stored at -80°C for long-term.

**Synthesis Specific compounds as potential PTP1B inhibitors.** $VO(acac)_2$ had been synthesized from vanadyl sulfate: $VOSO_4 + 2\ Hacac \rightarrow VO(acac)_2 + H_2SO_4$. The procedures were slightly modified from the reported in the literature[112]. Syntheses $VO(acac)_2$ complexes

was performed under a pure $N_2$ atmosphere using standard Schlenk and air-free manipulation techniques. The vanadium forms a square pyramidal structure. After the reaction, the complexes are purified using recrystallization.

**Crystallization.** The purified YczR was buffer exchanged into crystallization screening buffer (50 mM HEPES, pH 7.5, 300 nM NaCl) and concentrated to 10 mg/ mL using a 10,0000 molecular weight cutoff (MWCO) Amicon-Ultra centrifugal filter device (Millipore). PTP1B and $VO(acac)_2$ were pre-mixed before crystallization. The initial sparse matrix crystal screen was using the Gryphon crystallization robot (Art Robbins Instruments) to obtain initial crystallization hits, and further optimized manually to obtain crystals with the best size and morphology.

After the initial crystallization screen, crystal optimizations of PTP1B were obtained by the hanging drop method using protein at 10 mg/mL at room temperature. The optimized crystallization solution contained 20% (W/V) PEG6000, 100 mM HEPES, pH 7.5, 200 mM $MgCl_2$. The hanging drops were incubated at room temperature. Crystals appeared in 6 days and grew to maximum size in1-2 week.

Crystals with good size and morphology were picked from the hanging drops and flash-cooled in liquid nitrogen using 20% (w/v) glycerol in addition to the compounds of the reservoir solution as the cryoprotectant.

**Diffraction data collection, Structure Determination, and Refinement.** Diffraction data of PTP1B crystals were collected at Advanced Photon Source Structural Biology Center (SBC) 19ID and 19BM beamlines. (Argonne National Laboratory, IL). The crystal diffraction images were recorded by the 4×4 tiled 300×300 $mm^2$ sensitive area CCD detector from Area Detector

Systems Corp. (ADSC). The data collection was at a wavelength of 1.0 Å. The datasets were indexed, integrated and scaled using the HKL2000[74]. The datasets were all scaled in the space group P21 to resolutions of 2.2 Å respectively.

Using PHASER[75] from CCP4 program suite performed molecular replacement, using the previously reported PTP1B (PDB ID: 1BZJ) as the search model[117]. Rigid body refinement followed by restrained refinement was carried out in Refmac5 in the CCP4 software suits until no more side chains can be built in automatically[78].

The remainder of the model was built and modified manually in Coot based on electron density maps. Before built the ligand in the model, the electron density from the Fo-Fc map supported the existence of the modification of Cys-215. The chemical restraint of Vanadium compound was generated in the program JLigand from CCP4.

The ligand was manually fit in the model based the residual electron density in the deference (Fo-Fc) map. The PTP1B with Vanadium ligand model was further refined in Refmac 5. The R free and R factor for the inactivated structure were 0.226 and 0.177, respectively. The data collection and refinement statistics for the PTP1B are given in Table 2. All figures depicting crystal structures were prepared by using Chimera.

**PTP1B inhibition assay.** The PTP1B assay was used Synergy H1 hybrid multimode microplate reader (Biotek, USA) with 96 well plates (Greiner Bio-one, Monroe, NC). The PTP1B assay was carried out in a continuous manner using *para*-nitrophenylphosphate ($pNPP^{2-}$) as the substrate. The absorbance of the product *para*-nitrophenol (*pNP*) was measured at 405 nm using a microplate reader. All inhibitors were pre-incubated with PTP1B in a buffer containing

50 mM HEPES (pH7.5), and 5 mM DTT, 300 mM NaCl for 5 min at 25 °C. The reaction that was followed is given in equation.

$$pNPP^{2-}(aq) + H_2O(aq) \rightarrow PO_4^{3-}(aq) + pNP\,(aq) + H^+(aq)$$

Calculate enzyme activity from the data. The assay was carried out at 25 °C using the substrate $pNPP^{2-}$ in the presence of 2 mM EDTA in a 50 mM Bis-Tris (pH 6.3) buffer, which provides a constant ionic strength over a wide pH range.

Reactions were initiated by addition of enzyme into the microplate wells (final reaction volume of 300 uL). Inhibition constants ($K_i$) for PTP1B were determined by measuring initial rates of pNPP hydrolysis in the presence of a range of concentrations of inhibitor (0-3.33 uM) and [$pNPP^{2-}$] (0-20 mM) at pH 6.3 and 25 °C. Kinetic parameters were determined by a fit of the initial rate (v) versus [$pNPP^{2-}$] data to the Michaelis-Menten equation.

# CHAPTER SEVEN

# CONCLUSIONS

PTP1B is one of the most studied PTPs and plays a critical role in the regulation of the insulin-signaling pathway[84, 114]. The highly charged PTP1B catalytic site contains the common structural motif of PTPs. In the research pertaining to inhibition of PTP1B, specificity issues with other PTPs still present significant challenges, because of the high homogeneity to other cellular PTPs. It also remains a challenge in medicinal chemistry to design inhibitors with relatively uncharged outer surfaces and hydrophobic regions. This design would allow the inhibitor to cross membranes while maintaining relative efficacy once inside cells.

PTPs specifically catalyze the dephosphorylation of phosphotyrosyl residues of peptides and proteins[119-121]. Orthovanadate ($VO_4^{3-}$) is a phosphate analog that can quickly adopt a trigonal bipyramidal structure and could bind as a transition state analog to inhibit the PTP1B[122]. Previous research has proved this part of vanadate's insulin-mimetic effect due to its PTPs inhibition function[123].

Vanadate was reported to inhibit protein phosphotyrosine phosphatase and activate autophospohrylation of solubilized insulin receptors[124]. Vanadate is insulin-mimetic[125], and its insulin-mimetic properties have been researched in details[107]. It has been shown to be potentially useful in treating both noninsulin- and insulin-dependent diabetes mellitus in human clinical trials[126]. Vanadate has multitudinous effects on biological systems[123]. The biochemistry of vanadate behaves like phosphate due to their similar molecular geometry [127].

Vanadate may also have a strong affinity for the PTP1B active site[119-121, 128].

Vanadate would chelate many organic molecules and also reversibly coordinate free thiol and hydroxyl groups. The interaction of vanadate with assay components and buffers has also been monitored in research investigating the mechanism of vanadate inhibition [106]. HEPES was chosen for the studies because it is one of the few buffers that does not have appreciable interaction with vanadate[106].

After $VO(acac)_2$ enters the cell, it would be degraded to vanadyl (IV) ion, and got oxdiated to vanadate (V)[129-131]. Furthermore, vanadate often was treated as a simple phosphate analog. Its rich redox and coordination chemistry also need to be considered, including coordination with many reagents and varied oxidation states[107].

Here we report the X-ray structure of PTP1B with $VO(acac)_2$ complex at a 2.2Å resolution. The detailed PTP1B crystal structure may provide more information about the ability of $VO(acac)_2$ to bind the active site and inhibit PTP1B activity. The results suggest that the oxovanadium (IV) complex $VO(acac)_2$ dissociates and oxidizes to a vanadium (V) species vanadate ($VO_4^{3-}$) under crystallization conditions. This structure further suggests that better bioavailability of $VO(acac)_2$ is the primary reason for its enhanced inhibitory behavior compared to simple salts[132].

While the precise biochemical pathway of vanadate action is not yet known, these kinetics assay and X-ray crystallography studies will provide reliable proof of concept, validating the inhibition of PTP1B by $VO(acac)_2$. Catalytic cysteine exists as a thiolate anion, making it susceptible to oxidation and more reactive. The mechanism of decrease regulating PTP1B activity *in vivo* may occur because of the oxidation of the catalytic Cys-215 in the active site.

Dithiothreitol (DTT) or glutathione can reduce the disulphides with Cys-215 to resore PTP1B activity[111].

An essential cysteine residue at the catalytic site is the hallmark of PTPs. The cysteine would form a Thiol-phosphate intermediate during catalysis[133]. In this study, we set out to elucidate the mechanisms of vanadate inhibition of PTP1B.

Compared with vanadium compounds, zinc compounds have received less attention in anti-diabetic drug research. Zinc compounds also play essential structural roles in PTPs inhibition in vivo[134]. Previous research has shown that zinc compounds do not show inhibition of PTP1B at concentration lower than 10 uM[109].

Compared with previous research, this crystal structure also redefines the oxidation state of the catalytic Cys-215 in PTP1B. The nucleophilic Sg atom of Cys-215 attacks the backbone nitrogen atom of Ser-216 formed the sulphenyl-amide bond occurs through oxidation of Cys-215 to sulphenic acid. The hydrogen bond interaction between the Nd1 atom of the invariant His-214 side chain and the carbonyl oxygen atom of Cys-215 increases the partial charge on the backbone nitrogen atom of Ser-216 and enhance its reactivity. This assumption supports the nucleophilic substitution mechanism[114].

During the *in vitro* oxidation reactions of the cysteine and thiol residues, sulfenic acids are commonly generated as intermediates[135]. We used a small organic vanadate molecule as a useful model to research the redox-sensing assembly of the PTP1B active site functional groups. The possible interactions and reactions of vanadate must be taken into consideration when these inhibitors are used *in vivo* and *in vitro*. Our results may help to better understand the mechanism

of inhibition of PTP1B, This would then aid in the design of potential inhibitors against the crucial therapeutic target of diabetes treating.

# REFERENCE LIST


1. Ryan, K.J., C.G. Ray, and J.C. Sherris, *Sherris medical microbiology : an introduction to infectious diseases*. 4th ed. 2004, New York: McGraw-Hill. xiii, 979 p.

2. Fair, R.J. and Y. Tor, *Antibiotics and bacterial resistance in the 21st century.* Perspect Medicin Chem, 2014. **6**: p. 25-64.

3. Yigit, H., et al., *Novel carbapenem-hydrolyzing beta-lactamase, KPC-1, from a carbapenem-resistant strain of Klebsiella pneumoniae.* Antimicrob Agents Chemother, 2001. **45**(4): p. 1151-61.

4. Nordmann, P., G. Cuzon, and T. Naas, *The real threat of Klebsiella pneumoniae carbapenemase-producing bacteria.* Lancet Infect Dis, 2009. **9**(4): p. 228-36.

5. Schwaber, M.J. and Y. Carmeli, *Carbapenem-resistant Enterobacteriaceae: a potential threat.* JAMA, 2008. **300**(24): p. 2911-3.

6. Arnold, R.S., et al., *Emergence of Klebsiella pneumoniae carbapenemase-producing bacteria.* South Med J, 2011. **104**(1): p. 40-5.

7. Yinnon, A.M., et al., *Klebsiella bacteraemia: community versus nosocomial infection.* QJM, 1996. **89**(12): p. 933-41.

8. Skogberg, K., et al., *Increase in bloodstream infections in Finland, 1995-2002.* Epidemiol Infect, 2008. **136**(1): p. 108-14.

9. Uslan, D.Z., et al., *Age- and sex-associated trends in bloodstream infection: a population-based study in Olmsted County, Minnesota.* Arch Intern Med, 2007. **167**(8): p. 834-9.

10. Giani, T., et al., *Emergence in Italy of Klebsiella pneumoniae sequence type 258 producing KPC-3 Carbapenemase.* J Clin Microbiol, 2009. **47**(11): p. 3793-4.

11. Teo, J.W., et al., *Emergence of clinical Klebsiella pneumoniae producing OXA-232 carbapenemase in Singapore.* New Microbes New Infect, 2013. **1**(1): p. 13-5.

12. Bogaerts, P., et al., *Emergence of clonally related Klebsiella pneumoniae isolates of sequence type 258 producing KPC-2 carbapenemase in Belgium.* J Antimicrob Chemother, 2010. **65**(2): p. 361-2.

13. Samuelsen, O., et al., *Emergence of clonally related Klebsiella pneumoniae isolates of sequence type 258 producing plasmid-mediated KPC carbapenemase in Norway and Sweden.* J Antimicrob Chemother, 2009. **63**(4): p. 654-8.

14. Nastro, M., et al., *[Emergence of colistin-resistant Klebsiella pneumoniae. Microbiological and epidemiological characterization of the isolates producing and non-producing KPC-type carbapenemase]*. Rev Argent Microbiol, 2013. **45**(3): p. 185-90.

15. van Duin, D., et al., *Carbapenem-resistant Enterobacteriaceae: a review of treatment and outcomes.* Diagn Microbiol Infect Dis, 2013. **75**(2): p. 115-20.

16. Centers for Disease, C. and Prevention, *Guidance for control of infections with carbapenem-resistant or carbapenemase-producing Enterobacteriaceae in acute care facilities.* MMWR Morb Mortal Wkly Rep, 2009. **58**(10): p. 256-60.

17. Neuner, E.A., et al., *Treatment and outcomes in carbapenem-resistant Klebsiella pneumoniae bloodstream infections.* Diagn Microbiol Infect Dis, 2011. **69**(4): p. 357-62.

18. Kazakov, A.E., et al., *Transcription factor family-based reconstruction of singleton regulons and study of the Crp/Fnr, ArsR, and GntR families in Desulfovibrionales genomes.* J Bacteriol, 2013. **195**(1): p. 29-38.

19. Small, S.P., et al., *Struggling to understand: the experience of nonsmoking parents with adolescents who smoke.* Qual Health Res, 2002. **12**(9): p. 1202-19.

20. Seshasayee, A.S., K. Sivaraman, and N.M. Luscombe, *An overview of prokaryotic transcription factors : a summary of function and occurrence in bacterial genomes.* Subcell Biochem, 2011. **52**: p. 7-23.

21. Browning, D.F. and S.J. Busby, *The regulation of bacterial transcription initiation.* Nat Rev Microbiol, 2004. **2**(1): p. 57-65.

22. Haydon, D.J. and J.R. Guest, *A new family of bacterial regulatory proteins.* FEMS Microbiol Lett, 1991. **63**(2-3): p. 291-5.

23. Yoshida, K.I., Y. Fujita, and S.D. Ehrlich, *An operon for a putative ATP-binding cassette transport system involved in acetoin utilization of Bacillus subtilis.* J Bacteriol, 2000. **182**(19): p. 5454-61.

24. DiRusso, C.C., T.L. Heimert, and A.K. Metzger, *Characterization of FadR, a global transcriptional regulator of fatty acid metabolism in Escherichia coli. Interaction with the fadB promoter is prevented by long chain fatty acyl coenzyme A.* J Biol Chem, 1992. **267**(12): p. 8685-91.

25. Rigali, S., et al., *Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies.* J Biol Chem, 2002. **277**(15): p. 12507-15.

26. Vindal, V., K. Suma, and A. Ranjan, *GntR family of regulators in Mycobacterium smegmatis: a sequence and structure based characterization.* BMC Genomics, 2007. **8**: p. 289.

27. Lee, M.H., et al., *PlmA, a new member of the GntR family, has plasmid maintenance functions in Anabaena sp. strain PCC 7120.* J Bacteriol, 2003. **185**(15): p. 4315-25.

28. Zhou, X., Q. Yan, and N. Wang, *Deciphering the regulon of a GntR family regulator via transcriptome and ChIP-exo analyses and its contribution to virulence in Xanthomonas citri.* Mol Plant Pathol, 2017. **18**(2): p. 249-262.

29. Ji, L. and J. Xie, *GntR family regulators of the pathogen of fish tuberculosis Mycobacterium marinum.* Biochem Biophys Res Commun, 2011. **410**(4): p. 780-5.

30. Suvorova, I.A., Y.D. Korostelev, and M.S. Gelfand, *GntR Family of Bacterial Transcription Factors and Their DNA Binding Motifs: Structure, Positioning and Co-Evolution.* PLoS One, 2015. **10**(7): p. e0132618.

31. Jain, D., *Allosteric control of transcription in GntR family of transcription regulators: A structural overview.* IUBMB Life, 2015. **67**(7): p. 556-63.

32. Zheng, M., et al., *Structure of Thermotoga maritima TM0439: implications for the mechanism of bacterial GntR transcription regulators with Zn2+-binding FCD domains.* Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 4): p. 356-65.

33. Resch, M., et al., *Insight into the induction mechanism of the GntR/HutC bacterial transcription regulator YvoA.* Nucleic Acids Res, 2010. **38**(7): p. 2485-97.

34. Gao, Y.G., et al., *The structures of transcription factor CGL2947 from Corynebacterium glutamicum in two crystal forms: a novel homodimer assembling and the implication for effector-binding mode.* Protein Sci, 2007. **16**(9): p. 1878-86.

35. Belitsky, B.R. and A.L. Sonenshein, *GabR, a member of a novel protein family, regulates the utilization of gamma-aminobutyrate in Bacillus subtilis.* Mol Microbiol, 2002. **45**(2): p. 569-83.

36. Edayathumangalam, R., et al., *Crystal structure of Bacillus subtilis GabR, an autorepressor and transcriptional activator of gabT.* Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17820-5.

37. Wu, R., et al., *PLP and GABA trigger GabR-mediated transcription regulation in Bacillus subtilis via external aldimine formation.* Proc Natl Acad Sci U S A, 2017. **114**(15): p. 3891-3896.

38. Richard, J.P., et al., *Pyridoxal 5'-phosphate: electrophilic catalyst extraordinaire.* Curr Opin Chem Biol, 2009. **13**(4): p. 475-83.

39. Theriault, O., et al., *Pyridoxal-5'-phosphate (MC-1), a vitamin B6 derivative, inhibits expressed P2X receptors.* Can J Physiol Pharmacol, 2014. **92**(3): p. 189-96.

40. El Qaidi, S., et al., *The vitamin B(6) biosynthesis pathway in Streptococcus pneumoniae is controlled by pyridoxal 5'-phosphate and the transcription factor PdxR and has an impact on ear infection.* J Bacteriol, 2013. **195**(10): p. 2187-96.

41. Bothra, S., et al., *Applications of vitamin B6 cofactor pyridoxal 5'-phosphate and pyridoxal 5'-phosphate crowned gold nanoparticles for optical sensing of metal ions.* Spectrochim Acta A Mol Biomol Spectrosc, 2017. **174**: p. 1-6.

42. Percudani, R. and A. Peracchi, *A genomic overview of pyridoxal-phosphate-dependent enzymes.* EMBO Rep, 2003. **4**(9): p. 850-4.

43. Bilski, P., et al., *Vitamin B6 (pyridoxine) and its derivatives are efficient singlet oxygen quenchers and potential fungal antioxidants.* Photochem Photobiol, 2000. **71**(2): p. 129-34.

44. Lambrecht, J.A., J.M. Flynn, and D.M. Downs, *Conserved YjgF protein family deaminates reactive enamine/imine intermediates of pyridoxal 5'-phosphate (PLP)-dependent enzyme reactions.* J Biol Chem, 2012. **287**(5): p. 3454-61.

45. John, R.A., *Pyridoxal phosphate-dependent enzymes.* Biochim Biophys Acta, 1995. **1248**(2): p. 81-96.

46. Jansonius, J.N., *Structure, evolution and action of vitamin B6-dependent enzymes.* Curr Opin Struct Biol, 1998. **8**(6): p. 759-69.

47. Bateman, A., P. Coggill, and R.D. Finn, *DUFs: families in search of function.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2010. **66**(Pt 10): p. 1148-52.

48. Steczkiewicz, K., et al., *Eukaryotic domain of unknown function DUF738 belongs to Gcn5-related N-acetyltransferase superfamily.* Cell Cycle, 2006. **5**(24): p. 2927-30.

49. Goodacre, N.F., D.L. Gerloff, and P. Uetz, *Protein domains of unknown function are essential in bacteria.* MBio, 2014. **5**(1): p. e00744-13.

50. Carpenter, E.P., et al., *Overcoming the challenges of membrane protein crystallography.* Curr Opin Struct Biol, 2008. **18**(5): p. 581-6.

51. Milano, T., et al., *A Bioinformatics Analysis Reveals a Group of MocR Bacterial Transcriptional Regulators Linked to a Family of Genes Coding for Membrane Proteins.* Biochem Res Int, 2016. **2016**: p. 4360285.

52. Milano, T., et al., *A Comprehensive Computational Analysis of Mycobacterium Genomes Pinpoints the Genes Co-occurring with YczE, a Membrane Protein Coding Gene Under the Putative Control of a MocR, and Predicts its Function.* Interdiscip Sci, 2018. **10**(1): p. 111-125.

53. Bramucci, E., T. Milano, and S. Pascarella, *Genomic distribution and heterogeneity of MocR-like transcriptional factors containing a domain belonging to the superfamily of the pyridoxal-5'-phosphate dependent enzymes of fold type I.* Biochem Biophys Res Commun, 2011. **415**(1): p. 88-93.

54. Belitsky, B.R., *Bacillus subtilis GabR, a protein with DNA-binding and aminotransferase domains, is a PLP-dependent transcriptional regulator.* J Mol Biol, 2004. **340**(4): p. 655-64.

55. Belitsky, B.R., *Role of PdxR in the activation of vitamin B6 biosynthesis in Listeria monocytogenes.* Mol Microbiol, 2014. **92**(5): p. 1113-28.

56. Han, Q., et al., *Crystal structures of Aedes aegypti alanine glyoxylate aminotransferase.* J Biol Chem, 2006. **281**(48): p. 37175-82.

57. Sakuraba, H., et al., *Structure of an archaeal alanine:glyoxylate aminotransferase.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 6): p. 696-9.

58. Zhang, X., et al., *Crystal structure of alanine:glyoxylate aminotransferase and the relationship between genotype and enzymatic phenotype in primary hyperoxaluria type 1.* J Mol Biol, 2003. **331**(3): p. 643-52.

59. Meyer, P., et al., *Crystal structure and confirmation of the alanine:glyoxylate aminotransferase activity of the YFL030w yeast protein.* Biochimie, 2005. **87**(12): p. 1041-7.

60. Noguchi, T., et al., *Characteristics of hepatic alanine-glyoxylate aminotransferase in different mammalian species.* Biochem J, 1978. **169**(1): p. 113-22.

61. Chen, X., L. Levine, and P.Y. Kwok, *Fluorescence polarization in homogeneous nucleic acid analysis.* Genome Res, 1999. **9**(5): p. 492-8.

62. Eliot, A.C. and J.F. Kirsch, *Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations.* Annu Rev Biochem, 2004. **73**: p. 383-415.

63. Smith, D.L., et al., *2.8-A-resolution crystal structure of an active-site mutant of aspartate aminotransferase from Escherichia coli.* Biochemistry, 1989. **28**(20): p. 8161-7.

64. Hayashi, H., H. Mizuguchi, and H. Kagamiyama, *Rat liver aromatic L-amino acid decarboxylase: spectroscopic and kinetic analysis of the coenzyme and reaction intermediates.* Biochemistry, 1993. **32**(3): p. 812-8.

65. Hayashi, H., et al., *Acid-base chemistry of the reaction of aromatic L-amino acid decarboxylase and dopa analyzed by transient and steady-state kinetics: preferential binding of the substrate with its amino group unprotonated.* Biochemistry, 1999. **38**(47): p. 15615-22.

66. Chu, W.C. and D.E. Metzler, *Enzymatically active truncated cat brain glutamate decarboxylase: expression, purification, and absorption spectrum.* Arch Biochem Biophys, 1994. **313**(2): p. 287-95.

67. Olmo, M.T., et al., *Spectroscopic analysis of recombinant rat histidine decarboxylase.* J Biochem, 2002. **132**(3): p. 433-9.

68. Okuda, K., et al., *Role of the aminotransferase domain in Bacillus subtilis GabR, a pyridoxal 5'-phosphate-dependent transcriptional regulator.* Mol Microbiol, 2015. **95**(2): p. 245-57.

69. Metzler, D.E. and C.M. Metzler, *Biochemistry : the chemical reactions of living cells*. 2nd ed. 2001, San Diego, Calif.: Harcourt/Academic Press.

70. Nasir, M.S. and M.E. Jolley, *Fluorescence polarization: an analytical tool for immunoassay and drug discovery.* Comb Chem High Throughput Screen, 1999. **2**(4): p. 177-90.

71. Amidani, D., et al., *Study of DNA binding and bending by Bacillus subtilis GabR, a PLP-dependent transcription factor.* Biochim Biophys Acta, 2017. **1861**(1 Pt A): p. 3474-3489.

72. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.

73. Marck, C., *'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers.* Nucleic Acids Res, 1988. **16**(5): p. 1829-36.

74. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode.* Methods Enzymol, 1997. **276**: p. 307-26.

75. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

76. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

77. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

78. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography.* Acta Crystallogr D Biol Crystallogr, 1994. **50**(Pt 5): p. 760-3.

79. Silverman, R.B. and M.A. Levy, *Mechanism of inactivation of gamma-aminobutyric acid-alpha-ketoglutaric acid aminotransferase by 4-amino-5-halopentanoic acids.* Biochemistry, 1981. **20**(5): p. 1197-203.

80. Goodacre, N.F., D.L. Gerloff, and P. Uetz, *Protein domains of unknown function are essential in bacteria.* MBio, 2013. **5**(1): p. e00744-13.

81. Mehta, P.K. and P. Christen, *The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes.* Adv Enzymol Relat Areas Mol Biol, 2000. **74**: p. 129-84.

82. Alexander, F.W., et al., *Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific alpha, beta and gamma families.* Eur J Biochem, 1994. **219**(3): p. 953-60.

83. Schneider, G., H. Kack, and Y. Lindqvist, *The manifold of vitamin B6 dependent enzymes.* Structure, 2000. **8**(1): p. R1-6.

84. Chen, P.J., et al., *Protein tyrosine phosphatase 1B (PTP1B): A key regulator and therapeutic target in liver diseases.* Toxicology, 2015. **337**: p. 10-20.

85. Reaven, G.M. and Y.D. Chen, *Role of insulin in regulation of lipoprotein metabolism in diabetes.* Diabetes Metab Rev, 1988. **4**(7): p. 639-52.

86. Bakke, J. and F.G. Haj, *Protein-tyrosine phosphatase 1B substrates and metabolic regulation.* Semin Cell Dev Biol, 2015. **37**: p. 58-65.

87. Byon, J.C., A.B. Kusari, and J. Kusari, *Protein-tyrosine phosphatase-1B acts as a negative regulator of insulin signal transduction.* Mol Cell Biochem, 1998. **182**(1-2): p. 101-8.

88. Ullrich, A. and J. Schlessinger, *Signal transduction by receptors with tyrosine kinase activity.* Cell, 1990. **61**(2): p. 203-12.

89. Glenney, J.R., Jr., *Tyrosine-phosphorylated proteins: mediators of signal transduction from the tyrosine kinases.* Biochim Biophys Acta, 1992. **1134**(2): p. 113-27.

90. Asante-Appiah, E. and B.P. Kennedy, *Protein tyrosine phosphatases: the quest for negative regulators of insulin action.* Am J Physiol Endocrinol Metab, 2003. **284**(4): p. E663-70.

91. Lund, I.K., et al., *Mechanism of protein tyrosine phosphatase 1B-mediated inhibition of leptin signalling.* J Mol Endocrinol, 2005. **34**(2): p. 339-51.

92. Seale, A.P., et al., *Development of an automated protein-tyrosine phosphatase 1B inhibition assay and the screening of putative insulin-enhancing vanadium(IV) and zinc(II) complexes.* Biotechnol Lett, 2005. **27**(4): p. 221-5.

93. Trischitta, V. and G.M. Reaven, *Evidence of a defect in insulin-receptor recycling in adipocytes from older rats.* Am J Physiol, 1988. **254**(1 Pt 1): p. E39-44.

94. Fantus, I.G. and E. Tsiani, *Multifunctional actions of vanadium compounds on insulin signaling pathways: evidence for preferential enhancement of metabolic versus mitogenic effects.* Mol Cell Biochem, 1998. **182**(1-2): p. 109-19.

95. Johnson, T.O., J. Ermolieff, and M.R. Jirousek, *Protein tyrosine phosphatase 1B inhibitors for diabetes.* Nat Rev Drug Discov, 2002. **1**(9): p. 696-709.

96. Du, Z.D., et al., *Protein tyrosine phosphatase 1B regulates migration of ARPE-19 cells through EGFR/ERK signaling pathway.* Int J Ophthalmol, 2015. **8**(5): p. 891-7.

97. Zinker, B.A., et al., *PTP1B antisense oligonucleotide lowers PTP1B protein, normalizes blood glucose, and improves insulin sensitivity in diabetic mice.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11357-62.

98. Zhang, Z.Y., *Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development.* Annu Rev Pharmacol Toxicol, 2002. **42**: p. 209-34.

99. Scior, T., et al., *Chimeric design, synthesis, and biological assays of a new nonpeptide insulin-mimetic vanadium compound to inhibit protein tyrosine phosphatase 1B.* Drug Des Devel Ther, 2010. **4**: p. 231-42.

100. Tracey, A.S. and D.C. Crans, *Vanadium compounds : chemistry, biochemistry, and therapeutic applications*. ACS symposium series,. 1998, Washington, DC Cary, NC: American Chemical Society ; Distributed by Oxford University Press. xiii, 381 p.

101. Mehdi, M.Z. and A.K. Srivastava, *Organo-vanadium compounds are potent activators of the protein kinase B signaling pathway and protein tyrosine phosphorylation: mechanism of insulinomimesis.* Arch Biochem Biophys, 2005. **440**(2): p. 158-64.

102. Irving, E. and A.W. Stoker, *Vanadium Compounds as PTP Inhibitors.* Molecules, 2017. **22**(12).

103. Choi, J.S., et al., *Protein tyrosine phosphatase 1B inhibitory activity of alkaloids from Rhizoma Coptidis and their molecular docking studies.* J Ethnopharmacol, 2015. **171**: p. 28-36.

104. Crans, D.C., et al., *The chemistry and biochemistry of vanadium and the biological activities exerted by vanadium compounds.* Chem Rev, 2004. **104**(2): p. 849-902.

105. Basuki, W., et al., *Enhancement of insulin signaling pathway in adipocytes by oxovanadium(IV) complexes.* Biochemical and Biophysical Research Communications, 2006. **349**(3): p. 1163-1170.

106. Crans, D.C., et al., *Vanadate monomers and dimers both inhibit the human prostatic acid phosphatase.* Biochem Biophys Res Commun, 1989. **165**(1): p. 246-50.

107. Huyer, G., et al., *Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate.* J Biol Chem, 1997. **272**(2): p. 843-51.

108. Mustafi, D. and M.W. Makinen, *Structure and conformation of bis(acetylacetonato)oxovanadium(IV) and bis(maltolato)oxovanadium(IV) in solution determined by electron nuclear double resonance spectroscopy.* Inorg Chem, 2005. **44**(16): p. 5580-90.

109. Bellomo, E., et al., *Zinc ions modulate protein tyrosine phosphatase 1B activity.* Metallomics, 2014. **6**(7): p. 1229-39.

110. van Montfort, R.L., et al., *Oxidation state of the active-site cysteine in protein tyrosine phosphatase 1B.* Nature, 2003. **423**(6941): p. 773-7.

111. Denu, J.M. and K.G. Tanner, *Specific and reversible inactivation of protein tyrosine phosphatases by hydrogen peroxide: evidence for a sulfenic acid intermediate and implications for redox regulation.* Biochemistry, 1998. **37**(16): p. 5633-42.

112. Brandao, T.A., A.C. Hengge, and S.J. Johnson, *Insights into the reaction of protein-tyrosine phosphatase 1B: crystal structures for transition state analogs of both catalytic steps.* J Biol Chem, 2010. **285**(21): p. 15874-83.

113. Zhang, Z.Y., *Protein-tyrosine phosphatases: biological function, structural characteristics, and mechanism of catalysis.* Crit Rev Biochem Mol Biol, 1998. **33**(1): p. 1-52.

114. Barford, D., A.J. Flint, and N.K. Tonks, *Crystal structure of human protein tyrosine phosphatase 1B.* Science, 1994. **263**(5152): p. 1397-404.

115. Stuckey, J.A., et al., *Crystal structure of Yersinia protein tyrosine phosphatase at 2.5 A and the complex with tungstate.* Nature, 1994. **370**(6490): p. 571-5.

116. Fauman, E.B., et al., *The X-ray crystal structures of Yersinia tyrosine phosphatase with bound tungstate and nitrate. Mechanistic implications.* J Biol Chem, 1996. **271**(31): p. 18780-8.

117. Groves, M.R., et al., *Structural basis for inhibition of the protein tyrosine phosphatase 1B by phosphotyrosine peptide mimetics.* Biochemistry, 1998. **37**(51): p. 17773-83.

118. Brandao, T.A., S.J. Johnson, and A.C. Hengge, *The molecular details of WPD-loop movement differ in the protein-tyrosine phosphatases YopH and PTP1B.* Arch Biochem Biophys, 2012. **525**(1): p. 53-9.

119. Stone, R.L. and J.E. Dixon, *Protein-tyrosine phosphatases.* J Biol Chem, 1994. **269**(50): p. 31323-6.

120. Zhang, Z.Y. and J.E. Dixon, *Protein tyrosine phosphatases: mechanism of catalysis and substrate specificity.* Adv Enzymol Relat Areas Mol Biol, 1994. **68**: p. 1-36.

121. Barford, D., Z. Jia, and N.K. Tonks, *Protein tyrosine phosphatases take off.* Nat Struct Biol, 1995. **2**(12): p. 1043-53.

122. Percival, M.D., K. Doherty, and M.J. Gresser, *Inhibition of phosphoglucomutase by vanadate.* Biochemistry, 1990. **29**(11): p. 2764-9.

123. Shechter, Y., *Insulin-mimetic effects of vanadate. Possible implications for future treatment of diabetes.* Diabetes, 1990. **39**(1): p. 1-5.

124. Torossian, K., D. Freedman, and I.G. Fantus, *Vanadate down-regulates cell surface insulin and growth hormone receptors and inhibits insulin receptor degradation in cultured human lymphocytes.* J Biol Chem, 1988. **263**(19): p. 9353-9.

125. Heyliger, C.E., A.G. Tahiliani, and J.H. McNeill, *Effect of vanadate on elevated blood glucose and depressed cardiac performance of diabetic rats.* Science, 1985. **227**(4693): p. 1474-7.

126. Goldfine, A.B., et al., *In vivo and in vitro studies of vanadate in human and rodent diabetes mellitus.* Mol Cell Biochem, 1995. **153**(1-2): p. 217-31.

127. Hemrika, W., et al., *From phosphatases to vanadium peroxidases: a similar architecture of the active site.* Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2145-9.

128. Barford, D., *Protein phosphatases.* Curr Opin Struct Biol, 1995. **5**(6): p. 728-34.

129. Cantley, L.C., Jr. and P. Aisen, *The fate of cytoplasmic vanadium. Implications on (NA,K)-ATPase inhibition.* J Biol Chem, 1979. **254**(6): p. 1781-4.

130. Degani, H., et al., *Electron paramagnetic resonance studies and insulin-like effects of vanadium in rat adipocytes.* Biochemistry, 1981. **20**(20): p. 5795-9.

131. Willsky, G.R., D.A. White, and B.C. McCabe, *Metabolism of added orthovanadate to vanadyl and high-molecular-weight vanadates by Saccharomyces cerevisiae.* J Biol Chem, 1984. **259**(21): p. 13273-81.

132. Peters, K.G., et al., *Mechanism of insulin sensitization by BMOV (bis maltolato oxo vanadium); unliganded vanadium (VO4) as the active component.* J Inorg Biochem, 2003. **96**(2-3): p. 321-30.

133. Guan, K.L. and J.E. Dixon, *Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate.* J Biol Chem, 1991. **266**(26): p. 17026-30.

134. Coulston, L. and P. Dandona, *Insulin-like effect of zinc on adipocytes.* Diabetes, 1980. **29**(8): p. 665-7.

135. Claiborne, A., et al., *Protein-sulfenic acids: diverse roles for an unlikely player in enzyme catalysis and redox regulation.* Biochemistry, 1999. **38**(47): p. 15407-16.

# VITA

Yuanzhang Zheng was born and raised in Zoucheng, Shandong Province, China. He received his Bachelor degree of Biological technology from School of Life Science, Qingdao Agriculture University in 2009. He received his Master of Science in Zoology from School of Life Science, Lanzhou University in 2012. In 2012, he began the PhD program in Chemistry and Biochemistry at Loyola University Chicago.