6-26-2017

Biochemical characterization of ArsI: a novel C-As lyase for degradation of environmental organoarsenicals

Shashank Suryakant Pawitwar
spawi001@fiu.edu

DOI: 10.25148/etd.FIDC001926

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

BIOCHEMICAL CHARACTERIZATION OF ARSI: A NOVEL C-As LYASE FOR DEGRADATION OF ENVIRONMENTAL ORGANOARSENICALS

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

In

BIOMEDICAL SCIENCES

By

Shashank S Pawitwar

2017

To: Dean John A. Rock
Herbert Wertheim College of Medicine

This dissertation, written by Shashank S Pawitwar, and entitled Biochemical Characterization of ArsI: A Novel C-As Lyase for Degradation of Environmental Organoarsenicals, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Yong Cai

_____
Konstantinos Kavallieratos

_____
Jeremy Chambers

_____
Masafumi Yoshinaga, Co-Major Professor

_____
Barry P. Rosen, Co-Major Professor

Date of Defense: June 26, 2017

The dissertation of Shashank S Pawitwar is approved.

_____
Dean John A. Rock
Herbert Wertheim College of Medicine

_____
Andres G. Gil
Vice president for Research and economic development
And Dean of University Graduate School

Florida International University, 2017

DEDICATION

To my Family and friends for their support

ACKNOWLEDGMENTS

This dissertation is the result of many people and I am indebted to all them. I owe my deepest gratitude to my supervisor, Dr. Barry P Rosen, whose encouragement, guidance and support from the beginning to the end have enabled me to develop my project and how to approach a problem and look for a plausible solution. It has been a great learning experiencing under his mentorship. I am very grateful for all the scientific discussions and guidance I have received from Dr. Masafumi Yoshinaga, Co-advisor. I would also like to thank him for the countless hours he patiently helped me to figure out and fix the problems I had with the difficult experiments and in writing. I am also heartily thankful for my past and current committee members, Dr. Hiranmoy Bhattacharjee, Dr. Xiaotang Wang, Dr. Chambers, Dr. Konstantinos Kavallieratos and Dr. Yong Cai who assisted me in the right direction of my graduate research. I wish to thank Dr. Jitesh Pillai for his help during the initial stage of my Ph.D. with many techniques particularly ITC. I wish to recognize Dr. Charles Packianathan for help with protein purification. The model studies on BmArsI collaboration with Dr. Venkadesh Sarkarai Nadar, who solved TcArsI structure. Dr. Dharmendra S. Dheeman, Dr. Jian Chen and Dr. Hui Dong with analytical techniques and their sound advice. I would also like to thank all my present and past laboratory colleagues: Dr. Kavitha Marapakala, Dr. Yu Yan, Dr. Thiruselvam Viswanathan, Ms. Jiaojiao Li, Mr. Luis Garbrinski. I want to express my thanks to all the faculties and staffs in our department who continuously supported me with

advice and friendship. My special regards to Dr. Daniela Maizel for being a fabulous friend. My regards to Mr. Gajanan Bhalerao, Dr. Saurabh Joshi and Dr. Samihar Khandre for being great friends. My thanks to Dr. Vidya Sagar, Dr. Deepak Balasubramanian, Dr. Pratik Nyati, Ms. Parul Maheshwari, Dr. Mansi Sharma, Dr. Priyanka Kushwaha, Dr. Ketaki Deshpande, Dr. Gorakh Tatke, Dr. Rahul Dev Jayant, Dr. Ajeet Kaushik, Dr. Arti Vashishtha, Dr. Pandiya Raj, Ms. Supurna Dhar and Mr. Mayur Doke for their friendship and peer support.

Lastly, I am grateful to my parents Mr. Suryakant V Pawitwar, Mrs. Shailaja S Pawitwar, my brother and sister-in-law Mr. Vithal S Pawitwar and Mrs. Amruta Vithal Pawitwar for their unwavering support and inspiration, my dear sister Mrs. Supriya M Madrewar and brother-in-law Mr. Mayur Madrewar for her faith and confidence in me, and my nephew Mr. Shreeyan V Pawitwar for being stress buster during my Ph.D. journey.

ABSTRACT OF THE DISSERTATION

BIOCHEMICAL CHARACTERIZATION OF ARSI: A NOVEL C-As LYASE FOR DEGRADATION OF ENVIRONMENTAL ORGANOARSENICALS

by

Shashank S Pawitwar

Florida International University, 2017

Miami, Florida

Professor Barry P. Rosen, Co-Major Professor

Dr. Masafumi Yoshinaga, Co-Major Professor

Arsenic is considered by the U.S. Environmental Protection Agency (EPA) to be the most prevalent environmental toxin. Pentavalent organoarsenicals such as MSMA (monosodium methylarsenate) and roxarsone (4-hydroxy-3-nitrophenylarsonic acid) are currently used as herbicide and growth enhancers in animal husbandry, respectively. They undergo environmental degradation to more toxic inorganic arsenite (As(III)) that contaminate crops and drinking water supplies. Recently, our laboratory has identified a two-step pathway of degradation of MSMA to As(III) by microbial communities. First step includes reduction of MSMA to methylarsonous acid [MAs(III)] by some bacterial species. The second step is demethylation of MAs(III) to As(III) by other bacteria. Our laboratory identified the gene (*arsI*) responsible for MAs(III) demethylation from an environmental isolate, *Bacillus* sp. MD1. *arsI* encodes a non-heme iron-dependent dioxygenase that cleaves the carbon-arsenic bond. Purified ArsI catalyzes Fe(II)-dependent demethylation of the trivalent forms of MSMA and

roxarsone. ArsI is the first C-As lyase to be identified and the first enzyme shown to be involved in organoarsenical degradation. My objective was to characterize the mechanism of the ArsI reaction. To investigate the role of specific residues in ArsI demethylation catalysis reaction, amino acid residues lining the Fe(II)-binding site, and the substrate binding site have been altered by site-directed mutagenesis. The altered proteins purified, and their biochemical properties evaluated using a combination of fluorometry, isothermal titration calorimetry (ITC) and other biophysical techniques. Protein fluorescence studies of wild-type ArsI showed a greater affinity for the trivalent roxarsone. ArsI affinity varies with different trivalent substrate in the order Rox(III) > pASA(III) > PhAs(III) > MAs(III). The K_d values for Fe(II), PhAs(III), MAs(III) and As(III) determined by ITC are in the micro to nanomolar range. Rox(III) has a characteristic optical spectral change (blue shift) as a function of time in the presence of ArsI. ESI-MS data demonstrate that ArsI cleaves C-As bond and incorporates an atom of molecular oxygen into each carbon and arsenic of the target bond. It results in the formation of 4-hydroxy-2-nitrophenol as carbon product after Rox(III) degradation by ArsI. These data will lead to elucidation of the mechanism of ArsI catalysis, augmenting our understanding how microbes remodel the environment through biotransformation of organoarsenicals and complement our understanding of the arsenic biogeochemical cycle.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF FIGURES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

ABBREVIATION	FULL NAME
AQP	Aquaporin
ArsA	Arsenical ATPase
ArsB	Arsenical pump membrane protein
ArsD	Arsenical metallochaperon
ArsM	As(III) S-adenosylmethionine methyltransferase
ArsP	Arsenical membrane permease
ArsR	Arsenical repressor
As(III)	Trivalent arsenite
As(V)	Pentavalent arsenate
ATP	Adenosine triphosphate
ATSDR	Agency for Toxic Substances and Disease Registry
DMAs(III)	Trivalent dimethyl arsenite
DMAs(V)	Dimethylarsenate
EDTA	Ethylenediaminetetraacetic acid
EPA	Environmental Protection Agency
GlpF	Aquaglyceroporin
GLUT	Glucose permease
Grx	Glutaredoxin
MAs(III)	Trivalent methyl arsenite
MAs(V)	Pentavalent methylarsenate

MBP	Maltose binding protein
MOPS	3-(N-morpholino) propane sulfonic acid
Nitarsone	4-nitrophenyl arsonic acid
Roxarsone	4-hydroxy-3-nitrobenzenearsonic acid
TMAs(III)	Trivalent methyl arsenic
TMAs(V)O	Trimethylarsine oxide
HILIC	Hydrophilic Interaction Chromatography
NMR	Nuclear Magnetic Resonance
ESI-MS	Electron Spray Ionization Mass Spectrometry
ITC	Isothermal Titration Calorimetry

CHAPTER 1 INTRODUCTION

1.1 Arsenic: Evolution, Natural, and Anthropogenic Sources

Life arose when the atmosphere was neutral, and toxic metal ion concentrations were almost certainly higher than in present-day oceans (Zhu et al. 2014). The toxic metalloid arsenic is omnipresent in the environment, primarily from geothermal sources and bedrock. Arsenic exposure in Bangladesh and West Bengal in India by groundwater contaminated from bedrock is known as the arsenic largest mass poisoning in history (Smith et al. 2000). In the United States, a recent study on the high levels of arsenic in soil in the state of Ohio found that geological and soil processes are primarily responsible (Venteris et al. 2014).

1.2 Arsenic Distribution, Toxicity and Health Effects

The U.S. Environmental Protection Agency (EPA) and Agency for Toxic Substances and Disease Registry (ATSDR) consider arsenic to be the most prevalent environmental toxin and rank at the top of the US Priority List of Hazardous Substances (https://www.atsdr.cdc.gov/SPL/index.html). Because of its ubiquitous presence, arsenic enters to drinking water and food supply in many ways, including natural sources and anthropogenic activities such as mining, smelting, wood preservatives, herbicides, pesticides and in the poultry industry as antimicrobial growth promoters. The EPA ranks arsenic as the number 1 environmental toxic substance on the spell out (ASTDR) (web site), and the International Agency for Research on Cancer (IARC) classify arsenic as a group

1 human carcinogen . Consequently, exposure to arsenic causes many health problems including cancer, cardiovascular disease (States et al. 2009), diabetes (Drobna et al. 2013), hearing loss (Bencko et al. 1977), peripheral neuropathy (Mukharjee et al. 2003) and retarded neurological development (Tyler et al. 2014). Chronic arsenic exposure (ingestion and inhalation) can cause cancers of the bladder, liver, prostate, skin, lung, and angiosarcoma (Falk et al. 1981). The strongest association between chronic arsenic exposure and cancer is for skin, lung, and bladder cancer (IARC 2004). In Chile, arsenic-contaminated drinking water was associated with an increase of vasospastic changes (Raynaud's disease) and thickening of the small and medium-sized arteries in autopsied children (Garcia-Vargas et al. 1996). In high amounts arsenic is poisonous. For example, in nineteenth century Northern England, around 6000 individuals were poisoned by an arsenic-contaminated beer that caused 71 deaths (Dyer 2009; Copping et al. 2009). The U.S. government set the safe drinking water standard to 0.01 mg/L as maximum contamination level (MCL), which was lowered from the old level of 0.05 mg/L in 2002 by the EPA. Currently, more than 6 million people in the United States are exposed to 0.01mg/L arsenic through drinking water and 2.5 million above the MCL. Globally, 200 million people are exposed to arsenic levels above the EPA's MCL and the World Health Organization's (WHO) recommended maximal level (Ravenscroft 2009; WHO).

1.3 Arsenic Resistance Genes in *ars* Operon

To survive in the presence of arsenic, the most ubiquitous of all environmental toxins, the earliest cells must have had mechanisms for detoxifying arsenic. This intense selective pressure was the driving force for the evolution of bacterial arsenic resistance (*ars*) genes and pathways that are found in almost every extant organism. Without arsenic detoxification systems, life could not have arisen and would not exist today.

At a minimum *ars* operons encode an As(III)-responsive repressor (ArsR), which senses environmental inorganic arsenic (San Francisco et al. 1990) and one of two As(III) efflux permeases, ArsB (San Francisco et al. 1988) or Acr3 (Wysocki et al. 1997; Ghosh et al. 1999). In addition to the efflux system, there are a variety of *ars* gene products that expand the range or capacity for inorganic arsenic. The most common is ArsC, which is an arsenate reductase that transforms inorganic As(V) to As(III), the substrate of the efflux pumps (Mukhopadhyay and Rosen 2002; Rosen et al. 1988; Lin et al. 2006;). The ArsA As(III)-stimulated ATPase (Hsu 1999) and the ArsD As(III) metallochaperone (Lin) enhance the ability of ArsB to extrude As(III) and increase resistance.

Phosphate transporter systems are the major route of pentavalent arsenate uptake into most cells. Aquaglyceroporins such as GlpF (E. coli), AQP1 (leishmanial) and AQP9 (human) catalyze trivalent arsenite uptake in most cells (Liu et al. 2002, 2006 a,b; Bhattacharjee et al. 2004). These are all mechanisms for sensing, transforming and transporting inorganic arsenic.

All these genes were shown to be responsible for resistance to inorganic arsenic. In addition to inorganic arsenic, organic arsenicals are also present in the environment and used widely as herbicides and pesticides because pentavalent organoarsenicals have lower toxicity to humans than inorganic arsenic.. However, organic arsenicals are not as prevelant in the environment as inorganic arsenic. Is it because of degradation of organic arsenic into inorganic arsenic? Are there any specific genes responsible for its degradation or biotransformation? Only recently has the existence of a parallel cycle for biotransformations and detoxification of organoarsenicals been recognized. *ars* genes responsible for detoxification of organoarsenicals include *arsM, arsH*, *arsP*, and *arsI.*

The *arsM* gene encodes an As(III) S-adenosylmethionine methyltransferase (Qin, Rosen et al. 2006; Chen, Bhattacharjee et al. 2015; Chen, Madegowda et al. 2015), *arsI* encodes a non-heme iron-dependent C-As lyase, organoarsenical oxidase by *arsH*, and *arsP* encodes a methylarsenite efflux permease (Fig 1).

1.4 Arsenic Biotransformation and the Biogeochemical Cycle

Biotransformations of arsenic are carried out by many microorganisms, including bacteria and fungi. The cyanobacterium *Nostoc* sp. PCC 7120 has been shown to carry out both methylation and demethylation of MAs(III) by *arsM* and *arsI* genes, respectively (Yan et al. 2017). *Mycobacterium neoaurum* and *Candida humicola* have also been demonstrated to have demethylation activity

(Lehr et al. 2013; Cullen and McBride et al. 1979). *Mycobacterium neoaurum* can demethylate about 500 ppb MSMA into inorganic arsenic and utilize MSMA as a carbon source. Pentavalent organoarsenicals such as MSMA (monosodium methylarsenate or MAs(V)), roxarsone (4-hydroxy-3-nitrophenylarsonic acid) (Rox(V)) and phenylarsenate (PhAs(V)) are currently used as herbicides and growth enhancers in animal husbandry, respectively. They undergo environmental degradation to more toxic inorganic arsenite (As(III)) that contaminates crops and drinking water supplies (Fan et al. 2009, Yoshinaga, and Rosen, 2014). The half-life of MSMA was determined to be about 22 days, but it was still detectable in the soil after 1.5 years (Woolson et al. 1982). Moreover, a quantitative study of DMAs shows that demethylation is far more important than transformation into gaseous arsines (Gao and Burau 1997). Two *Pseudomonas putida* species capable of MSMA demethylation were isolated from the chemical warfare agent-contaminated soil of Ohkunoshima Island in Japan (Maki et al. 2006). Thus, there are many reports of bacterial demethylation of organoarsenicals. Recently, our laboratory has identified a two-step pathway of degradation of MSMA to As(III) by microbial communities. First step includes reduction of MSMA to methylarsonous acid [MAs(III)] by some bacterial species. The second step is demethylation of MAs(III) to As(III) by other bacteria (Yoshianga et al. 2011). We also cloned the *arsI* gene responsible for MAs(III) demethylation from an environmental isolate, *Bacillus* sp. MD1. *arsI* encodes a non-heme iron-dependent extradiol dioxygenase that cleaves the carbon-arsenic bond.

1.5 Dioxygenases: History, Structure, and Function

Respiration is an essential activity in living organisms. In aerobic respiration, oxygen plays an important role as a terminal electron acceptor that yields ATP, the major energy currency in a cell. In addition to oxidative phosphorylation, dioxygen serves as a source of oxygen for the synthesis of many essential molecules such as aromatic amino acids, neurotransmitters, and steroids, to name an only few. Additionally, it is involved in the synthesis of vital biomolecules, for instance, synthesis of the antibiotic penicillin. Oxygen is required not only in biosynthesis but also in biodegradation of biomolecules; oxygen is incorporated into the substrate to carry out degradation reaction(s). Also, molecular oxygen serves as a hydrogen acceptor in biological oxidation processes, where oxygen is either reduced to water or hydrogen peroxide. Oxidase enzymes are involved in these process(es), whereas oxygenases are more likely to be involved in converting nutrients into biologically important substances (Kovaleva and Lipscomb 2008; Nozaki 1979).

In 1955, two groups of investigators discovered oxygenases independently. The Mason group discovered phenolase, which carries out oxidation of 3,4-dimethylphenol to dimethyl catechol. At the same time, the Hayaishi group found pyrocatechase, which incorporates two atoms of molecular oxygen into catechol. In both cases, oxygen atoms are incorporated into the substrate. Therefore, these enzymes were termed “oxygenases” (Fetzner 2012). There are two types of oxygenases, monooxygenases, and dioxygenases.

Monooxygenases carry out substrate oxidation by incorporating one atom of oxygen into the substrate and the other atom to form water, whereas dioxygenases incorporate both oxygen atoms into the substrate. Both oxygenases required cofactors for catalysis (Armstrong 2000).

Microorganisms have been shown to utilize ring-cleaving dioxygenases to degrade both aromatic and heterocyclic compounds (Fetzner 2012). Dioxygenases add two oxygen atoms into a substrate to catalyze the oxidative degradation of a double bond of aromatic compounds. Ring-cleaving dioxygenases are divided into two groups, intradiol and extradiol dioxygenases, based on scission of the aromatic ring. Intradiol enzymes use a non-heme Fe(III) to carry out cleavage of the carbon-carbon [C-C] bond between two hydroxyl groups (ortho-cleavage). Extradiol dioxygenases use a non-heme Fe(II) to cleave the aromatic ring between a hydroxylated carbon and an adjacent non-hydroxylated carbon (meta-cleavage) (Costas et al. 2004).

There are no significant sequence or structural similarities between intradiol and extradiol dioxygenases. They are evolutionarily distinct classes of proteins. Intradiol enzymes belong to a single evolutionary family, and different enzymes in this family share a common structural fold (Vettting and Ohlendorf 2000). There are three evolutionarily independent lineages for extradiol and extradiol-type dioxygenases (Eltis and Bolin 1996; Dunwell et al. 2001).

1) Type I extradiol dioxygenases form the vicinal oxygen chelate (VOC) superfamily (Armstrong, 2000; Gerlt and Babbitt, 2001).

2) Type II extradiol dioxygenases are multimers with one or two different subunits (Vaillancourt et al. 2006).

3) The third family of extradiol-type dioxygenases is the cupin superfamily (Dunwell et al. 2001).

1.5.1 Extradiol Dioxygenases

A metalloenzymatic superfamily member carries out different sets of reactions with distinct transition states and yet share a more common mechanism. The vicinal oxygen chelates (VOC) superfamily is made up of structurally-related proteins with a paired βαβββ motif which provides two or three readily accessible metal coordination environments. These coordination sites provide metals with direct electrophilic participation in catalysis. There are five distinct members of the VOC superfamily classified based on functional analysis, three-dimensional structure, and sequence alignment. The extradiol dioxygenases are one of the members (Armstrong, 2000). Extradiol dioxygenases typically act on two hydroxyl groups of an aromatic substrate by ring fission between carbon-carbon bonds proximal to one of the hydroxyl groups. Catechol 2,3-dioxygenase [metapyrocatechase] was the first aromatic-ring-cleavage dioxygenase obtained in a crystalline form and is encoded by the *xylE* gene present on the TOL plasmid (Nozaki 1979; Haryama et al. 1992). The primary structure shows that it has four identical 32 kDa subunits, and each contains a catalytically required Fe(II) site. Although metapyrocatechase is colorless, with no absorption in the visible range, its reaction product gives an

intense yellow color. So, the *xylE* gene was used as a marker and reporter in molecular biology (Hahn et al. 1991).

In brief, the catechol 2,3- dioxygenase mechanism of the enzyme reaction is as (follows):

1) The substrate catechol binds to the enzyme metapyrocatechase.

2) Dioxygen binds to the enzyme-substrate complex and forms a ternary complex.

3) The catechol substrate is broken down into the product 2-hydroxymuconic semialdehyde (Nozaki, 1979).

In addition to catechol, metapyrocatechase can oxidize derivatives of catechol such as methyl catechol, ethyl catechol, and chloro- catechol. Magnetic circular dichroism in the near-infrared spectrum showed a ferrous electronic structure in the active site of metapyrocatechase (Mabrouk et al. 1991). The substrate catechol acts as a bidentate ligand and cause(s) conformational changes upon binding to metapyrocatechase (Fetzner, 2012). Although the enzyme requires Fe(II), a few Mn(II) dependent extradiol dioxygenases have also been identified (Costas et al. 2004).

The proposed mechanism of extradiol catechol dioxygenase and is based on the following structural and spectroscopic data (Fig. 2):

A) The substrate binds bidentate to the enzyme by displacing two water molecules and forms a five coordinate Fe(II) species with monoanionic catecholate (Fig. 2B). These changes likely decrease the redox potential of the iron(II) center, leading to reactivity with molecular oxygen.

B) As a result, oxygen binds to metal center and transfers an electron from (the) metal to oxygen, which results in the formation of a Fe(III)-superoxide complex (Fig. 2C).

C) Transfer of one more electrons from the substrate to metal follows, forming a semiquinonatoiron (II)-superoxide species (Fig. 2D).

D) An alkyl peroxy intermediate is generated after a nucleophilic attack of the superoxide on the aromatic ring (Fig. 2E).

E) A Criegee rearrangement of the formed alkyl peroxy intermediate forms a seven-membered α-oxolactone ring (Fig. 2F) that produces a yellow 2-hydroxymuconic semialdehyde product (Kovaleva and Lipscomb, 2008; Fetzner, 2012).

1.5.2 Type I Extradiol Dioxygenases

The vicinal oxygen chelates (VOC) superfamily consists of two copies of βαββββ structure (Armstrong, 2000; Gerlt and Babbitt, 2001). The α helix and β sheet are the most prominent types of protein secondary structure (Nelson and Cox 2008). There are five members of VOC superfamily: 4-hydroxyphenylpyruvate dioxygenase, fosfomycin resistance proteins, type I

extradiol dioxygenases, glyoxalases I and methyl malonyl-CoA epimerases (Serre et al. 1999; Bernat, 1997; Cameron et al. 1997; McCarthy et al. 2001). Most members utilize divalent cations to facilitate metal ion chelation of the substrate or an intermediate by adjacent oxygens in the reaction. The type I dioxygenases share a single-domain ancestor based on a phylogenetic study. Thus, it appears that two duplication events in the evolution of type I extradiol dioxygenases resulted in one and two-domain enzymes. Later the two-domain family was sub-divided into families based on substrate specificity. For instance, 2,3-dihydroxybiphenyl 1,2-dioxygenase belongs to a family that prefers bicyclic substrates, whereas catechol 2,3-dioxygenase belongs to a family with a preference for monocyclic substrates (Eltis and Bolin 1996; Harayama and Rekik,1989; Kita et al. 1999).

The bleomycin resistance proteins are a sixth family within the VOC superfamily (Dumas P et al. 1994). It differs from the other members of the superfamily as it does not bind a metal ion and accumulates bleomycin and related compounds. It neither transforms nor degrades bleomycin-related compounds. It is believed that to accommodate the antibiotic, the metal ions may have been shed during the evolution of this family (Vaillancourt et al. 2006).

1.5.3 Type III Extradiol Dioxygenases

The cupin superfamily is a group of structurally related proteins that share a β-barrel fold ('*cupa*' is the Latin word for a small barrel). These metalloproteins show either monocupin or bicupin metal-binding sites in a monomer. The metal

binding site is composed of residues from both monomers of a homodimer. Cupin enzymes have two conserved motifs based on widespread sequence analysis. For motif 1, it is $G(X)_5HXH(X)_{3,4}E(X)_6G$, and in motif 2 the characteristic conserved sequence is $G(X)_5PXG(X)_2H(X)_3N$ (Dunwell et al. 2001). These motifs are distinct. The first motif consists of the first two β sheets and the last two β sheets in motif 2. The two middle β strands and the lesser conserved loop play a major role in different classes of cupin superfamily. Some proteins in this family consist of a single domain, for instance, 3-hydroxyanthranilate dioxygenase and oxalate oxidase. These proteins function against stress and pathogens in plants. Other cupin superfamily proteins consist of two copies of the domain, which occurred most likely by a gene duplication process. Gentisate 1,2-dioxygenase and oxalate decarboxylase are examples of a protein family with two copies of the domain (Dunwell et al. 2000, 2008).

Cupin superfamily proteins use different metal ions such as iron in gentisate 1,2-dioxygenase, cysteine dioxygenase, and 1-hydroxy-2-naphthoate dioxygenase. In contrast oxalate oxidase and germin-like proteins utilize manganese as a metal ion; phosphomannose isomerase contains zinc, and nickel is utilized by acireductone dioxygenase (Harpel and Lipscomb 1990; Iwabuchi and Harayama, 1998; Anand et al. 2002; Cleasby et al. 1996; Pochapsky et al. 2002).

1.5.4 Structural Aspects of Extradiol Dioxygenases

Information about crystal structure of Fe(II)-dependent extradiol dioxygenases in the Fe(II) active form is available. C23O and HPCD are two examples of crystal structures of Fe(II)-dependent extradiol dioxygenases (Kita et al. 1999; Vetting et al. 2004). These extradiol dioxygenases share similar active site architecture and use iron. The structural motif of Fe(II)-dependent extradiol dioxygenases is composed of a triad of two histidine residues and one carboxylate (Titus et al. 2000). Although there are few other conserved residues based on sequence alignment, through examination will be necessary to determine their role in the catalysis.

1.5.5 Type I Extradiol Dioxygenases

Type I extradiol dioxygenases are comprised of a single type of subunit in all identified enzymes to date. A typical size of subunits of the one and two-domain proteins is 21 and 32.5 kDa respectively. To date, most of type I enzymes known are two-domain, implying an advantage of catalytically inactive N-domain to the host. The N-terminal domains of type I enzymes have been shown not to be directly involved in type I enzyme catalysis because the enzymes are still active without an N-terminal domain. Vaillancourt and colleagues measured steady-state catalysis of 2,3-dihydroxybiphenyl 1,2-dioxygenase and showed that the two-domain enzyme is more catalytically efficient than the one-domain enzyme (Vaillancourt et al. 2003). This suggests that the N-domain may contribute by providing additional stability to the C-

domain, resulting in a more efficient active site. Again, additional characterization is required to examine this question. The two-domain enzymes exist in a range of oligomeric states. For instance, 2,3-dihydroxybiphenyl 1,2-dioxygenase and catechol 2,3-dioxygenase are octameric and tetrameric respectively, whereas 2,2′,3-trihydroxybiphenyl 1,2-dioxygenase is monomeric (Kita et al. 1999; Han et al. 1995; Happe et al. 1993). In contrast, the one-domain proteins are normally dimeric. The physiological significance of these different oligomeric states is yet unclear. Han and colleagues solved the crystal structure of substrate-free 2,3-dihydroxybiphenyl 1,2-dioxygenase [DHBD] with a resolution of 1.9 Å.

Each monomer is composed of one chain of 297 residues and has two domains with similar structures. The Fe(II) metal ion is located in the C-domain. The orange sphere in Figure 3 represents Fe(II), the sticks represents the two histidine and glutamate triad metal binding site. Two βαββββ modules make one domain. The N-domain and C-domain contain modules 1 and 2, and modules 3 and 4 respectively. The N- and C-domain has a funnel shape space. A ferrous ion lies within a funnel-shaped space in the C-domain. The central funnel of the C-domain is slightly larger than that of the N-domain. Therefore, the inactive site within the N-domain of the two domain proteins could be an evolutionary variation (Eltis and Bolin,1996; Han et al. 1995).

The ferrous ion active site is in the 20 Å long funnel of the C-terminal domain. Both ends of this funnel are open. The large opening is 10 Å wide, and the small opening is 6 Å wide. Thus, the ferrous ion is accessible to catechol

substrates only from the wide opening; however, from either end, water or O_2 can access the iron. The iron shows a well-defined square pyramid coordination geometry, with His146 as the axial ligand, and His210, Glu260, and two water molecules as equatorial ligands in the basal plane (Fig 3A). Additionally, spectroscopic studies of substrate-free DHBD also provide confirmation for the five-coordinate, square pyramidal geometry of the iron (Eltis and Bolin, 1996; Davis et al. 2003). DHBD crystal structures with various catechol substrates show that substrate binds to the iron inside the funnel-shaped cavity. The catechol substrate is highly complementary in both size and shape. One hydroxyl group of the substrate binds by displacing the two ordered water ligands in trans to His 146 and His 210 (Fig 3B).

Interestingly, water forms a bridge between Asp244 and the 3-hydroxyl group. In some of the enzyme-substrate [ES] complexes, hydrogen-bonded water bridges between His195 and the iron. Moreover, the DHBD binds its substrate dihydroxy biphenyl complex in a 5-coordinate structure, as indicated by spectroscopic studies. The ternary complex containing DHBD bound to DHB and NO (DHNB: DHB: NO) represents the site that is more likely for O_2 binding (Davis et al. 2003; Vaillancourt et al. 2002; Sato et al. 2002).

The crystal structure of two homoprotocatechuate 2,3-dioxygenase (HPCD) have been reported. Although there is 83% amino acid sequence identity between these enzymes, they use different metals. The HPCD from *Arthrobacter globiformis* and *Brevibacterium fuscum* utilizes Mn(II) and Fe(II), respectively.

These enzymes possess two histidines and one glutamate triad in the active site. However, depending on the crystal form of the enzymes, metals have two or three solvent ligands. The ES complex in homoprotocatechuate binds to the metal in a bidentate fashion. The acetate moiety of the substrate interacts with the conserved residues. Unlike another type I dioxygenases, a lid domain does not close over the substrate binding. The structures do not reveal the basis for metal selectivity (Eltis and Bolin, 1995; Vetting et al. 2004).

Sequence alignments reveal that the C-domain of type I extradiol dioxygenases and one-domain enzymes are structurally similar. A few one-domain enzymes are larger than the typical C-domain of two-domain enzymes by approximately 65 residues. The structure and function of these residues are unidentified.

Sequence alignment of type I extradiol dioxygenases sequence suggest that they do not possess an active site in the N-terminal domain. However, in the evolution of hydroquinone dioxygenase (HQO), a module-exchange event may have occurred. Precisely, the first iron ligand in type I dioxygenases, histidine is not conserved in HQO at the first β - strand of module 3 (His146 in DHBD). However, a histidine is conserved at the first β -strand of module 1 in a similar position. It is possible that the two domains of the HQO subfamily of type I enzymes comprise modules 1 and 4 which contains the active site and modules 2 and 3, respectively (Eltis and Bolin, 1995).

1.5.6 Type II Extradiol Dioxygenases

In type II extradiol dioxygenases, the structures of the ferric form of protocatechuate 4,5-dioxygenase as the free-enzyme and in enzyme complex with the protocatechuate substrate structures have been solved. The larger β subunit has 302 amino acids and is composed of 11 β stands, 9 α helices, and one 3_{10} helix. The smaller α subunit contains 139 residues that form 10 α-helices. The α2/β2 dimer stabilized by the α–β contacts (Sugimoto et al. 1999).

As mentioned in a previous section, type II extradiol dioxygenases have been characterized as multimers with one or two different types of the subunit. For example, 2,3-dihydroxyphenylpropionate 1,2-dioxygenase is a homooligomers, whereas protocatechuate 4,5-dioxygenase (PCD) and 2-aminophenol 1,6-dioxygenases (APD) are heterodimers composed of $\alpha_2\beta_2$ (Spence et al. 1996).

The α subunit covering the active site is in a cleft in the β subunit. Thus, the catalytic iron is buried approximately 15 Å from the surface of the enzyme. The iron is coordinated in a distorted trigonal pyramidal geometry by two histidine residues (β12 and β61)-1-Glu (β242) triad and one water molecule in the substrate-free form of the enzyme. Asn β59, a potential weak fifth ligand is located trans to the water at 2.9 Å. Binding of the substrate (protocatechuate) displaces the water and involves both hydroxyl groups.

1.5.7 Non-heme Fe(II) Oxygenases and 2-His-1-Carboxylate Motif

Enzymes employ different strategies to activate oxygen and aromatic compounds during catalysis, and the use of a metal center in the active site is a well-known mechanism. In various mononuclear non-heme Fe(II) oxygenases, a facial triad 2-his-1-carboxylate forms the canonical metal-coordinating motif in the catalytic center. Enzymes with a facial triad of 2-his-1-carboxylate include catechol-1,3-dioxygenases, Rieske dioxygenases, and ketoacid dependent dioxygenases. The 2-his-1-carboxylate facial triad is a well-known motif for Fe(II) active sites that carry out oxygen dependent oxygenation (Straganz and Nidetzky, 2006).

1.5.8 Specificity of Extradiol Dioxygenase for Divalent Cations in Catalysis

Although many members of the VOC superfamily can use more than one divalent cations to carry out a catalysis reaction, a few members such as extradiol dioxygenase cannot carry out the reaction with alternative cations. This difference in selectivity of VOC superfamily members depends on the reaction. If the reaction only needs an electrophilic center to stabilize the intermediates, then those members could function in the presence of more than one metal. However, in some enzymatic reaction catalysis, redox chemistry plays an important role. The redox potential of the metal center and that of the substrate must match. Thus, extradiol dioxygenases strictly depend on redox-active Fe(II) for catalysis and cannot function in the presence of the non-redox metal Mn(II) (Costas et al. 2004).

1.6 Pathways for Detoxification of MAs(III)

We propose that ArsM originally evolved in an anoxic environment to produce toxic MAs(III), which has antibiotic properties. Once the atmosphere became oxidizing, MAs(III) was converted to nontoxic MAs(V), and ArsM became a detoxification process for inorganic arsenic. However, some soil bacteria can reduce MAs(V) to MAs(III), recreating the toxic antibiotic even in aerobic soil (Qin et al. 2006). As described earlier, other soil bacteria have *arsI*, *arsH* and *arsP* genes that encode MAs(III) detoxification proteins. The presence of these genes in extant bacteria would not occur if they those bacteria were not under the present-day selective pressure of MAs(III) production in soil.

1.6.1 ArsI, a C-As Bond Lyase

If microbes expressing As(III) SAM methyltransferases are common, why is most of the arsenic in the environment not methylated? Inorganic arsenic continuously enters the environment from geothermal sources, and some methylated species are degraded abiotically, but there are less environmental methylarsenicals than might be expected considering how widespread *arsM* genes are. Thus, even though the steady-state levels of methylated arsenicals in soil is not high, its continual synthesis suggests the existence of biological methylarsenical breakdown pathways. The answer is the presence of microbial pathways for MAs(III) degradation such as the ArsI C-As lyase, which cleaves the C-As bond, converting MAs(III) into As(III) (Yoshinaga and Rosen, 2014). Although microbial degradation of environmental organoarsenicals has been

documented, molecular details of the reaction were unknown (Stolz et al. 2007; Feng et al. 2005; Lehr et al. 2003; Nadar et al. 2014). Microbial communities in Florida golf course soil were shown to carry out a two-step pathway of MSMA reduction and demethylation (Yoshinaga et al. 2011). An environmental MAs(III)-demethylating bacterium *Bacillus* sp. MD1 was isolated from Florida golf course soil, and the *arsI* gene was cloned by selection for MAs(III) (Yoshinaga and Rosen, 2014). ArsI is non-heme Fe(II)-dependent dioxygenase with C-As lyase activity. A thermophilic ArsI was cloned from *Thermomonospora curvata* and crystallized [Nadar et al. 2016]. The structure of TcArsI was recently solved at 1.97 Å (Fig 4) (Rutherford et al. 2003). The enzyme was crystallized with Ni(II) bound to the three-amino acid triad that forms the metal binding site in other dioxygenases. The enzyme has a cysteine pair that was predicted to be the MAs(III) binding site (Garbarino et al. 2003). Although the enzyme had not previously been crystallized with bound MAs(III), a structure with 2-mercaptoethanol bound to one of the two cysteines and to the inner shell of the Fe(II) binding site was determined, strongly supporting the role of the cysteine pair in binding MAs(III). The two cysteines are present in a flexible loop. In a series of structures the loop could be observed to move from solvent to the metal binding site, suggesting a mechanism in which MAs(III) is bound in the bulk solvent and progressively moved into proximity of the Fe(II), where one atom of dioxygen could be added to the arsenic atom of MAs(III) and the other oxygen atom to the carbon, forming the products $As(OH)_3$ and formaldehyde. ArsI can also cleave the C-As bond of aromatic arsenicals. Large amounts of organic

arsenicals have used as herbicides and growth promoters in animal husbandry, as well as in various industrial activities such as the semiconductor industry (Stolz et al. 2007; Matteson et al. 2014).

The structure of ArsI from *T. curvata* was solved at a final resolution of 1.46 Å (PDB ID: 5C68) (Fig. 4). Atoms are colored as in Figure 1. Fe(II) (orange sphere) is bound to Gln8, His65, and Glu117. The blue spheres are oxygen atoms. The organoarsenic binding site is composed of Cys98 and Cy99, with MAs(III) modeled into the site. Both inorganic and organic arsenicals have been used in agriculture in the United States for over a century (Stellman et al. 2003). For example, MSMA (MAs(V)) was used for decades for weed control and is still in use today as an herbicide for turf maintenance on golf courses, sod farms, highway rights of way and weed control in cotton fields (Stellman et al. 2003). DMAs(V) (Agent Blue) was one of the rainbow herbicides used during the war in Vietnam (Makris et al. 2008). The pentavalent aromatic arsenicals Roxarsone (Rox(V)), Nitarsone (Nit(V) (4-nitrophenyl)arsenate) and p-ASA (p-aminophenyl arsenate) have been used since the 1940s as antimicrobial growth promoters for chickens, turkeys and pigs to control *Coccidioides* infections, improve weight gain, feed efficiency and meat pigmentation (Stolz et al, 2007; Feng et al. 2005). Pfizer has voluntarily suspended production of these aromatic arsenicals in the United States, but they are still produced worldwide and used in poultry farms in other countries. These aromatic arsenicals are largely excreted intact by animals, but their excrements are used as fertilizers. These pentavalent organoarsenicals

are relatively benign and less toxic than inorganic arsenicals. However, both aromatic (Stolz et al. 2007; Feng et al. 2005; Von et al. 1968) and methyl (Lehr et al. 2003; Chen et al. 2014) arsenicals are reduced to the highly toxic trivalent forms (Neyt et al. 2002), which are degraded into toxic inorganic forms in the environment, where they contaminate foods and water supplies. This suggests that ArsI-catalyzed C-As bond cleavage of these aromatic arsenicals by microbial communities transform aromatic arsenicals into inorganic arsenicals that contaminate our food and water supplies.

These structural features strongly indicate that ArsI proteins catalyze As-C bond cleavage by utilizing a mechanism like that of ring-cleavage by AkbC. In *ars* gene products, conserved cysteine residues function as arsenic binding sites and are essential for the activity (Shi et al. 1996; Qin et al. 2007; Marapakala et al. 2012).

In brief, the proposed mechanism of ArsI dioxygenase (Fig 5), C-As bond cleavage of the substrate methyl arsenite [MAs(III)] [upper] and trivalent roxarsone [Rox(III)] [lower] shown with Fe(II) in the catalytic center ligated with 2-his-1-carboxylate triad and three water molecules. After catalysis of this substrate by ArsI dioxygenase, the products are inorganic arsenite [As(III)] and predicted carbon compounds as formaldehyde for MAs(III). The produce of Rox(III) cleavage is uncertain. If the ring is cleaved, it may be 4-hydroxy-5-nitro-hexa-2,4-dienal or a similar open chain compound. If the ring is not cleaved, it may be 4-hydroxy-2-nitrophenol.

1.7 Figures

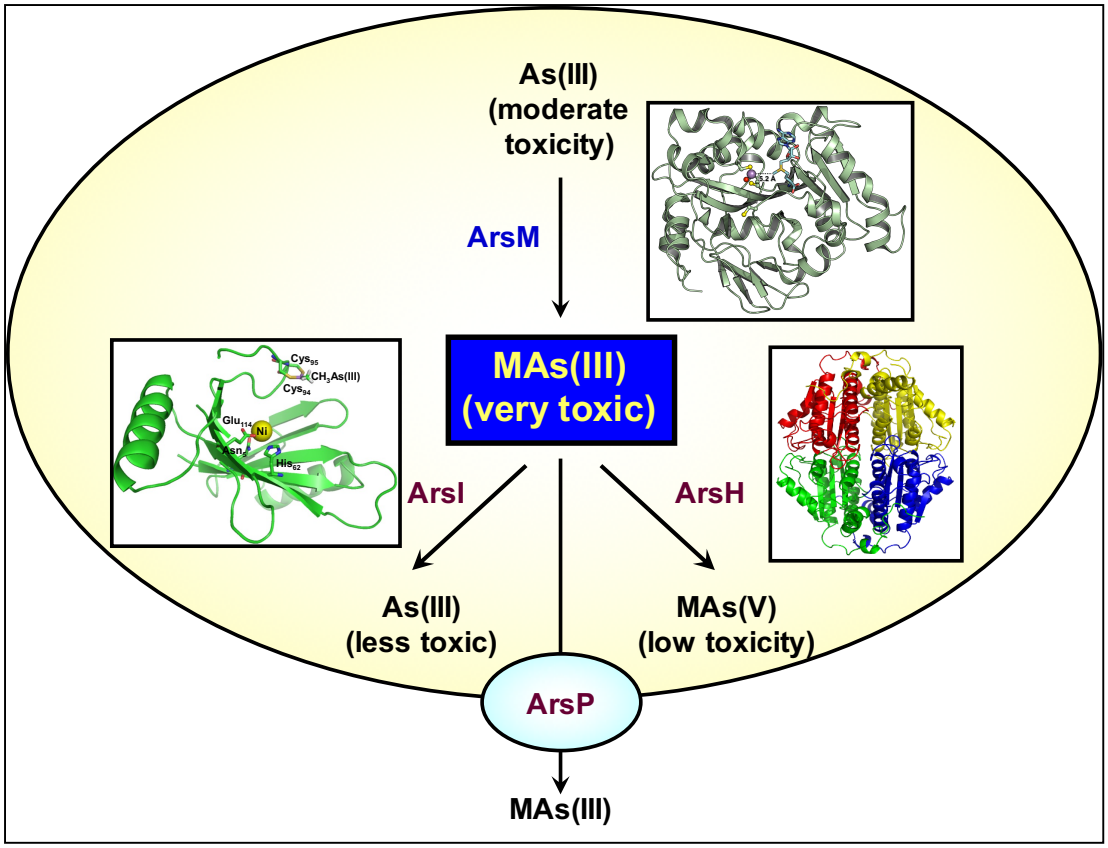


Figure 1. Fate of As(III) in bacteria by Ars proteins and the crystal structures of the relevant enzymes. In communities of soil microbes, some bacteria such as *Rhodopseudomonas palustris* carry an *arsM* gene for the As(III) SAM methyltransferase, producing highly toxic MAs(III). This trivalent organoarsenical has antibiotic-like properties. Other soil bacteria carry genes for MAs(III) resistance. Others such as *Bacillus* MD1 have the *arsI* gene for the ArsI C-As lyase gene that confers resistance to MAs(III) by degrading it into As(III) and formaldehyde. Yet other soil bacteria such as *Pseudomonas putida* have a gene encoding *ArsH*, a flavoprotein that uses $NADP^+$ to oxidize highly toxic MAs(III) to relatively nontoxic MAs(V), thus conferring resistance. Finally, other bacteria such as *Campylobacter jejuni*, which inhabits the intestinal tract of poultry and other farm animals, carry the *arsP* gene. ArsP is a MAs(III) efflux permease that extrudes trivalent organoarsenicals from cells, conferring resistance. The crystal structures of the relevant enzymes are shown next to their respective reactions (Reproduced by permission of The Royal Society of Chemistry).

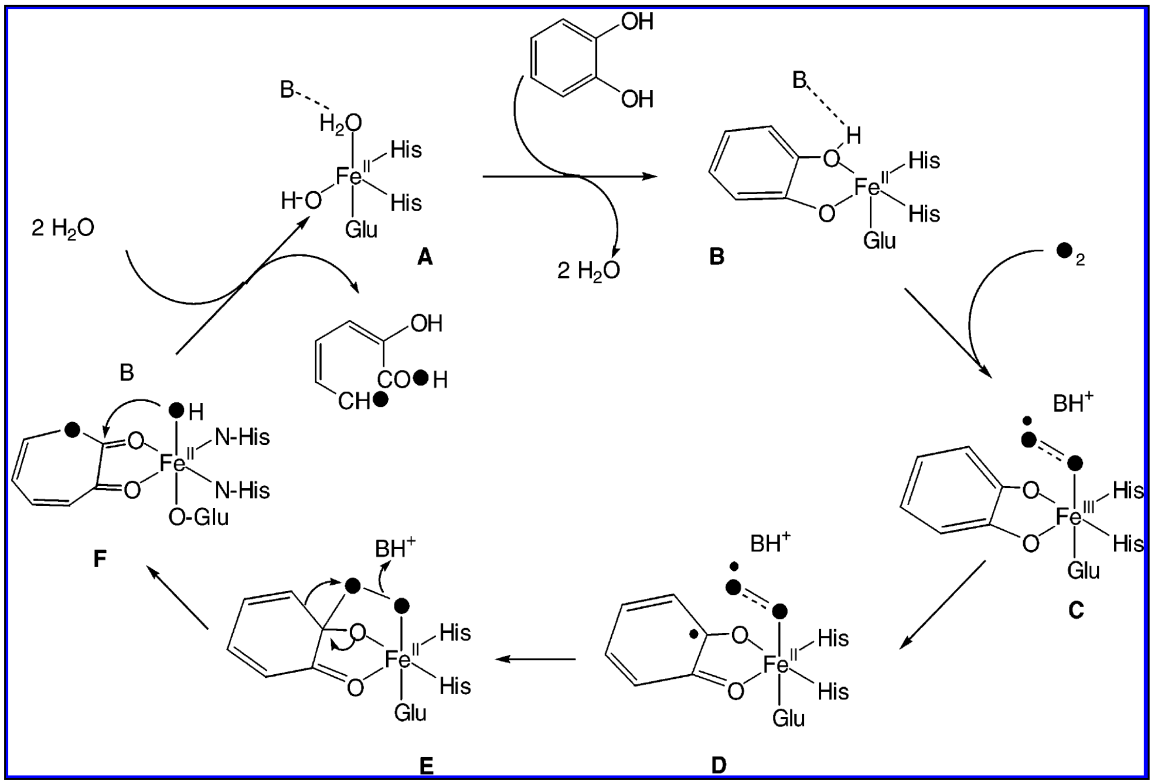


Figure 2. The proposed mechanism of extradiol catechol dioxygenase.

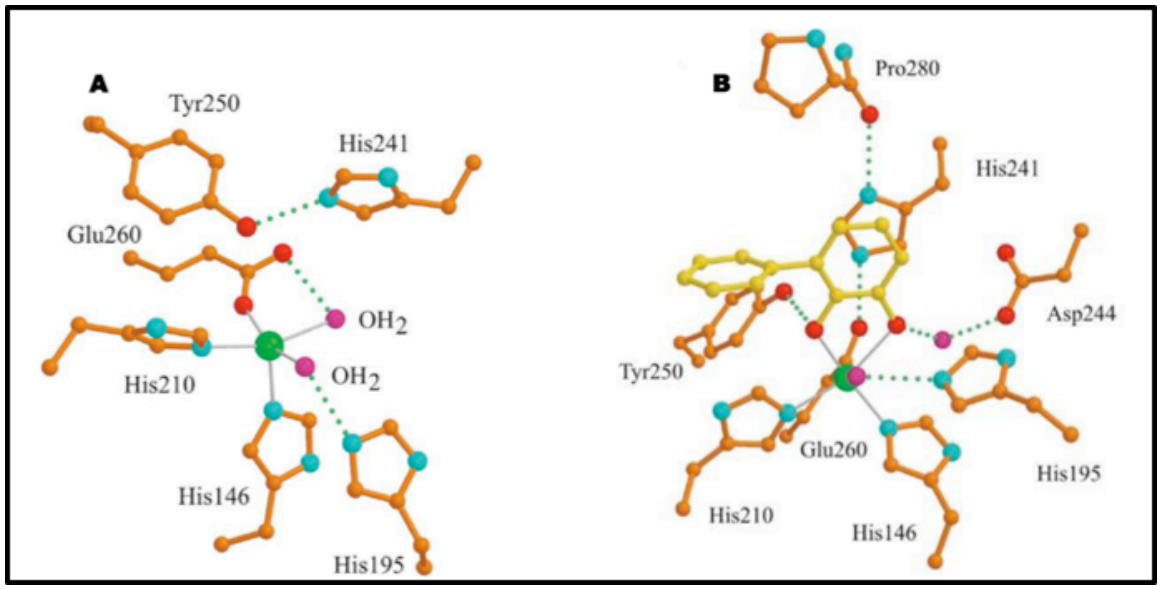


Figure 3. The structure of the active site of substrate-free DHBD (A) and bound to DHB (B).

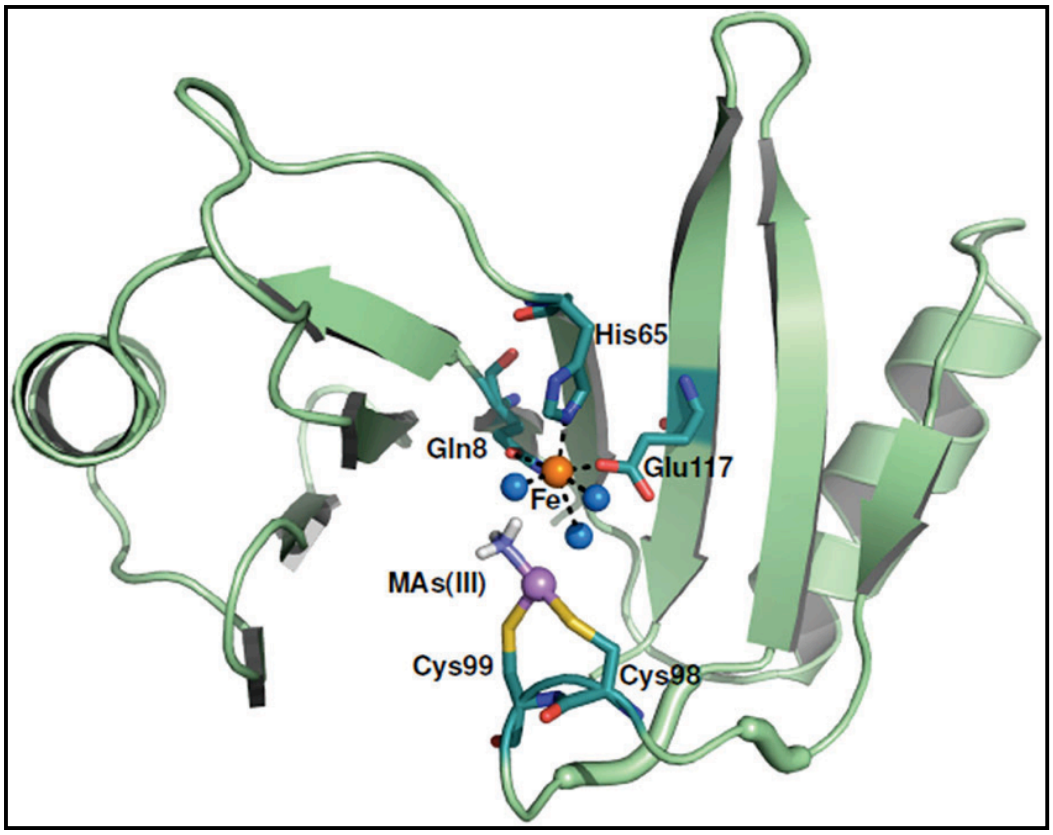


Figure 4. The structure of the ArsI C-As lyase.

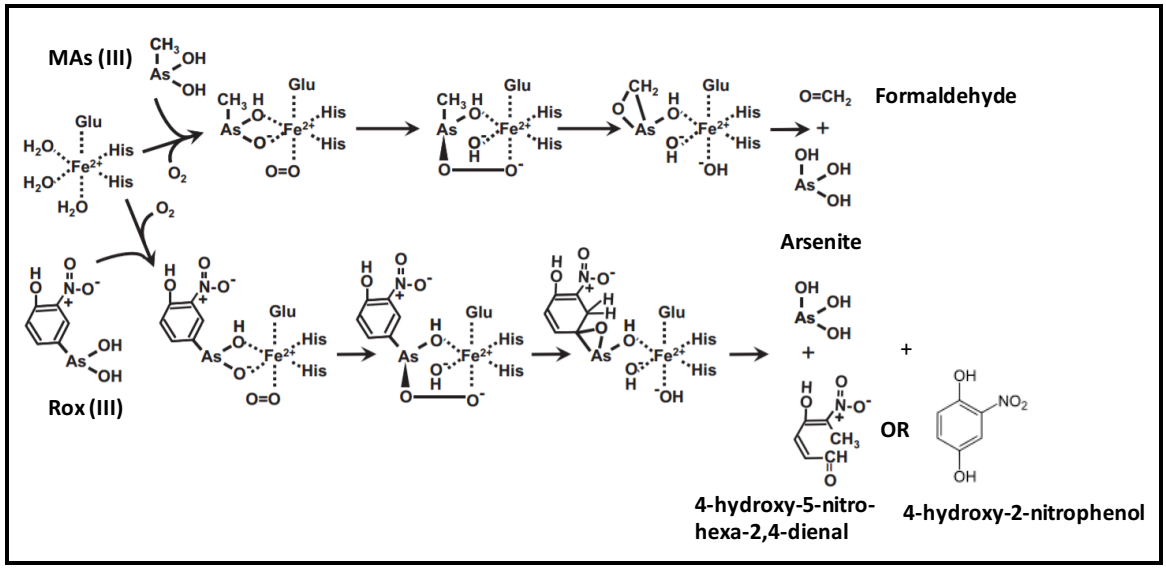


Figure 5. Proposed mechanism and products for ArsI catalyzed the degradation of MAs(III) and Rox(III).

1.8 References

Anand, R., P. C. Dorrestein, et al. (2002). "Structure of Oxalate Decarboxylase from *Bacillus subtilis* at 1.75 Å Resolution," Biochemistry **41**(24): 7659-7669.

Armstrong, R. N. (2000). "Mechanistic Diversity in a Metalloenzyme Superfamily " Biochemistry **39**(45): 13625-13632.

Bencko, V. and K. Symon (1977). "Test of environmental exposure to arsenic and hearing changes in exposed children." Environmental Health Perspectives **19**: 95.

Bernat, B. A., L. T. Laughlin, et al. (1997). "Fosfomycin Resistance Protein (FosA) Is a Manganese Metalloglutathione Transferase Related to Glyoxalase I and the Extradiol Dioxygenases." Biochemistry **36**(11): 3050-3055.

Bhattacharjee, H., J. Carbrey, et al. (2004). "Drug uptake and pharmacological modulation of drug sensitivity in leukemia by AQP9." Biochemical and Biophysical Research Communications **322**(3): 836-841.

Cameron, A. D., B. Olin, et al. (1997). "Crystal structure of human glyoxalase I—evidence for gene duplication and 3D domain swapping." The EMBO Journal **16**(12): 3386-3395.

Chen, J., H. Bhattacharjee, et al. (2015). "ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone." Molecular Microbiology **96**(5): 1042-1052.

Chen, J., M. Madegowda, et al. (2015). "ArsP: a methylarsenite efflux permease." Molecular Microbiology **98**(4): 625-635.

Chen, J., J. Qin, et al. (2013). "Engineering the soil bacterium *Pseudomonas putida* for arsenic methylation." Applied and Environmental Microbiology **79**(14): 4493-4495.

Chen, J. and B. P. Rosen (2014). "Biosensors for inorganic and organic arsenicals." Biosensors **4**(4): 494-512.

Cleasby, A., A. Wonacott, et al. (1996). "The X-ray crystal structure of phosphomannose isomerase from Candida albicans at 1.7 A resolution." Nature Structural & Molecular Biology **3**(5): 470-479.

Copping, M. (2009). "Death in the beer glass: the Manchester arsenic-in-beer epidemic of 1900–1 and the long-term poisoning of beer." Journal of the Brewery History Society **132**: 31-57.

Costas, M., M. P. Mehn, et al. (2004). "Dioxygen Activation at Mononuclear Nonheme Iron Active Sites: Enzymes, Models, and Intermediates." Chemical Reviews **104**(2): 939-986.

Cullen, W. R., B. C. McBride, et al. (1979). "The transformation of arsenicals by Candida humicola." Canadian Journal of Microbiology **25**(10): 1201-1205.

Davis, M. I., E. C. Wasinger, et al. (2003). "Spectroscopic and Electronic Structure Studies of 2,3-Dihydroxybiphenyl 1,2-Dioxygenase: O2 Reactivity of the Non-Heme Ferrous Site in Extradiol Dioxygenases." Journal of the American Chemical Society **125**(37): 11214-11227.

Drobná, Z., L. M. Del Razo, et al. (2013). "Environmental exposure to arsenic, AS3MT polymorphism and prevalence of diabetes in Mexico." Journal of Exposure Science and Environmental Epidemiology **23**(2): 151-155.

Dumas, P., M. Bergdoll, et al. (1994). "Crystal structure and site-directed mutagenesis of a bleomycin resistance protein and their significance for drug sequestering." The EMBO Journal **13**(11): 2483-2492.

Dunwell, J. M., A. Culham, et al. (2001). "Evolution of functional diversity in the cupin superfamily." Trends in Biochemical Sciences **26**(12): 740-746.

Dunwell, J. M., S. Khuri, et al. (2000). "Microbial Relatives of the Seed Storage Proteins of Higher Plants: Conservation of Structure and Diversification of Function during Evolution of the Cupin Superfamily." Microbiology and Molecular Biology Reviews **64**(1): 153-179.

Eltis, L. D. and J. T. Bolin (1996). "Evolutionary relationships among extradiol dioxygenases." Journal of Bacteriology **178**(20): 5930-5937.

Falk, H., G. Caldwell, et al. (1981). "Arsenic-related hepatic angiosarcoma." American Journal of Industrial Medicine **2**(1): 43-50.

Feng, M., J. E. Schrlau, et al. (2005). "Arsenic transport and transformation associated with MSMA application on a golf course green." Journal of Agricultural and Food Chemistry **53**(9): 3556-3562.

Fetzner, S. (2012). "Ring-Cleaving Dioxygenases with a Cupin Fold." Applied and Environmental Microbiology **78**(8): 2505-2514.

Gao, S. and R. G. Burau (1997). "Environmental factors affecting rates of arsine evolution from and mineralization of arsenicals in soil." Journal of Environmental Quality **26**(3): 753-763.

Garbarino, J. R., A. J. Bednar, et al. (2003). "Environmental fate of roxarsone in poultry litter. I. Degradation of roxarsone during composting." Environmental Science & Technology **37**(8): 1509-1514.

García-Vargas, G. G. and A. Hernández-Zavala (1996). "Urinary Porphyrins and Heme Biosynthetic Enzyme Activities Measured byHPLC in Arsenic Toxicity." Biomedical Chromatography **10**(6): 278-284.

Gerlt, J. A. and P. C. Babbitt (2001). "Divergent Evolution of Enzymatic Function: Mechanistically Diverse Superfamilies and Functionally Distinct Suprafamilies." Annual Review of Biochemistry **70**(1): 209-246.

Ghosh, M., J. Shen, et al. (1999). "Pathways of As(III) detoxification in Saccharomyces cerevisiae." Proceedings of the National Academy of Sciences **96**: 5001-5006.

Hahn, D. R., P. J. Solenberg, et al. (1991). "Tn5099, a xylE promoter probe transposon for Streptomyces spp." Journal of Bacteriology **173**(17): 5573-5577.

Han, D., G. J. Handelman, et al. (1995). "Analysis of reduced and oxidized lipoic acid in biological samples by high-performance liquid chromatography." Methods in Enzymology **251**: 315-325.

Happe, B., L. D. Eltis, et al. (1993). "Characterization of 2,2',3-trihydroxybiphenyl dioxygenase, an extradiol dioxygenase from the dibenzofuran- and dibenzo-p-dioxin-degrading bacterium Sphingomonas sp. strain RW1." Journal of Bacteriology **175**(22): 7313-7320.

Harayama, S., M. Kok, et al. (1992). "Functional and evolutionary relationships among diverse oxygenases." Annual Reviews in Microbiology **46**(1): 565-601.

Harayama, S. and M. Rekik (1989). "Bacterial aromatic ring-cleavage enzymes are classified into two different gene families." Journal of Biological Chemistry **264**(26): 15328-15333.

Harpel, M. R. and J. D. Lipscomb (1990). "Gentisate 1,2-dioxygenase from pseudomonas. Purification, characterization, and comparison of the enzymes from Pseudomonas testosteroni and Pseudomonas acidovorans." Journal of Biological Chemistry **265**(11): 6301-6311.

Iwabuchi, T. and S. Harayama (1998). "Biochemical and Molecular Characterization of 1-Hydroxy-2-naphthoate Dioxygenase from Nocardioides sp. KP7." Journal of Biological Chemistry **273**(14): 8332-8336.

Kita, A., S.-i. Kita, et al. (1999). "An archetypical extradiol-cleaving catecholic dioxygenase: the crystal structure of catechol 2,3-dioxygenase (metapyrocatechase) from *Pseudomonas putida mt-2*." Structure **7**(1): 25-34.

Kovaleva, E. G. and J. D. Lipscomb (2008). "Versatility of biological non-heme Fe(II) centers in oxygen activation reactions." Nature Chemical Biology **4**(3): 186-193.

Lehr, C., E. Polishchuk, et al. (2003). "Demethylation of methylarsenic species by Mycobacterium neoaurum." Applied Organometallic Chemistry **17**(11): 831-834.

Lehr, C. R., E. Polishchuk, et al. (2003). "Demethylation of methylarsenic species by Mycobacterium neoaurum." Applied Organometallic Chemistry **17**(11): 831-834.

Lin, Y. F., A. R. Walmsley, et al. (2006). "An arsenic metallochaperone for an arsenic detoxification pump." Proceedings of the National Academy of Sciences **103**(42): 15617-15622.

Liu, Z., J. Shen, et al. (2002). "Arsenite transport by mammalian aquaglyceroporins AQP7 and AQP9." Proceedings of the National Academy of Sciences **99**(9): 6053-6058.

Liu, Z., M. Styblo, et al. (2006). "Methylarsonous acid transport by aquaglyceroporins." Environmental Health Perspectives **114**(4): 527-531.

Mabrouk, P. A., A. M. Orville, et al. (1991). "Variable-temperature variable-field magnetic circular dichroism studies of the iron(II) active site in metapyrocatechase: implications for the molecular mechanism of extradiol dioxygenases." Journal of the American Chemical Society **113**(11): 4053-4061.

Maki, T., N. Takeda, et al. (2006). "Isolation of monomethylarsonic acid-mineralizing bacteria from arsenic contaminated soils of Ohkunoshima Island." Applied Organometallic Chemistry **20**: 538-544.

Makris, K. C., S. Quazi, et al. (2008). "Fate of arsenic in swine waste from concentrated animal feeding operations." Journal of Environmental Quality **37**(4): 1626-1633.

Marapakala, K., J. Qin, et al. (2012). "Identification of catalytic residues in the As(III) S-adenosylmethionine methyltransferase." Biochemistry **51**(5): 944-951.

McCarthy, A. A., H. M. Baker, et al. (2001). "Crystal Structure of Methylmalonyl-Coenzyme A Epimerase from P. shermanii: a Novel Enzymatic Function on an Ancient Metal Binding Scaffold." Structure **9**(7): 637-646.

Mukherjee, S. C., M. M. Rahman, et al. (2003). "Neuropathy in arsenic toxicity from groundwater arsenic contamination in West Bengal, India." Journal of Environmental Science and Health, Part A **38**(1): 165-183.

Mukhopadhyay, R. and B. P. Rosen (2002). "Arsenate reductases in prokaryotes and eukaryotes." Environ Health Perspect **110 Suppl 5**: 745-748.

Nadar, S. V., M. Yoshinaga, et al. (2014). "Crystallization and preliminary X-ray crystallographic studies of the ArsI C-As lyase from Thermomonospora curvata." Acta Crystallographica Section F: Structural Biology Communications **70**(Pt 6): 761-764.

Nadar, V. S., M. Yoshinaga, et al. (2016). "Structure of the ArsI C–As Lyase: Insights into the Mechanism of Degradation of Organoarsenical Herbicides and Growth Promoters." Journal of Molecular Biology **428**(11): 2462-2473.

Nozaki, M. (1979). "Oxygenases and dioxygenases." Biochemistry: 145-186.

Petrick, J. S., F. Ayala-Fierro, et al. (2000). "Monomethylarsonous Acid (MMAIII) Is More Toxic Than Arsenite in Chang Human Hepatocytes." Toxicology and Applied Pharmacology **163**(2): 203-207.

Pochapsky, T. C., S. S. Pochapsky, et al. (2002). "Modeling and experiment yields the structure of acireductone dioxygenase from Klebsiella pneumoniae." Nature Structural & Molecular Biology **9**(12): 966-972.

Qin, J., H. L. Fu, et al. (2007). "Convergent evolution of a new arsenic binding site in the ArsR/SmtB family of metalloregulators." Journal of Biological Chemistry **282**(47): 34346-34355.

Qin, J., B. P. Rosen, et al. (2006). "Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine

methyltransferase." Proceedings of the National Academy of Sciences **103**(7): 2075-2080.

Ravenscroft, P., H. Brammer, et al. (2009). Arsenic pollution: a global synthesis, John Wiley & Sons.

Rosen, B. P., U. Weigel, et al. (1988). "Molecular characterization of an anion pump. The *arsA* gene product is an arsenite(antimonate)-stimulated ATPase." Journal of Biological Chemistry **263**(7): 3067-3070.

San Francisco, M. J., C. M. Chen, et al. (1988). "Identification of the membrane component of the anion pump encoded by the arsenical resistance operon of R-factor R773." Progress in Clinical and Biological Research **252**: 311-316.

San Francisco, M. J., C. L. Hope, et al. (1990). "Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon." Nucleic Acids Research **18**(3): 619-624.

Sato, N., Y. Uragami, et al. (2002). "Crystal Structures of the Reaction Intermediate and its Homologue of an Extradiol-cleaving Catecholic Dioxygenase." Journal of Molecular Biology **321**(4): 621-636.

Serre, L., A. Sailland, et al. (1999). "Crystal structure of Pseudomonas fluorescens 4-hydroxyphenylpyruvate dioxygenase: an enzyme involved in the tyrosine degradation pathway." Structure **7**(8): 977-988.

Shi, W., J. Dong, et al. (1996). "The role of arsenic-thiol interactions in metalloregulation of the *ars* operon." Journal of Biological Chemistry **271**(16): 9291-9297.

Smith, A. H., E. O. Lingas, et al. (2000). "Contamination of drinking-water by arsenic in Bangladesh: a public health emergency." Bulletin of the World Health Organization **78**(9): 1093-1103.

Spence, E. L., G. J. Langley, et al. (1996). "Cis−Trans Isomerization of a Cyclopropyl Radical Trap Catalyzed by Extradiol Catechol Dioxygenases: Evidence for a Semiquinone Intermediate." Journal of the American Chemical Society **118**(35): 8336-8343.

States, J. C., S. Srivastava, et al. (2009). "Arsenic and Cardiovascular Disease." Toxicological Sciences **107**(2): 312-323.

Stellman, J. M., S. D. Stellman, et al. (2003). "The extent and patterns of usage of Agent Orange and other herbicides in Vietnam." Nature **422**(6933): 681-687.

Stolz, J. F., E. Perera, et al. (2007). "Biotransformation of 3-nitro-4-hydroxybenzene arsonic acid (roxarsone) and release of inorganic arsenic by Clostridium species." Environmental Science & Technology **41**(3): 818-823.

Straganz, G. and B. Nidetzky (2006). "Variations of the 2-His-1-carboxylate Theme in Mononuclear Non-Heme FeII Oxygenases." Chembiochem **7**(10): 1536-1548.

Sugimoto, K., T. Senda, et al. (1999). "Crystal structure of an aromatic ring opening dioxygenase LigAB, a protocatechuate 4,5-dioxygenase, under aerobic conditions." Structure **7**(8): 953-965.

Thomas, D. J. and B. P. Rosen (2013). Arsenic methyltransferases Encyclopedia of Metalloproteins. R. H. Kretsinger, V. N. Uversky and E. A. Permyakov. New York, Springer New York**:** 140-145.

Titus, G. P., H. A. Mueller, et al. (2000). "Crystal structure of human homogentisate dioxygenase." Nature Structural & Molecular Biology **7**(7): 542-546.

Tsagkaris, A. S., C. A. Papachristidis, et al. (2017). "Adulteration Stories." Food Authentication: Management, Analysis and Regulation: 423.

Tyler, C. R. and A. M. Allan (2014). "The effects of arsenic exposure on neurological and cognitive dysfunction in human and rodent studies: a review." Current Environmental Halth Reports **1**(2): 132-147.

Vaillancourt, F. H., C. J. Barbosa, et al. (2002). "Definitive Evidence for Monoanionic Binding of 2,3-Dihydroxybiphenyl to 2,3-Dihydroxybiphenyl 1,2-Dioxygenase from UV Resonance Raman Spectroscopy, UV/Vis Absorption Spectroscopy, and Crystallography." Journal of the American Chemical Society **124**(11): 2485-2496.

Vaillancourt, F. H., J. T. Bolin, et al. (2006). "The ins and outs of ring-cleaving dioxygenases." Critical Reviews in Biochemistry and Molecular Biology **41**(4): 241-267.

Vaillancourt, F. H., M.-A. Haro, et al. (2003). "Characterization of Extradiol Dioxygenases from a Polychlorinated Biphenyl-Degrading Strain That

Possess Higher Specificities for Chlorinated Metabolites." Journal of Bacteriology **185**(4): 1253-1260.

van Lis, R., W. Nitschke, et al. (2013). "Arsenics as bioenergetic substrates." Biochimica et Biophysica Acta (BBA)-Bioenergetics **1827**(2): 176-188.

Venteris, E. R., N. T. Basta, et al. (2014). "Modeling spatial patterns in soil arsenic to estimate natural baseline concentrations." Journal of Environmental Quality **43**(3): 936-946.

Vetting, M. W. and D. H. Ohlendorf (2000). "The 1.8 Å crystal structure of catechol 1,2-dioxygenase reveals a novel hydrophobic helical zipper as a subunit linker." Structure **8**(4): 429-440.

Vetting, M. W., L. P. Wackett, et al. (2004). "Crystallographic Comparison of Manganese- and Iron-Dependent Homoprotocatechuate 2,3-Dioxygenases." Journal of Bacteriology **186**(7): 1945-1958.

Von Endt, D. W., P. C. Kearney, et al. (1968). "Degradation of monosodium methanearsonic acid by soil microorganisms." Journal of Agricultural and Food Chemistry **16**: 17-20.

Woolson, E. A., N. Aharonson, et al. (1982). "Application of the high-performance liquid chromatography-flameless atomic absorption method to the study of alkyl arsenical herbicide metabolism in soil." Journal of Agricultural and Food Chemistry **30**(3): 580-584.

Wysocki, R., P. Bobrowicz, et al. (1997). "The *Saccharomyces cerevisiae ACR3* gene encodes a putative membrane protein involved in arsenite transport." Journal of Biological Chemistry **272**(48): 30061-30066.

Yan, Y., X.-M. Xue, et al. (2017). "Co-expression of cyanobacterial genes for arsenic methylation and demethylation in Escherichia coli offers insights into arsenic resistance." Frontiers in Microbiology **8**.

Yang, H.-C. and B. P. Rosen (2016). "New mechanisms of bacterial arsenic resistance." Biomedical Journal **39**(1): 5-13.

Yoshinaga, M., Y. Cai, et al. (2011). "Demethylation of methylarsonic acid by a microbial community." Environmental microbiology **13**(5): 1205-1215.

Yoshinaga, M. and B. P. Rosen (2014). "AC· As lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters." Proceedings of the National Academy of Sciences **111**(21): 7701-7706.

Yoshinaga, M. and B. P. Rosen (2014). "A CAs lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters." Proceedings of the National Academy of Sciences **111**(21): 7701-7706.

Zhu, Y. G., M. Yoshinaga, et al. (2014). "Earth abides arsenic biotransformations." Annual Review of Earth and Planetary Sciences **42**: 443-467.

CHAPTER 2. PROBLEM STATEMENT AND OBJECTIVES

Arsenic is an omnipresent environmental toxin. Additionally, anthropogenic activities also such as the use of MSMA as herbicide, roxarsone in poultry industry contribute to arsenic contamination in drinking water bodies and food. Arsenic toxicity varies with its form. The order of toxicity is MMA(III) > As(III) > As(V) > MMAs(V) = DMAs(V) (Petrik et al. 2000). To come up with a solution to arsenic toxicity, more knowledge of the arsenic biogeocycle is needed. Fig, 6 shows the biogeochemical cycle of arsenic in the environment. So far, arsenic is known to interchange between trivalent and pentavalent inorganic arsenic, and arsenic methylation was well known. Recently, our lab has found both soil isolate and gene involved in arsenic demethylation cycle as *Bacillus* sp. MD1 and *arsI* respectively. ArsI is a non-heme iron-dependent C-As lyase. It belongs to type-I extradiol dioxygenase and first reported enzyme to break C-As bond. It is a new enzyme in enzyme biology, and there is a necessity to characterize this enzyme to understand the molecular mechanism of its catalysis. Also, genome sequencing of soil isolates *Burkholderia* sp. MR, which carries out the reduction of pentavalent MSMA into trivalent form, is an objective with the goal of identifying the gene for the reductase.

Project objectives are as:

- Structure-function analysis of ArsI by mutagenesis
- ArsI thermodynamic studies and real-time binding assay of trivalent arsenicals
- Determination of the carbon product of trivalent organoarsenicals after ArsI catalysis
- Draft genome sequence of *Burkholderia* sp. MR1, a methylarsenate-reducing bacterial isolate from Florida golf course soil

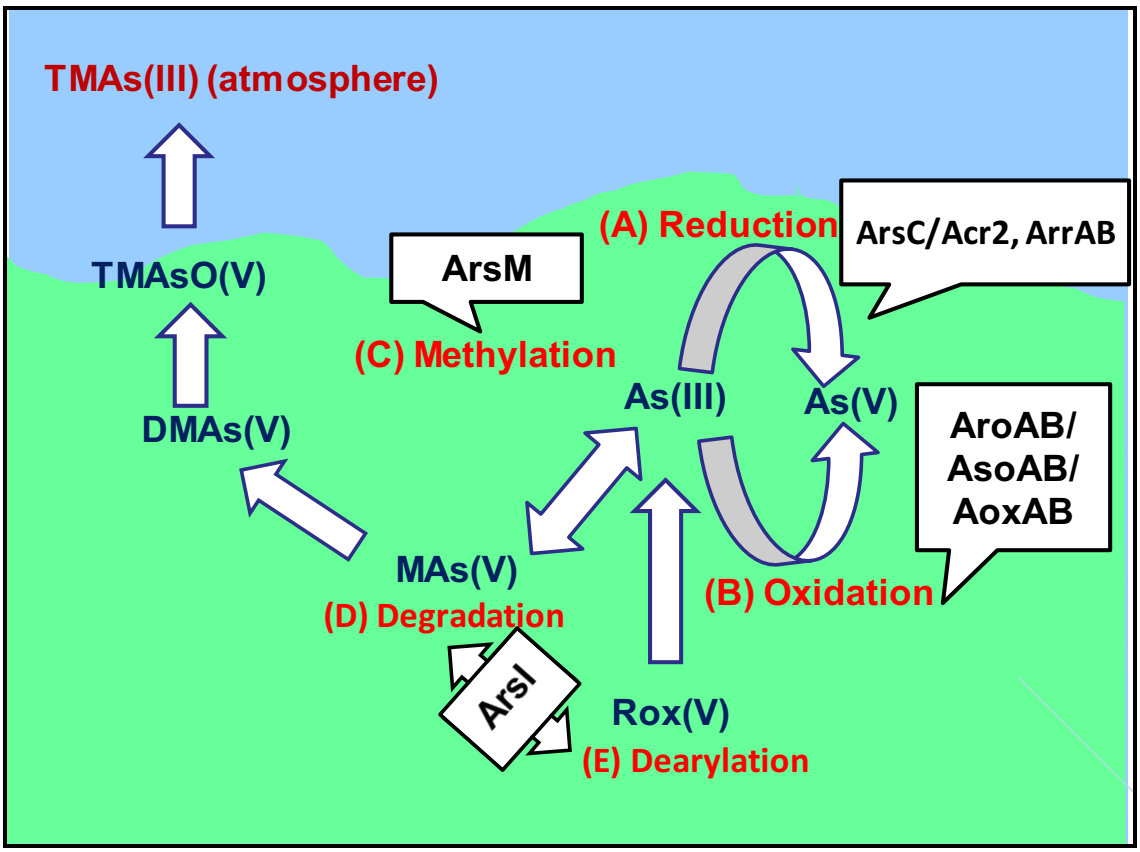


Figure 6. The biogeochemical cycle of arsenic in the environment.

CHAPTER 3. STRUCTURE-FUNCTION ANALYSIS OF ARSI BY MUTAGENESIS

3.1 Introduction

A novel ArsI, a C-As lyase was identified recently from a Florida golf course soil isolate Bacillus species and designated as *Bacillus* sp. MD1 is representing its MAs(III) demethylation capability. ArsI is a type-I extradiol dioxygenase and uses a divalent metal ion, a Fe(II)-dependent MAs(III) demethylase (Yoshinaga and Rosen 2014). As of April 25, 2017, there are nearly 1305 putative ArsI orthologs identified in 963 aerobic bacterial species by a Basic Local Alignment Search Tool Link to Protein and Structures search (BLAST Link or BLink), suggesting that *arsI* genes are widely distributed in the aerobic environment.

Our recent crystallographic study demonstrated that TcArsI, the ArsI ortholog from the thermophile *Thermomonospora curvata* DSM 42183 (Nadar, Yoshinaga et al. 2016), shares the overall structure of the C-terminal catalytic domain of LapB from *Pseudomonas* sp. KL28, a representative of type I extradiol ring-cleaving dioxygenases (Cho, Jung et al. 2009). The crystal structures obtained with divalent metals Ni(II) and Co(II) suggests that TcArsI forms the active site by coordinating Fe(II) via the triad of residues Gln8-His65-Glu117 (His5-His62-Glu115 in the *Bacillus* ArsI) conserved both in ArsI orthologs and the type-I extradiol ring-cleaving dioxygenases (Fig 7). The type-I ring-cleaving dioxygenases bind the substrates directly via Fe(II) in the active sites (Nozaki 1979). In contrast, we proposed that ArsI proteins have unique substrate binding

sites that are quite different from those in the type-I ring-cleaving dioxygenases (Nadar, Yoshinaga et al. 2016). In Ars proteins, the products of arsenic resistance (*ars*) genes, conserved cysteine residues often function as binding sites for substrates of trivalent arsenicals (Marapakala, Qin et al. 2012). There is a vicinal cysteine pair (Cys96-Cys97 in the *Bacillus* ArsI) conserved in all putative ArsI orthologs (Fig 7), which we predicted is the substrate binding site. TcArsI crystal structures showed that the cysteine pair is in a flexible loop, and recently TcArsI bound to Rox(III) crystal structure shows the conserved cysteine pair is the real substrate binding site of ArsI. TcArsI structure suggests a loop-gating mechanism for organoarsenicals degradation. Depending on the presence or absence of ligand, ArsI forms a closed or open state, respectively. In the absence of an organoarsenical substrate, the flexible loop is exposed to solvent, and moves towards the active site in the presence of a substrate and carries out organoarsenicals cleavage by dioxygen addition (Nadar, Yoshinaga et al. 2014; Nadar, Yoshinaga et al. 2016).

3.2 Materials and Methods

3.2.1 Reagents

Roxarsone (4-hydroxy-3-nitrophenylarsonic acid) [Rox(V)] and phenylarsonic acid [PAs(V)] were purchased from Acros organics (Geel, Belgium) p-arsanilic acid [p-ASA(V)] was purchased from Alfa Aesar (Ward Hill, MA). Unless otherwise mentioned, all other chemicals were obtained from Sigma-Aldrich (St. Louis, MO). Trivalent forms of roxarsone [Rox(III)], methylarsonic acid [MAs(III)]

and p-arsanilic acid [p-ASA(III)] were prepared using the procedure described previously (Yoshinaga and Rosen 2014) and used for all the experiments except isothermal titration calorimetry (ITC). For ITC assays, phenylarsine oxide (PAO) from Sigma-Aldrich and methylarsonous acid iodide derivative [MAs(III)I2] synthesized following the previous method (Stice 2014) were used as phenylarsonous acid [PhAs(III)] and MAs(III), respectively, to analyze substrate binding.

3.2.2 Preparation of Mutants

Escherichia coli strain TOP10 (Invitrogen) was used for molecular cloning. A derivative of *arsI* from *Bacillus* sp. MD1 lacking the C-terminus from Glu125 to the end that is not essential for enzyme activity (Yoshinaga and Rosen 2014), $arsI_{124}$, was amplified using forward primer 5′-GGGGCATATGAAATATGCGCATGTGGGTCTT-3’ (NdeI site underlined) and reverse primer 5′-AAAAGGAATTCAACAGTTGTCTTTGTGTAG-3′ (EcoRI site underlined), double-digested by NdeI and EcoRI (New England BioLabs) and cloned into vector plasmid pET28a (Novagene), generating pET28a-$arsI_{124}$ encoding C-terminal truncated $ArsI_{124}$ with an N-terminal 6-histidine tag (Table 1). Mutations in the gene $arsI_{124}$ were generated by a QuickChange™ Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) using pET28a-$arsI_{124}$ as a template with primers listed in Table 2. The codons for the Fe(II)-binding residues (His5, His62, and Glu115) were individually mutated to alanine codons, generating H5A, H62A, and E115A. The codons for the substrate-binding residues (Cys96 and Cys97) were individually mutated to serine codons, generating C96S and

C97S. The C96·97S double substitution was also introduced into *arsI*$_{124}$. All constructed mutants were confirmed by sequencing.

3.2.3 Purification of ArsI

Cells of *E. coli* BL21(DE3) (Novagen) bearing pET28a-*arsI*$_{124}$ with or without mutation(s) were grown in Luria–Bertani (LB) medium supplemented with 50 μg/ml kanamycin at 37 °C and induced at A_{600} of 0.6 with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After incubation of additional 4 h at 37 °C, the cells were harvested and suspended in 5 mL per gram of wet cells in buffer A [50mM morpholinopropane-1-sulfonic acid (MOPS), 0.5 M NaCl, 20% glycerol (vol/vol), 20 mM Imidazole and 1 mM *tris*(2-carboxyethyl) phosphine (TCEP), pH 7.0]. Cells were broken by a single passing through a French pressure cell at 20,000 psi, treated immediately with diisopropyl fluorophosphate (2.5 μL per gram wet cell) and pelleted by ultracentrifugation at 125,000 x *g* for an hour. The supernatant was loaded onto a HisPur^{TM} Ni–NTA column (Thermo Scientific, Rockford, IL) with a flow rate of 0.5 ml/min. The column was washed with 100 ml of buffer A. Buffer B, buffer A with 0.2 M imidazole, and was used to elute ArsI$_{124}$ proteins. Eluted proteins were first treated with 2 mM ethylenediaminetetraacetic acid (EDTA) to get rid of Ni ions co-eluted from Ni-NTA column, then concentrated and buffer exchanged into buffer C, buffer A with no imidazole. The purities were analyzed by SDS-PAGE and the concentrations were estimated by Bradford method. Purified ArsI$_{124}$ proteins were divided into small fractions, rapidly frozen using liquid nitrogen and stored at -80 °C until required.

3.2.4 Chemical Modifications of ArsI with Residue-specific Reagents

Inactivation of purified wild-type $ArsI_{124}$ with amino acid residue-specific reagents was examined as described previously (Nogales, Canales et al. 2011). 1 μM $ArsI_{124}$ was incubated in an assay buffer [50 mM MOPS, 150 mM KCl, 3 mM TCEP, 1 mM cysteine, 0.1 mM Fe(II)] containing the cysteine-specific reagents *N*-ethylmaleimide (1 mM) or iodoacetamide (5 mM), the histidine-specific reagent diethyl pyrocarbonate (DEPC) (5 mM), or the glutamic/aspartic acid-specific reagent 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (1 mM) at 25 °C for 30 min, following which 1 μM MAs(III) was added to the protein solution. After incubated at 30°C for an hour with shaking at 200 rpm, MAs(III) demethylation in each solution was analyzed as described above.

3.2.5 MAs(III) Demethylation

Demethylation of MAs(III) was analyzed both *in vivo* using cells of *E. coli* expressing wild-type or mutant *arsI*$_{124}$ genes and *in vitro* with purified proteins. For *in vivo* analysis, cells of *E. coli* strain BL21(DE3) bearing pET28a-*arsI*$_{124}$ derivatives described above were cultured in LB medium supplemented with 50 μg/mL kanamycin at 37 °C. Cells at late exponential phase were diluted 100-fold in LB medium containing 50 μg/ml kanamycin and 0.1 mM IPTG. After 4 h of induction at 37 °C, the cells were transferred to ST^{-1} medium (Maki, Hirota et al. 2009) and incubated with 1 μM MAs(III) for 1h at 25 °C. After incubation, the cells were centrifuged and the resulting supernatant was passed through a 3-kDa cut-off Amicon Ultrafilter (Millipore, Billerica, MA) and arsenic species in the filtrates were analyzed by high-pressure liquid chromatography (HPLC) (Series 2000,

PerkinElmer, Waltham, MA) connected to inductively coupled plasma mass spectroscopy (ICP-MS) (ELAN DRC-e, PerkinElmer) using a C18 reverse phase column (250 x 4.6 mm; 5μ; Thermo Scientific) with mobile-phase consisting of 3 mM malonic acid and 5% (v/v) methanol (pH 5.9, adjusted with tetrabutylammonium hydroxide) at a flow rate of 1 ml/min. MAs(III)-demethylating activities with purified wild-type and mutant $ArsI_{124}$ proteins were assayed in a buffer consist of 0.1 M MOPS, 0.15 M NaCl (pH 7.2) containing 1 mM cysteine, 3 mM TCEP and 0.1 mM Fe(II). The reaction was initiated by addition of 5 μM of MAs(III), incubated at 30°C with shaking at 200 rpm for 1 h and finally terminated with EDTA.

Table 1. Strains and Plasmids

***E. coli* strains**	**Genotype/Description**	**Reference**
TOP10	F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] Δ hsdS λ DE3 =λ sBamHIo Δ EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δ nin5*	(Green M. R 2012)
Plasmids		
pET28a	Expression vector for His-tagged proteins controlled by T7 promoter (Kn^r)	Novagen
pET28a-*arsI*$_{124}$	*arsI*$_{124}$ (encoding C-terminus truncated $ArsI_{124}$) inserted into pET28a at NdeI and EcoRI sites (Kn^r)	This Study

Table 2. Oligonucleotide Primers for Mutagenesis

Primer	Sequence (5'-3')	Result
H5A Fw	ATGAAATATGCGGCTGTGGGTCTT	His5 changed to Ala in wild-type background
H5A Rv	GGATTTCTCCAGATTCGTTACATTAAGACCCACAGCCGCATATTTCAT	
H62A Fw	AAATGGGAATCAAGTCGGGGCTTTCG	His62 changed to Ala in wild-type background
H62A Rv	TTCCACTTGAATCCCGAAAGCCCCGA	
E115A Fw	CGGATGGAAATGAGTGGGCGTTTTTC	Glu115 changed to Ala in wild-type background
E115A Rv	GTCTTTGTGTAGAAAAACGCCCACTC	
C96S Fw	GAAATCAACACTACCTCCTGCTATGC	Cys96 changed to Ala in wild-type background
C96S Rv	CTGGAGTGCATAGCAGGAGGTAGTG	
C97S Fw	ATCAACACTACCTGCTCCTATGCACT	Cys97 changed to Ala in wild-type background
C97S Rv	GTCCTGGAGTGCATAGGAGCAGGTA	

3.3 Results

3.3.1 Chemical Modification of ArsI by Site-specific Reagents

Specific chemical modifiers of His (DEPC), Asp/Glu (EDC) and Cys (*N*-ethylmaleimide or iodoacetamide) residues all inhibited the MAs(III)-demethylating activity by $ArsI_{124}$ completely (Fig 8). In addition to the conserved Fe(II)-binding His5-His-62-Glu115 triad, $ArsI_{124}$ contains multiple non-conserved histidines and acidic residues (Fig 7). Thus, the inhibition data by DEPC and EDC only partially supports the participation of the triad in the catalysis. In contrast, other than the predicted substrate binding residues Cys96-Cys97, $ArsI_{124}$ has no additional cysteine residues (Fig 7). Therefore, the inhibition of the demethylation activity by cysteine-specific modifiers supports our proposal that the cysteine pair is involved in the catalytic mechanism.

3.3.2 Effect of Substitution of Conserved Residues Cys96, Cys97, and Cys96/Cys97, and Fe(II) Binding Conserved Residues His5, His62 and Glu115 on ArsI Demethylation Activity

Our recent crystallographic studies demonstrated that TcArsI, the ArsI ortholog from the thermophile *Thermomonospora curvata* DSM 42183 (36), shares the overall structure of the C-terminal catalytic domain of LapB from *Pseudomonas* sp. KL28, a representative of type I extradiol ring-cleaving dioxygenases (Fig 7). ArsI from *Bacillus*. sp MD1 (ArsI) has four cysteine pairs. Only the first cysteine pair in ArsI is conserved in all ArsI orthologs. Conserved cysteine residues in arsenic detoxification proteins such as ArsM (Marapakala, Qin et al. 2012), ArsP (Chen, Madegowda et al. 2015) and ArsD (Lin, Yang et al. 2007) form binding sites for arsenic. Therefore, it was reasonable to predict the

conserved Cys96-Cys97 pair is a trivalent organoarsenicals binding site in ArsI. Since the C-terminal cysteine pairs of ArsI are not conserved in other orthologs, they were not considered further in this study. Non-conserved cysteine residues in other arsenic detoxification proteins including ArsR (Xu and Rosen 1997), ArsD (Abdul Ajees, Yang et al. 2011) and ArsM (Marapakala, Qin et al. 2012) are not necessary for their catalytic function, and removal of these non-conserved residues helps in improvement of protein production and crystallization. Therefore, in this study, a truncated ArsI lacking the sequences for residues from Gln125 to the C-terminus, including three nonconserved cysteine pairs, was used as wild-type. To explore roles of the conserved cysteine pair Cys96-Cys97 in ArsI demethylation, we individually or entirely altered the conserved cysteine residues to alanine by site-directed mutagenesis to create C96A, C97A, and C96·97A derivatives. MAs(III) demethylating activity in the single and double cysteine mutants was examined both *in vivo* with *E. coli* cells and *in vitro* with purified proteins. It was reasonable to consider that the mutations have structural changes and it could affect ArsI demethylation activity. Therefore, each cysteine mutants were produced in amounts comparable to that of wild-type ArsI. All purified mutants were soluble in the cytosol, reflecting its normal folding. MAs(III)-demethylation activity of *E. coli* cells expressing ArsI124 and cysteine mutants was determined. All cysteine mutants lost their demethylation activity, and it is reasonable, therefore to infer loss resistant to toxic MAs(III) (Fig 9A). Purified wild-type and mutant ArsI proteins incubated at 30 °C for an hour, only parental

ArsI conferred MAs(III) demethylation and like in *E. coli* cells, all purified cysteine mutant proteins lost MAs(III) demethylation activity (Fig 9B).

ArsI is a type-I extradiol dioxygenase and uses a divalent metal ion, a Fe(II)-dependent MAs(III) demethylase. AkbC and BphC, belonging to the same type-I extradiol dioxygenases contain a triad of three residues as a divalent metal binding site. BLAST identified a triad of three putative metal-binding site residues His5-His62-Glu115 in ArsI. The H5A, H62A and E115A substitutions were individually introduced into ArsI124 amino acid residues lining the Fe(II)-binding site were altered by site-directed mutagenesis. ArsI catalysis activity in *E. coli* cells and with purified putative Fe(II)-binding residue mutants was determined. All single mutants' loss MAs(III) demethylation activity. There is a small arsenic contamination in the MAs(III) product (Fig 9 A-B).

Thus, none of the cysteine and Fe(II) binding mutations confer resistance to toxic MAs(III) in *E. coli* cells and were shown to lose demethylation activity with purified proteins, indicating both conserved cysteine and Fe(II) binding residues are required for ArsI catalysis. Note that two conserved residues could provide two sulfur ligands necessary for trivalent organoarsenical MAs(III) binding. While it is possible that MAs(III) can bind to single-cysteine mutant and Fe(II) to single metal binding mutant, it is not sufficient for ArsI catalysis. Therefore, it was reasonable to study real time arsenic binding assay by using intrinsic tryptophan fluorescence spectroscopy.

3.4 Figures

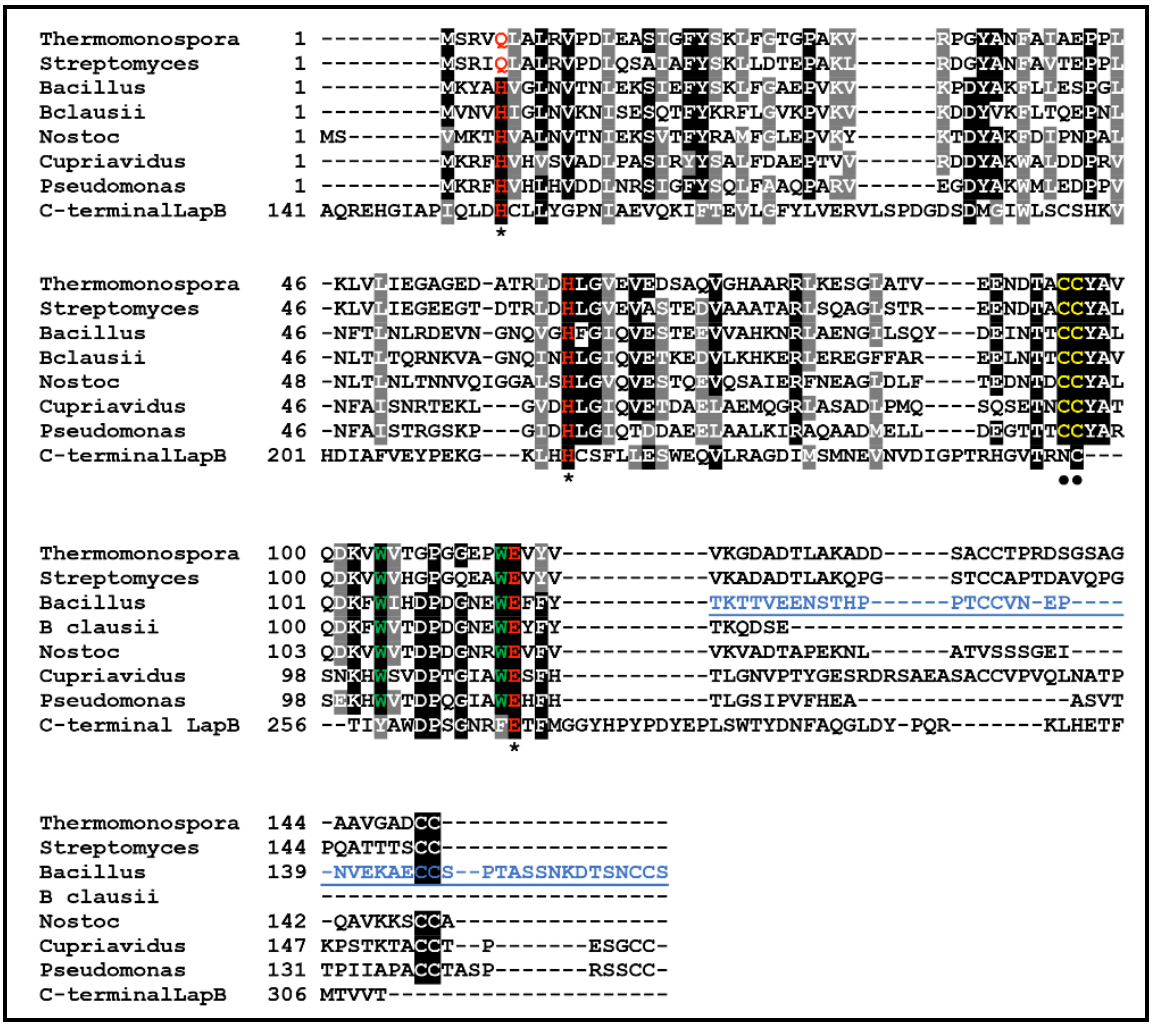


Figure 7. Multiple alignments of *Bacillus* sp. MD1 ArsI homologs. Identities highlighted in black and the conservative replacement in gray. The residues of the Fe(II) binding site are in red with an asterisk. The MAs(III) binding cysteines are in yellow with a circle. Tryptophans are highlighted by green. The underlined blue residues containing three nonconserved cysteine residues indicates removed C-terminal of Bacillus ArsI to construct $ArsI_{124}$ (termed simply ArsI for the purposes of this study).

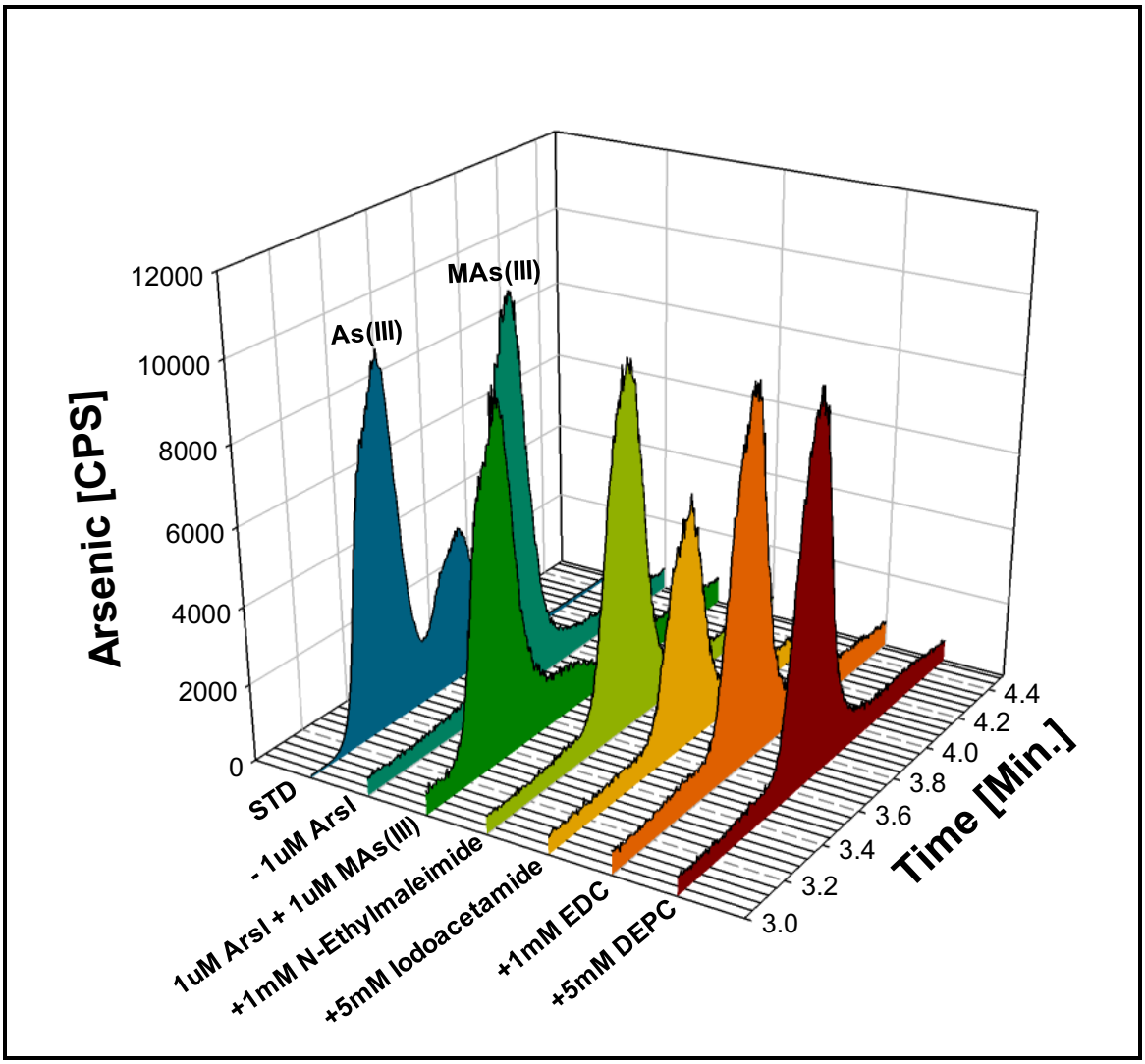


Figure 8. Effect of residue specific chemical modifiers on demethylation activity of ArsI.

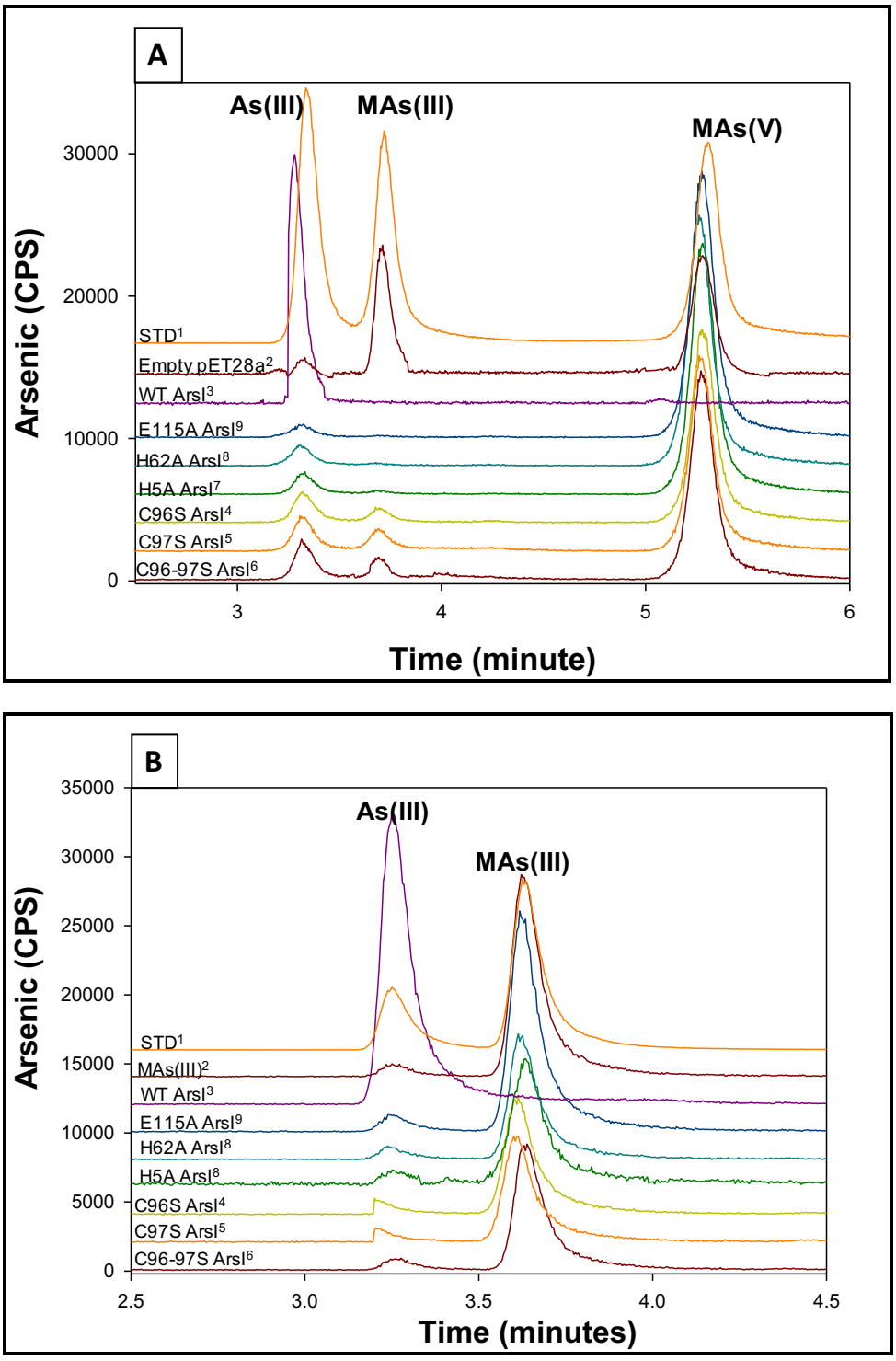


Figure 9. All single mutations in ArsI in the predicted Fe(II) and substrate-binding sites results in loss of activity in both *in vivo* (A) and *in vitro* (B) analyzed by HPLC-ICP-MS.

3.5 Discussion

ArsI is a newly identified enzyme in biology which cleaves a carbon-arsenic bond of trivalent organoarsenicals. The TcArsI metal-bound structure provided information about residues involved in the metal binding site. A new structure with bound Rox(III) was recently determined. Conserved cysteine pairs are the binding sites for arsenicals in many of arsenic resistance proteins. Based on this information, we performed primary studies to determine the importance of metal binding and substrate binding residues using residue specific chemical modifiers (Fig 8). We observed loss of ArsI demethylation activity in the presence of chemical modifiers specific for both metal binding residues (histidine and carboxylic acids) and substrate binding residues (cysteines). These results supported our prediction of conserved residues role in ArsI catalysis. However, chemical modifier results for metal binding residues are not specific since ArsI has additional carboxylic acids.

We conducted a more detailed structure-function analysis of ArsI by site-directed mutagenesis and constructed single metal binding residues mutants (H5A, H62A, and E115A), and single and double substrate mutants (C96S, C97S, C96-97S). ArsI demethylation activity was absent both *in vivo* (in *E. coli* cells) and *in vitro* (with purified proteins) (Fig 9 A-B). The results of mutagenesis suggest that ArsI requires all conserved metal and substrate binding residue for catalysis. While mutation of these residues results in loss of ArsI demethylation

activity, they are negative results that support but do not confirm their roles in catalysis.

3.6 References

Abdul Ajees, A., J. Yang, et al. (2011). "The ArsD As(III) metallochaperone." Biometals **24**(3): 391-399.

Chen, J., M. Madegowda, et al. (2015). "ArsP: a methylarsenite efflux permease." Molecular Microbiology **98**(4): 625-635.

Cho, J.-H., D.-K. Jung, et al. (2009). "Crystal structure and functional analysis of the extradiol dioxygenase LapB from a long-chain alkylphenol degradation pathway in Pseudomonas." Journal of Biological Chemistry **284**(49): 34321-34330.

Green M. R, S. J., and Sambrrok J. (2012). Molecular cloning–A Laboratory Manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Lin, Y. F., J. Yang, et al. (2007). "ArsD: an As(III) metallochaperone for the ArsAB As(III)-translocating ATPase." Journal of Bioenergetics and Biomembranes **39**(5-6): 453-458.

Maki, T., W. Hirota, et al. (2009). "Seasonal dynamics of biodegradation activities for dimethylarsinic acid (DMA) in Lake Kahokugata." Chemosphere **77**(1): 36-42.

Marapakala, K., J. Qin, et al. (2012). "Identification of catalytic residues in the As(III) S-adenosylmethionine methyltransferase." Biochemistry **51**(5): 944-951.

Nadar, S. V., M. Yoshinaga, et al. (2014). "Crystallization and preliminary X-ray crystallographic studies of the ArsI C-As lyase from Thermomonospora curvata." Acta Crystallographica Section F: Structural Biology Communications **70**(Pt 6): 761-764.

Nadar, V. S., M. Yoshinaga, et al. (2016). "Structure of the ArsI C–As Lyase: Insights into the Mechanism of Degradation of Organoarsenical Herbicides and Growth Promoters." Journal of Molecular Biology **428**(11): 2462-2473.

Nogales, J., Á. Canales, et al. (2011). "Unravelling the gallic acid degradation pathway in bacteria: the gal cluster from Pseudomonas putida." Molecular Microbiology **79**(2): 359-374.

Nozaki, M. (1979). "Oxygenases and dioxygenases." Biochemistry: 145-186.

Stice, S. A. (2014). "Speciation, Metabolism, Toxicity, and Protein-binding of Different Arsenic Species in Human Cells."

Xu, C. and B. P. Rosen (1997). "Dimerization is essential for DNA binding and repression by the ArsR metalloregulatory protein of Escherichia coli." Journal of Biological Chemistry **272**(25): 15734-15738.

Yoshinaga, M. and B. P. Rosen (2014). "A CAs lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters." Proceedings of the National Academy of Sciences **111**(21): 7701-7706.

CHAPTER 4. ArsI THERMODYNAMIC STUDIES AND REAL-TIME BINDING ASSAYS OF TRIVALENT ARSENICALS

4.1 Introduction

Arsenic is present in air, water, and soil. The major route of arsenic introduction into the environment is from the earth's crust. Additionally, anthropogenic activities including usage of monosodium methylarsenate [MSMA] as a pesticide and roxarsone in animal husbandry, which contribute to contaminating soil and nearby water bodies (https://www.epa.gov/ingredients-used-pesticide-products/monosodium-methanearsonate-msma-organic-arsenical). According to an EPA 2007 report, MSMA rank fourth among the most commonly used conventional herbicides during the 2006 – 2007 period, with millions of pounds of active ingredient usage in industry, government and commercial market sector (https://www.epa.gov/sites/production/files/2015-10/documents/market_estimates2007.pdf). Historically the use of inorganic arsenic as herbicides and pesticides was largely replaced by MSMA, DSMA, and cacodylic acid. MSMA has been used on cotton fields, golf courses, lawns and residential backyard. However, this less toxic methylarsenical is largely biotransformed into the more toxic inorganic form after introduction into the environment, which leads to environmental pollution. Water percolating through the soil of simulated golf course greens on with MAs(V) shown to be demethylated to As(V) (Feng, Schrlau et al. 2005). The unrestricted use of DSMA (also called cacodylic acid) was canceled in 2009 by the EPA. However, the EPA still allows its use for sod farms, golf courses and highway rights-of-way uses, in

addition to cotton. This practice will be considered during the registration review of MSMA that began in March 2013 and is scheduled to be completed in 2019 (https://www.epa.gov/ingredients-used-pesticide-products/monosodium-methanearsonate-msma-organic-arsenical). Thus, the use of these organoarsenicals is posing a threat to environment and humans which cannot be unnoticed.

4.2 Materials and Methods

4.2.1 Isothermal Titration Calorimetric (ITC) Studies

The binding properties of $ArsI_{124}$ for Fe(II) and substrates were characterized by isothermal titration calorimetry (ITC). For Fe(II) binding, to prevent Fe(II) from oxidation, both protein solution (70 μM $ArsI_{124}$) and ligand solution [1 mM Fe(II)] were prepared in ITC Buffer (50 mM MOPS, 0.5 M NaCl, 5 mM TCEP, pH 7.2) deoxygenated using a Schlenk line and experiments were performed with a VP-ITC microcalorimeter (MicroCal, LCC., Morthampton, MA) in anaerobic condition. The ligand solution was titrated into the adiabatic cell containing the protein solution (2.2 mL) in 2- (initial injection) and 7-μL (subsequent 24 injections) volumes at 10-min intervals. Data were collected at 30 °C with stirring speed set at 350 rpm and analyzed using the Origin 7.0 software (TA instruments, New Castle, DE). For substrates binding, because reagents used to prepare trivalent organoarsenicals by reduction give the high background of heat generation during titrations, $MAs(III)I_2$ and PAO, both of which are free of the reducing reagents, were used as MAs(III) and PhAs(III), respectively. Protein solution (70

μM $ArsI_{124}$) and ligand solutions [0.5 mM MAs(III) or PhAs(III)] were prepared in ITC Buffer degassed by bubbling with pure argon for 30 min. Because these trivalent organoarsenicals are relatively stable compared to Fe(II), binding properties of $ArsI_{124}$ on these substrates were carried out in aerobic condition using MicroCal iTC_{200} (GE Healthcare Bio Sciences, Pittsburgh PA). Each ligand solution was titrated into the protein solution (300 μL) in the adiabatic cell in 0.4- (initial injection) and 1-μL (subsequent 24 injections) at 4-min intervals. Data were collected at 30 °C with stirring speed set at 1000 rpm and analyzed using the Origin 7.0 software provided by MicroCal. and was fit to a one-site binding model, which calculates the stoichiometry of binding and dissociation constant parameters. All experiments were performed in duplicate on independent samples to ensure data reproducibility.

4.2.2 Fluorescence Assays

A temperature-controlled QuantaMaster UV-Vis QM-4 steady state spectrofluorometer (Photon Technology International, Birmingham, NJ) was used for fluorescence measurements. Indicated $ArsI_{124}$ protein derivative (5 μM) in a MOPS buffer consisting (50 mM MOPS (pH 7.2), 0.5 M NaCl, 1 mM TCEP) was excited at 295 nm and its intrinsic tryptophan fluorescence was monitored at 340 nm. 25 μM of arsenical [As(III), MAs(III), MAs(V), *p*-ASA(III), *p*-ASA(V), Rox(III) or Rox(V)] or Fe(II) was added and the resulting fluorescence quenching was analyzed. For all data collections, the temperature was set at 23 °C and spectrum of buffer alone was subtracted to correct background fluorescence and Raman scattering.

4.3 Results

4.3.1 Fe(II) and Substrate Binding Properties of ArsI

In this study, isothermal titration calorimetry (ITC) was performed to characterize the binding properties of wild-type ArsI for Fe(II) (Fig 10) and arsenicals (Fig 11 A-C). To avoid Fe(II) oxidation, binding of Fe(II) was measured anaerobically, whereas substrate binding was measured under aerobic conditions. Fe(II) bound to ArsI with a K_d of 4.14 ± 0.08 μM at a stoichiometry ratio (metal/protein) of 0.96 ± 0.03. These values are comparable with those for other non-heme Fe(II)-dependent dioxygenases (Leitgeb, Straganz et al. 2009). MAs(III) and PhAs(III) bound to ArsI with a similar stoichiometry ratio (n); (MAs(III): 0.65 ± 0.02, PhAs(III): 0.70 ± 0.05), while the dissociation constant of PhAs(III) was lower than that of MAs(III) by three orders of magnitude (MAs(III): 0.30 ± 0.09 μM, PhAs(III): 0.62 ± 0.06 nM), suggesting aromatic arsenicals are preferable substrates. As(III), the product of ArsI reaction, poorly bound to ArsI (n: 0.010 ± 0.002) and showed a lower affinity (K_d: 50.7 ± 0.1 μM) by three to six orders of magnitude compared to MAs(III) and PhAs(III), the ArsI substrates. The affinity difference between the substrate and product may control the steps for substrate binding, product release and catalytic turnover of ArsI.

4.3.2 Tryptophan Fluorescence Spectroscopy

The participation of the residues in Fe(II) and substrate binding during catalysis was further investigated by site-directed mutagenesis. Each of the Fe(II)-binding residues was changed to alanine, and the cysteine residues in the

predicted substrate binding site were individually mutated to serine. All single mutations in the Fe(II)- and substrate-binding residues resulted in the loss of activity *in vivo* (Fig 9A) and *in vitro* (Fig 9B), implying that all those residues are involved in the ArsI catalysis.

We employed ligand-dependent changes in intrinsic tryptophan fluorescence, by which binding of substrates in various arsenic detoxification proteins has been extensively characterized (Zhou and Rosen 1997; Zhou, Shen et al. 2002; Yang, Rawat et al. 2010; Marapakala, Qin et al. 2012), to examine the effects of those single mutations on Fe(II) and substrate binding in ArsI. Wild-type ArsI has two conserved tryptophan residues (Trp105 and Trp114) (Fig 7), which produce intrinsic fluorescence that is quenched by additions of various trivalent organoarsenical substrates (Fig 12) and Fe(II) (Fig 13A). Inorganic As(III) has no effect. These results support the idea that ArsI binds Fe(II) and trivalent organoarsenicals, which is consistent with the ITC results.

All the single mutations in Fe(II)-binding residues resulted in the loss of fluorescence quenching (Fig 13A), suggesting that all the residues in His5-His62-Glu115 triad contribute to Fe(II) coordination in ArsI and are essential to maintaining Fe(II) in the active site. In contrast, mutations in the Fe(II)-binding site showed no effect on binding of arsenical substrates (Fig 14 A-C). These results support our hypothesis that ArsI has an independent substrate binding site, and Fe(II) binding is not required for substrate binding in ArsI, which differs from type I extradiol ring-cleaving dioxygenases that require Fe(II) to bind substrates (Nozaki 1979).

The level of the fluorescence quenching was different for different substrates. Aromatic arsenicals induced larger quenching than the methylated species MAs(III), which is probably because aromatic arsenicals bind to ArsI more tightly than MAs(III), consistent with the results of the ITC experiments (Table 3). Rox(III) and *p*-ASA(III) showed relatively stronger quenching, and among these two, Rox(III) caused faster quenching than *p*-ASA(III) (Fig 12). Even the pentavalent species Rox(V) and *p*-ASA(V), but not either MAs(V) or PhAs(V), induced a small quenching (Fig 12). It is possible that the substituent groups in the benzene rings of those species have some affinity with ArsI, which may be synergistic with binding to the cysteine pair that comprises the binding site for trivalent organoarsenicals. Rox(V) and *p*-ASA(V), but not PhAs(V), contain substituent groups in their benzene rings: Rox(V) has nitro and hydroxyl groups at the meta and para positions, respectively, (Fig 15), and *p*-ASA(V) has an amino group at the para position that may contribute to binding and quenching.

The single mutations in the substrate binding site caused up to 70% decrease of the fluorescence quenching induced by MAs(III) (Fig 13B), PhAs(III) (Fig 13C) and *p*-ASA(III) (Fig 13D), which suggests that both the cysteine residues participate in the substrate binding. Although the magnitude is less than that of wild-type, single mutants in the substrate binding residues still response to these substrates and show smaller but clear fluorescence quenching, suggesting that they retain the affinity as substrates to some extent. In contrast, neither of those single mutations showed an effect on fluorescence quenching induced by

Rox(III) (Fig 13E), which may be because of interactions of the substituent groups in Rox(III) with ArsI that may complement the decrease of the affinity caused by the single mutations. The cysteine double mutant ArsI derivative C96·97S was created, and substrate binding was further investigated. The C96·97S double mutation resulted in complete loss of the quenching induced by MAs(III) (Fig 13B) and PhAs(III) (Fig 13C). A similar tendency was also observed also for *p*-ASA(III) (Fig 13D) and Rox(III) (Fig 13E), the magnitude of fluorescence quenching was less in the cysteine double mutant compared to those in the single cysteine mutants. The small fluorescence quenching by addition of *p*-ASA(III) and Rox(III) was observed in the double mutant, which is again may be because of the interaction of their substituent groups with ArsI. ITC analysis indicated that C96·97S has no affinity with PhAs(III) (Fig 16). These results support our hypothesis that the conserved cysteine pair is the substrate binding of ArsI.

4.4 Figures

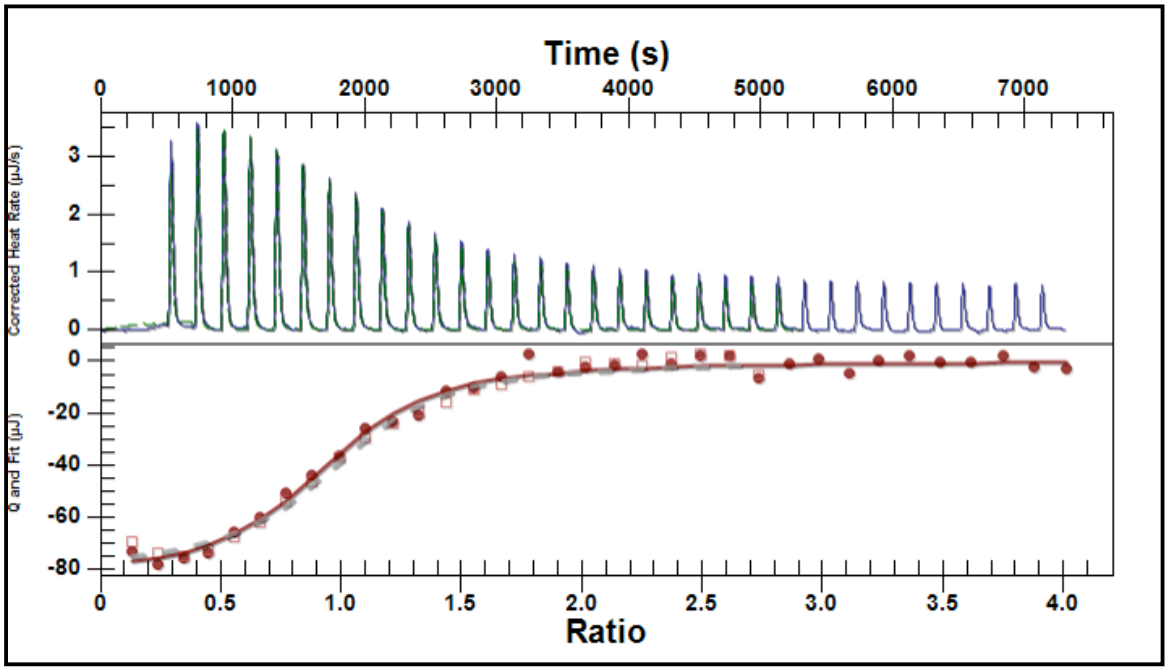


Figure 10. Isothermal titration calorimetric data for ArsI. Raw data isothermal titration calorimetry data and the binding isotherm for ArsI with Fe(II) ligand.

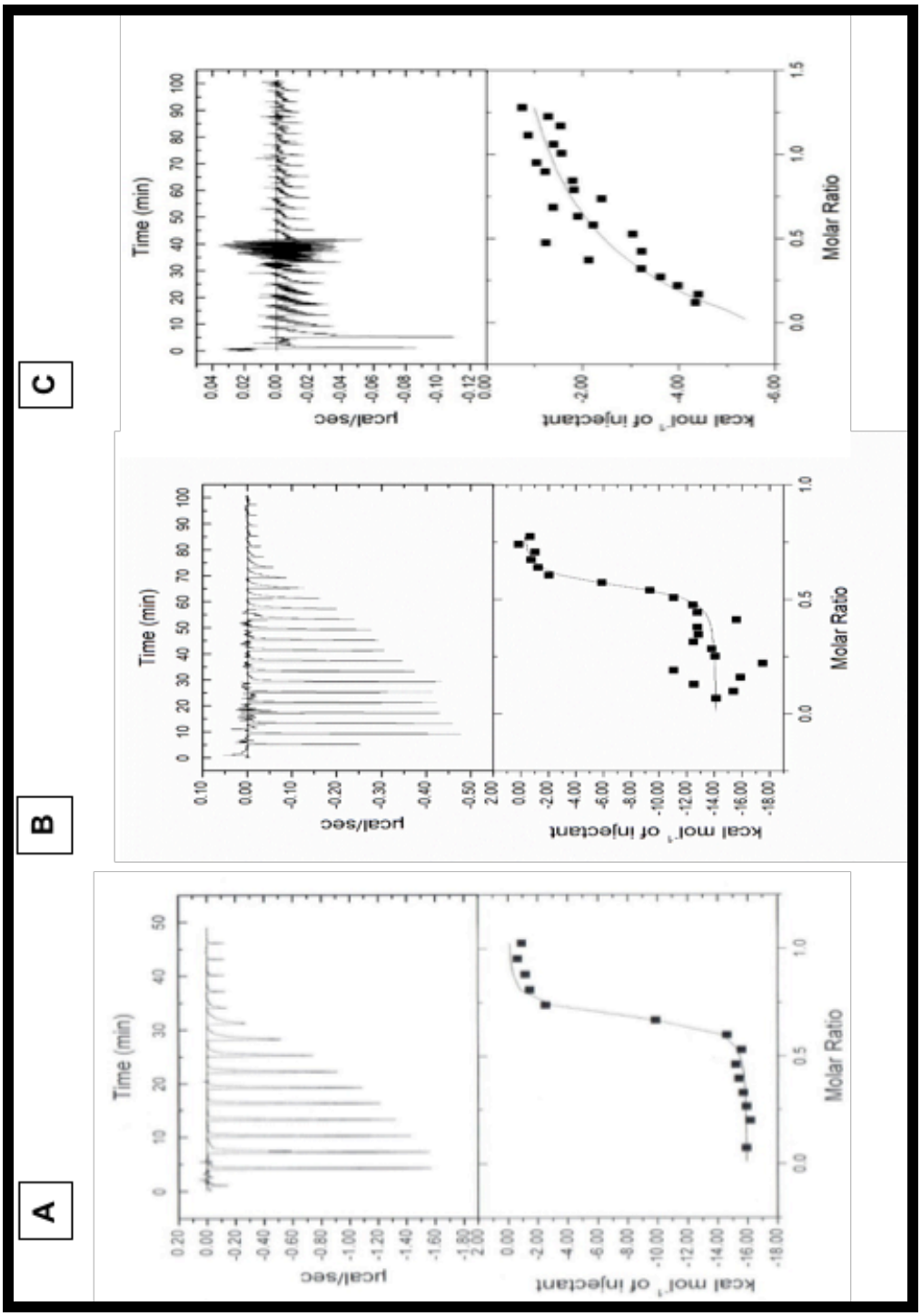


Figure 11. Isothermal titration calorimetric data for ArsI. Raw data isothermal titration calorimetry data and the binding isotherm for ArsI with PhAs(III) (A), MAs(III) (B), and inorganic As(III) (C).

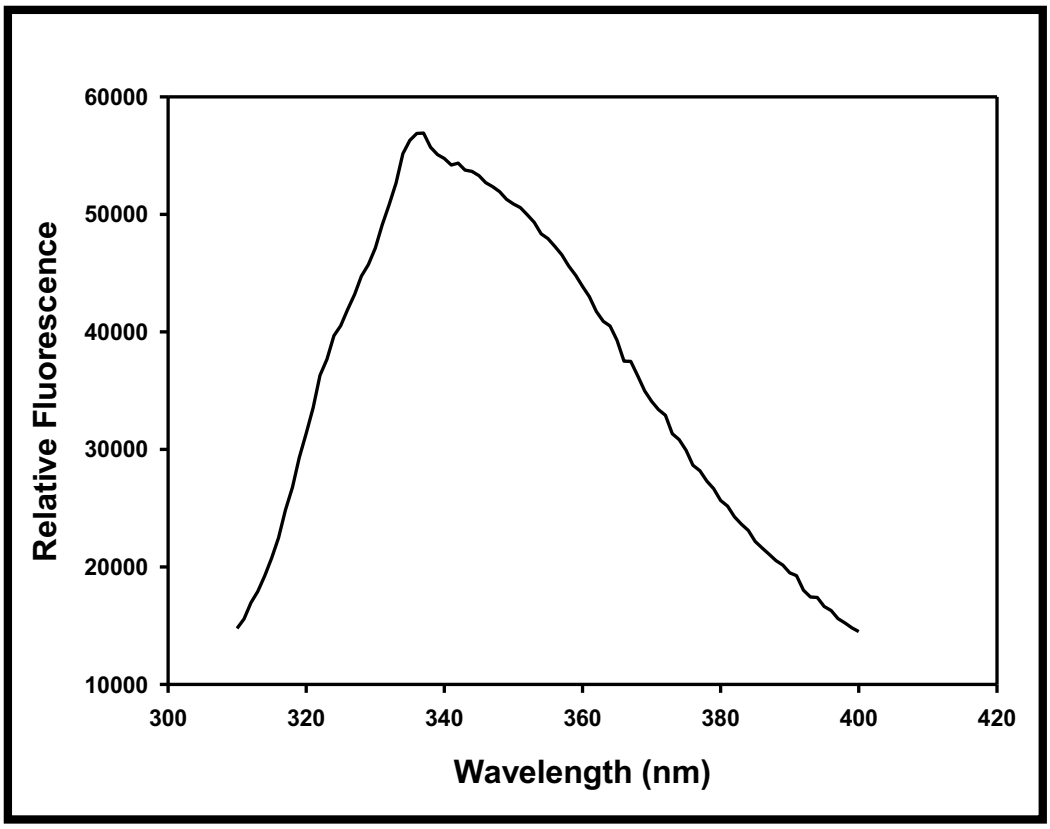
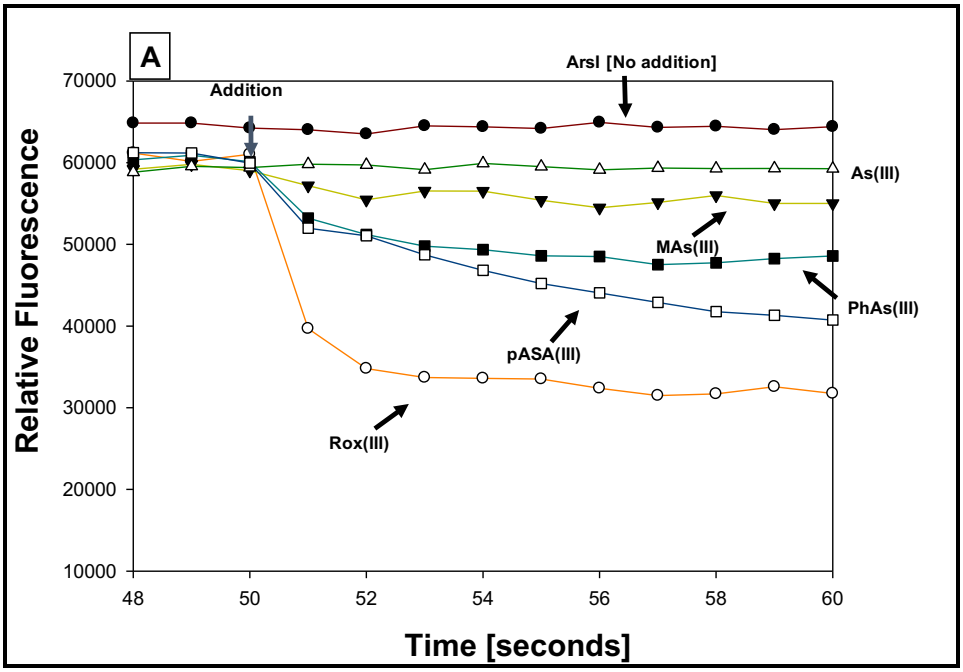


Figure 12. Emission spectra of ArsI. Samples were excited at 295 nm, and emission spectra were acquired at 23 °C as described in Materials and Methods.



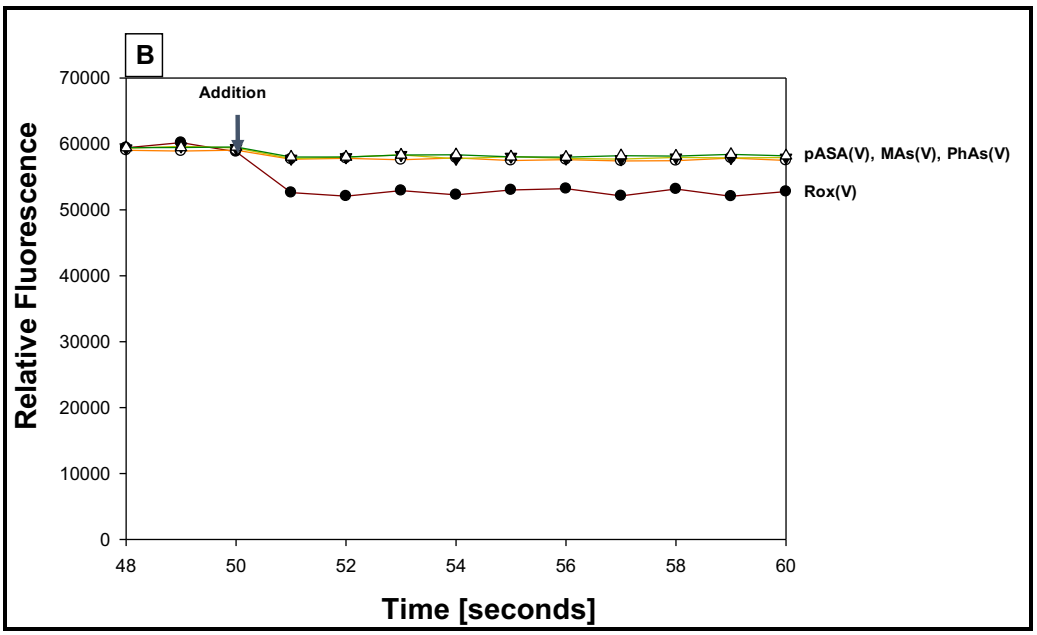
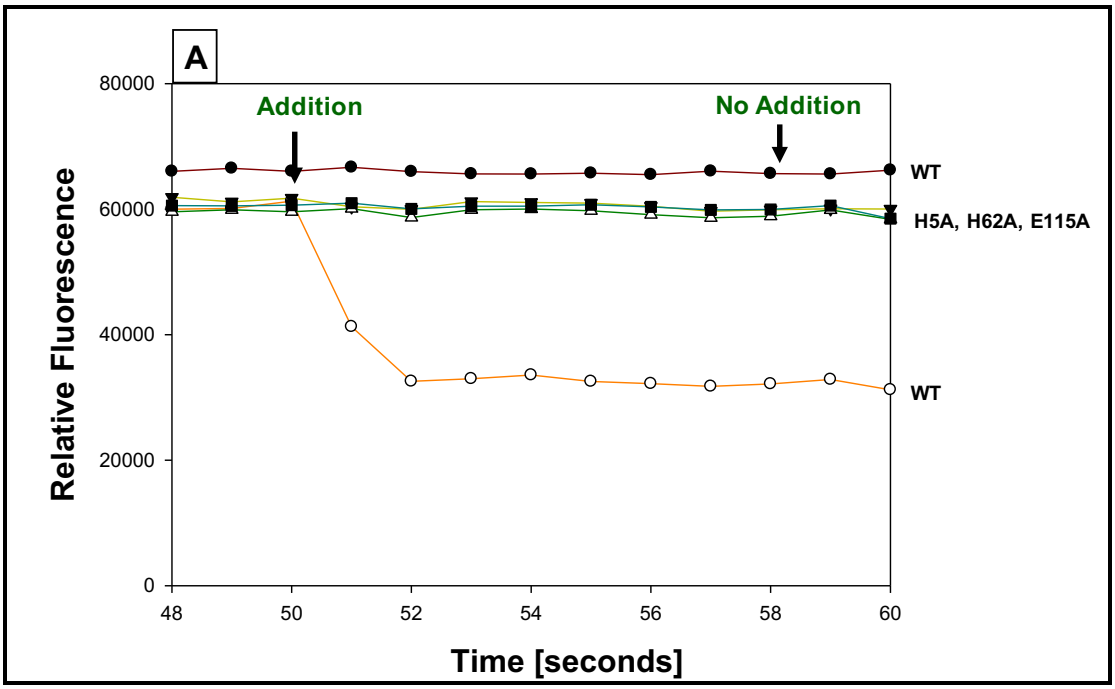


Figure 13. Ligand-dependent quenching of wild-type ArsI protein fluorescence in the presence of trivalent arsenicals (A) and pentavalent arsenicals (B).



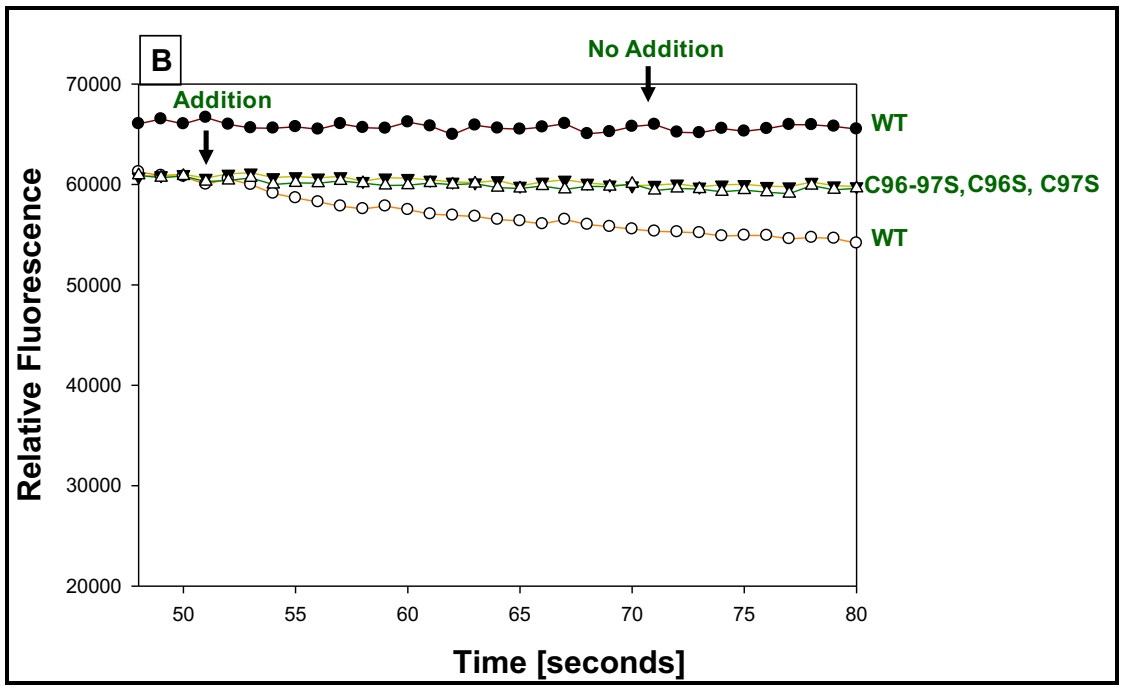
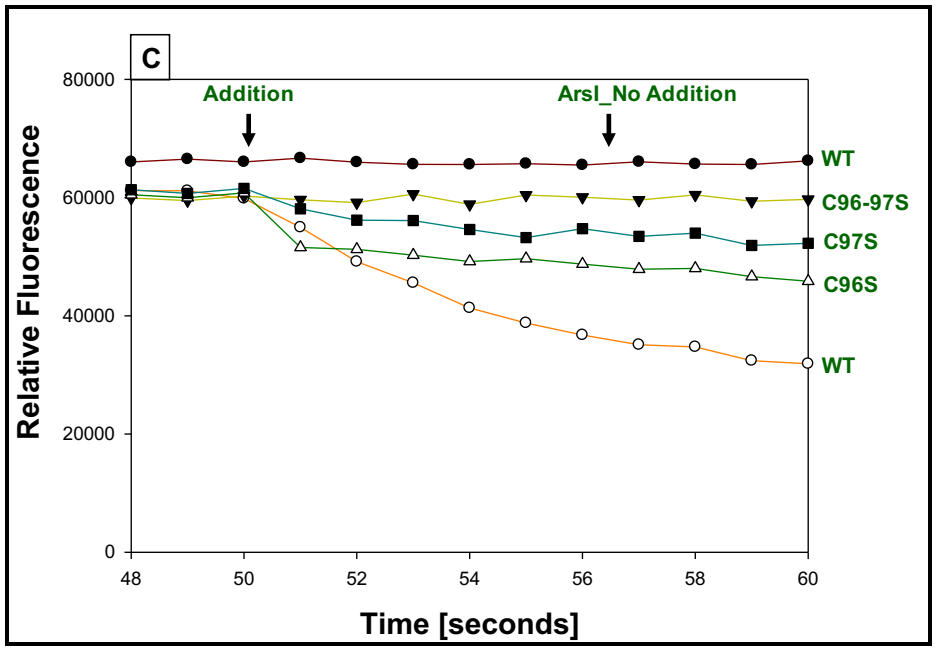


Figure 14. (A) Fe(II) dependent quenching of wild-type and metal binding site mutant ArsI protein fluorescence (B) quenching in the presence of methylarsenite.



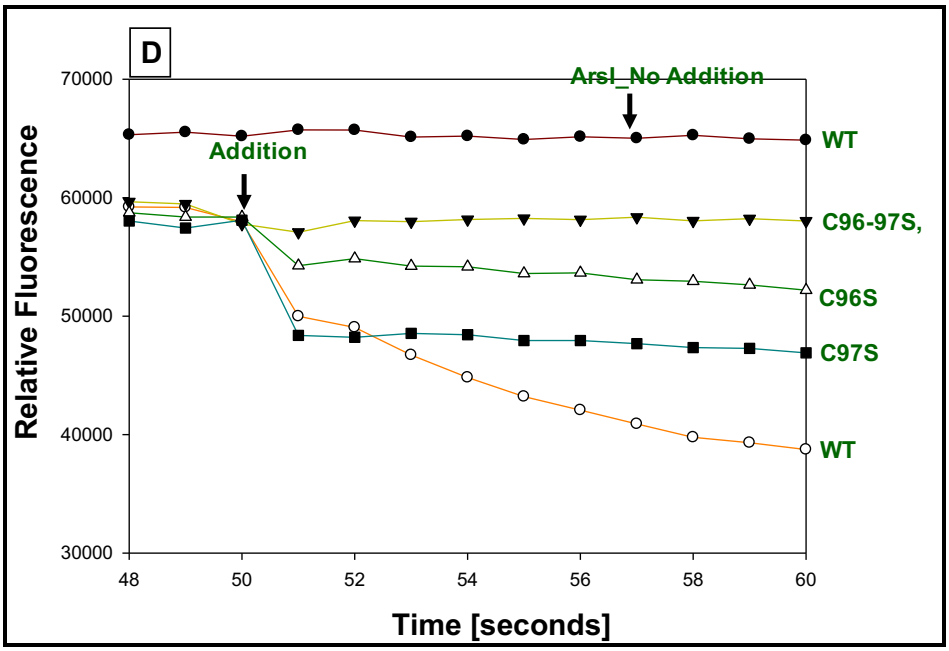


Figure 14. Phenylarsine oxide (C) ligand-dependent quenching of wild-type and cysteine mutants ArsI fluorescence. (D) Ligand-dependent quenching of wild-type and cysteine mutants ArsI protein fluorescence in the presence of trivalent *para*-arsanilic acid.

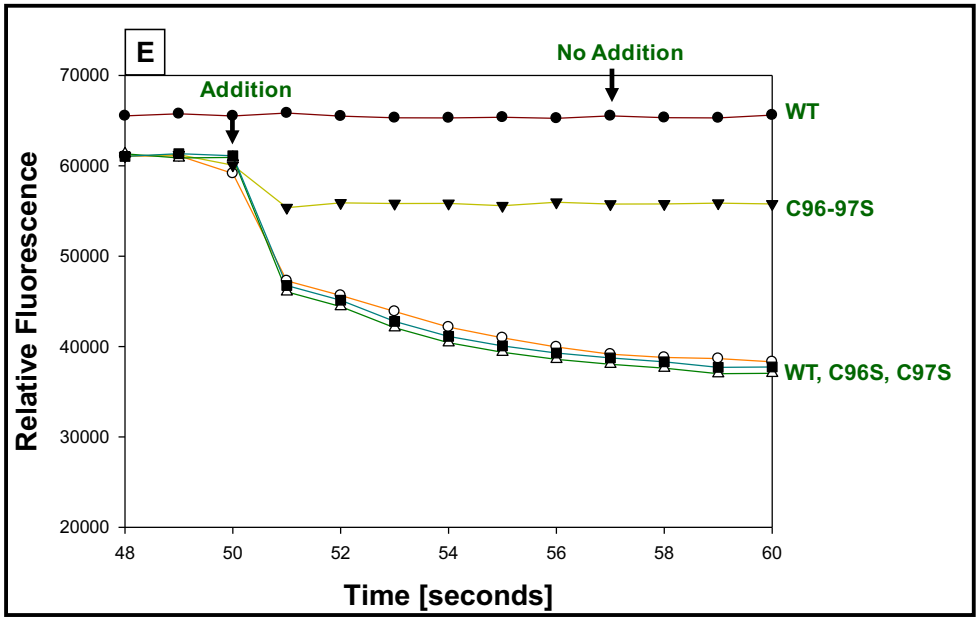


Figure 14. Effect of trivalent roxarsone (E) on the fluorescence of wild-type and cysteine mutants ArsI.

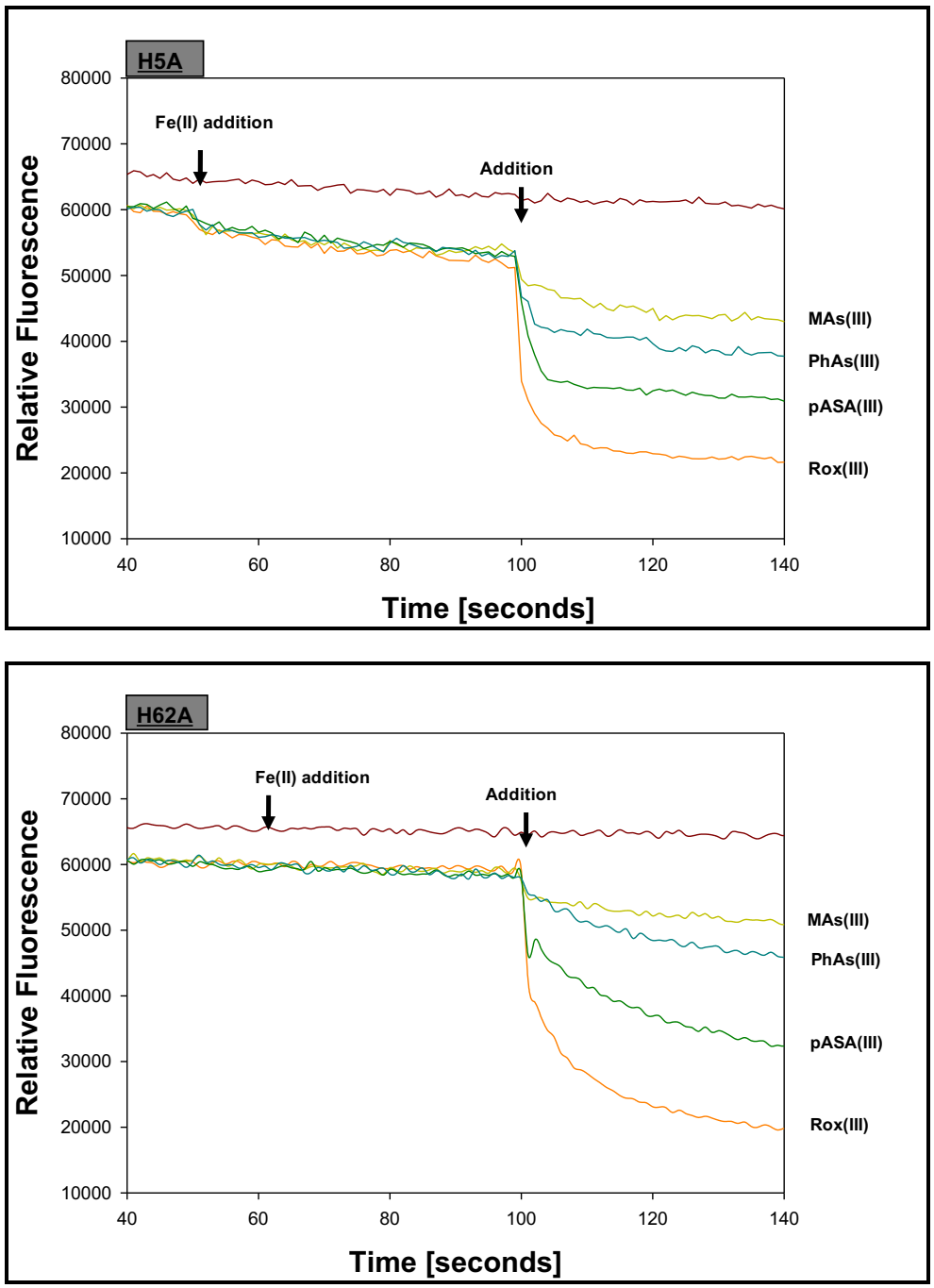


Figure 15. Fe(II) dependent quenching of wild-type and H5A mutant (A) and H62A mutant (B) ArsI protein fluorescence.

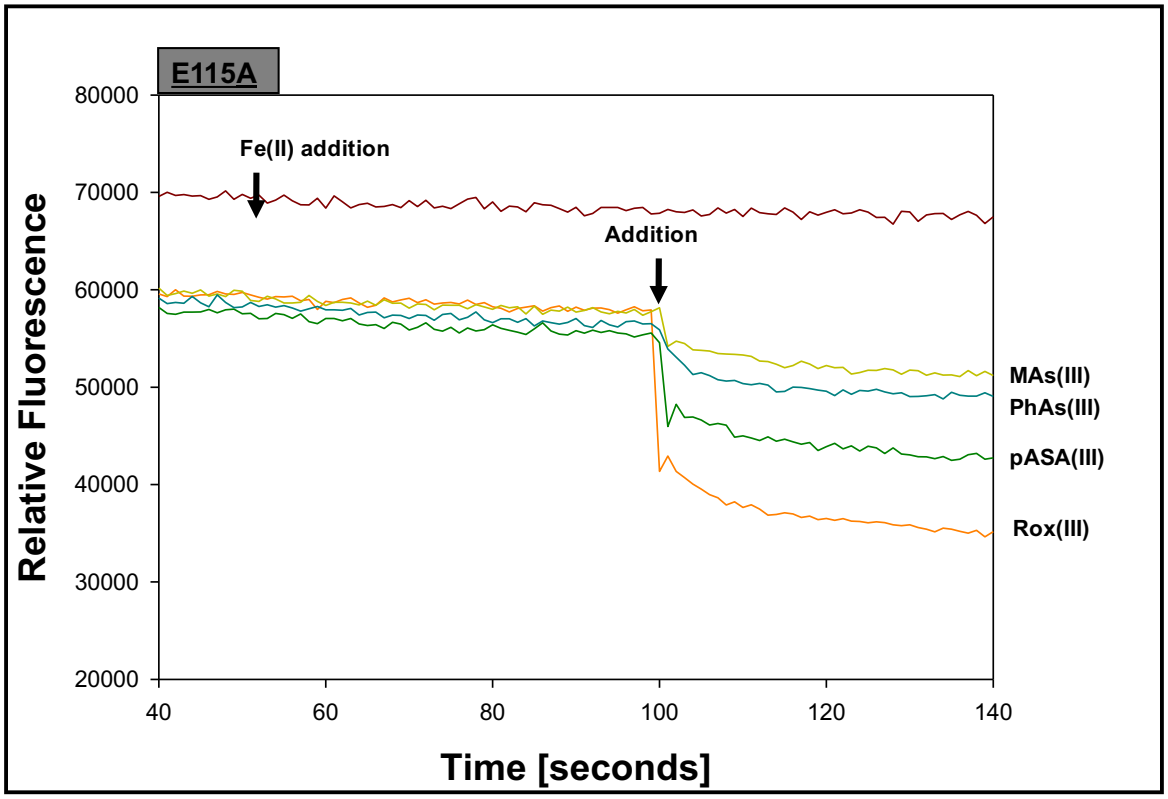


Figure 15. Fe(II) dependent quenching of wild-type and E115A mutant (C) ArsI protein fluorescence.

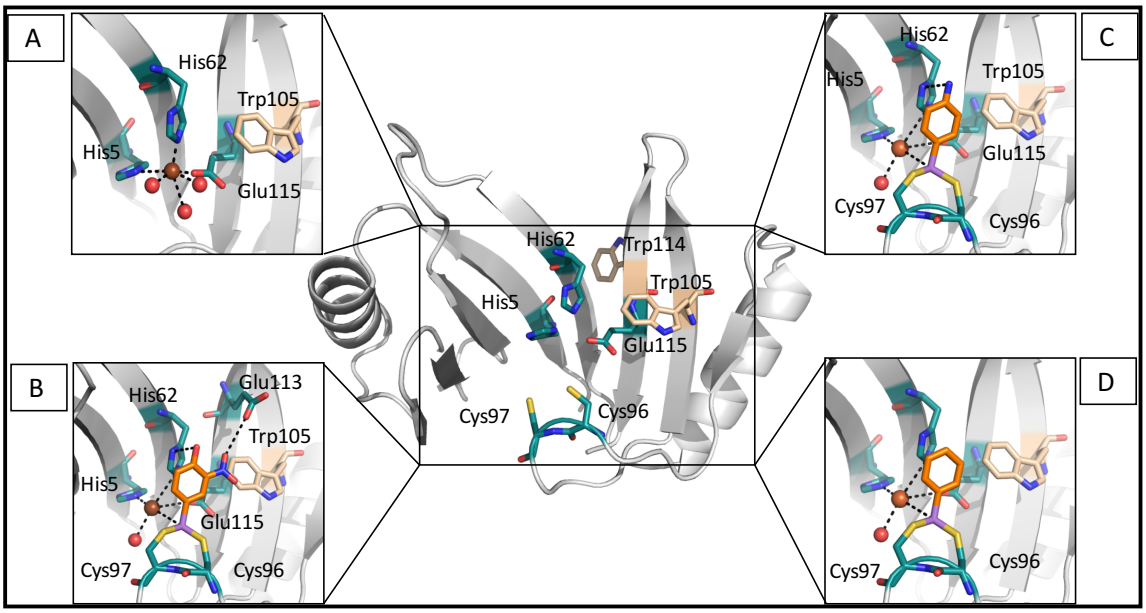


Figure 16. Model of interaction of Bacillus ArsI with (A) Fe(II), (B) roxarsone, (C) para arsanilic acid (D) phenylarsine oxide. The metal and arsenical binding residues are highlighted in cyan color and tryptophan residues are in tan color. (A). Interaction of ArsI with Fe atom (brown), it directly coordinates with protein through His5, His62 and Glu115 residues and water molecules (red)(B). The arsenic atom of Rox(III) binds with cysteine residues (Cys96, Cys97) and it also contacts Fe atom (distance 3.5 Å). The aromatic ring of Rox(III) stacks the five-membered ring of His62 and hydrophobic side chain of Trp105. The hydroxyl and nitro group contact His62 (2.8 Å) and Glu113 (3.4 Å) respectively. (C) pASA(III) also interact similar manner but it loses the contact with GLu113. (D) PhAs(III) loses both Glu113 and His62 contacts. Trp105 is in both metal and arsenical binding sites. From these models, we predicted as the binding affinity of Rox(III) is more the pASA(III) and PhAs(III).

Table 3. Thermodynamic parameters of ArsI determined by ITC

BmArsI-Ligand	K_d	N
Fe(II)	4.14 ± 0.08 μM	0.96 ± 0.03
PhAs(III)	0.62 ± 0.06 nM	0.85 ± 0.05
MAs(III)	0.3 ± 0.09 μM	0.65 ± 0.02
As(III)	50.7 ± 0.08 μM	0.01 ± 0.002

4.5 Discussion

As reported recently, a novel two-step MSMA degradation pathway carried out by a bacterial community from Florida golf course soil. The first step is a reduction of MAs(V) into MAs(III) catalyzed by *Burkholderia* sp. MR1 and the second step is demethylation of MAs(III) to trivalent inorganic arsenic [As(III)] carried out by *Streptomyces* sp. MD1 and *Bacillus* sp. MD1. The trivalent intermediate MAs(III) and the final product As(III) are more toxic than the initial form, MAs(V). The *in vivo* order of toxicity is MAs(III) > As(III) > As(V) > MAs(V). In addition, recently the gene responsible for the second step demethylation of MAs(III) to As(III), termed *arsI*, has been identified (Yoshinaga, Cai et al. 2011; Yoshinaga and Rosen 2014). ArsI has catalytic activity against wide range of trivalent organoarsenicals including MSMA, roxarsone, nitarsone, *para*-arsanilic acid. I performed an ITC study of ArsI with metal Fe(II) and trivalent organoarsenicals. The ITC study showed a stoichiometry of 1 (n=1) for all ligands. The affinity range of ligands is from micro to nanomolar. ArsI has nanomolar affinity for the trivalent aromatic arsenical PhAs(III) by ITC (Table 3). Overall, this study sheds light on the thermodynamic study of ArsI. The other studies of dioxygenases, for instance, Dke1 dioxygenase belongs to a cupin superfamily from *Acinetobacter johnsonii* reported a K_d value of 5 μM for Fe(II) in Dke1. The K_d value is close to what we observed in ITC for ArsI to Fe(II), which is 4 μM (Leitgeb, Straganz et al. 2009). Although ArsI and Dke1 are dioxygenases, ArsI belongs to type-1 extradiol dioxygenase and Dke1 from a cupin superfamily.

Recently, structures of TcArsI with and without metals Ni(II) or Co(II) were solved. Metal ions bind to the protein at the Gln5-His-62-Glu115 triad. TcArsI is an ortholog of *Bacillus* ArsI, and His5-His62-Glu115 are the corresponding triad. So, we proposed that these residues comprise the metal binding site in *Bacillus* ArsI. A new structure with bound PhAs(III) defines the structure of the substrate binding site in ArsI. Multiple sequence alignment shows cysteine pair at residues 96 and 97 is conserved in all ArsI orthologs. Other arsenic resistance proteins have conserved cysteine residues that act as arsenicals binding site (Dheeman, Packianathan et al. 2014). Based on this information, we made a C96-97S double mutant and compared PhAs(III) binding affinity with wild-type ArsI. The ITC study shows no binding of PhAs(III) to cysteine double mutant (Fig 16). These data support our hypothesis that the conserved Cys96-97 pair is part of the substrate binding site.

Loss of ArsI catalysis in mutants is consistent with the contribution of these residues in ArsI activity, but mutagenesis is intrinsically a negative result, and mutants do not provide positive information on the binding of ligands to ArsI. To elucidate the effect of ArsI mutation on the binding of ligands in more detail, real-time fluorescence spectroscopy was performed. ArsI wild-type and mutants showed informative results on the binding pattern of both metal and substrate ligands. We observed that ArsI has the highest affinity for aromatic arsenicals, with the order of affinity: Rox(III) > pASA(III) > PhAs(III) > MAs(III) ≅ Rox(V) (Fig 12). All single cysteine mutants showed decreased in the magnitude of the

fluorescence compared to wild-type ArsI (Fig 13 B-E). Single amino acid substitutions to alanine in the metal binding site resulted in decreased fluorescence quenching with Fe(II), but there is little or no effect on organoarsenical quenching (Fig 14 A-C). However, there was a small quenching with p-ASA(III) and Rox(III) observed with cysteine double mutants. Ideally, in the absence of substrate binding site, there should not be any fluorescence quenching expected. A *Bacillus* ArsI model based on crystal structure of TcArsI suggests an explanation for a fluorescence quenching with cysteine double mutant by aromatic arsenicals Rox(III), Rox(V) and pASA(III). The model shows the functional group [amino in p-ASA(III) and hydroxyl in Rox(III), Rox(V)] at the C4 position of these arsenicals interacts with His62 and Glu113. ArsI has two tryptophans at position 105 and 114. Also, Trp105 can stack with the aromatic group of the substrate. It also indicates that the reason for no fluorescence quenching with phenylarsenite, another trivalent aromatic arsenical, is that there is no functional group at C4 position to interact with His62 or Glu113 (Fig 15). Thus, protein fluorescence quenching data suggest the observed decreased in fluorescence is because of the cumulative interaction of aromatic arsenicals with cysteine and interaction of side group interactions. The fluorescence data of trivalent and pentavalent roxarsone in the presence of iron shows decreased in cysteine double mutants quenching because the interaction of roxarsone hydroxyl group with His62 is minimized by Fe(II) binding. It is important to note that His62 is a conserved residue in the Fe(II) binding site His5-His62-Glu115 of *Bacillus* ArsI.

To summarize, this study contributed thermodynamic data by ITC which provided ArsI dissociation constants and stoichiometry (Table 3). This study also characterized the role of the conserved metal binding site triad (His5-His62-Glu115) and catalytic residues (Cys96-97) in ArsI enzymatic activity.

4.6 References

Dheeman, D. S., C. Packianathan, et al. (2014). "Pathway of human AS3MT arsenic methylation." Chemical Research in Toxicology **27**(11): 1979-1989.

Feng, M., J. E. Schrlau, et al. (2005). "Arsenic transport and transformation associated with MSMA application on a golf course green." Journal of Agricultural and Food Chemistry **53**(9): 3556-3562.

Leitgeb, S., G. D. Straganz, et al. (2009). "Biochemical characterization and mutational analysis of the mononuclear non-haem Fe^{2+} site in Dke1, a cupin-type dioxygenase from Acinetobacter johnsonii." Biochemical Journal **418**(2): 403-411.

Marapakala, K., J. Qin, et al. (2012). "Identification of catalytic residues in the As(III) S-adenosylmethionine methyltransferase." Biochemistry **51**(5): 944-951.

Nozaki, M. (1979). "Oxygenases and dioxygenases." Biochemistry: 145-186.

Yang, J., S. Rawat, et al. (2010). "Arsenic binding and transfer by the ArsD As(III) metallochaperone." Biochemistry **49**(17): 3658-3666.

Yoshinaga, M., Y. Cai, et al. (2011). "Demethylation of methylarsonic acid by a microbial community." Environmental Microbiology **13**(5): 1205-1215.

Yoshinaga, M. and B. P. Rosen (2014). "A CAs lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters." Proceedings of the National Academy of Sciences **111**(21): 7701-7706.

Zhou, T. and B. P. Rosen (1997). "Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase." Biological Chemistry **272**(32): 19731-19737.

Zhou, T., J. Shen, et al. (2002). "Unisite and multisite catalysis in the ArsA ATPase." Journal of Biological Chemistry **8**(6): 23815-23820.

CHAPTER 5. DETERMINATION OF THE CARBON PRODUCT OF ARSI DEGRADATION OF ORGANOARSENICALS

5.1 Introduction

Organoarsenicals such as monosodium methylarsenate and dimethylarsinic acid has been used as an herbicide and pesticide in agriculture for a decade (Bednar, Garbarino et al. 2002). Moreover, animal drugs such as roxarsone, nitarsone used as a feed additive to prevent chicken parasitic disease, coccidiosis, support animal growth, tissue pigmentation; to treat histomoniasis in turkey and chickens [(Chapman and Johnson 2002) https://www.fda.gov/animalveterinary/safetyhealth/productsafetyinformation/ucm257540.htm]. Despite these advantages in agriculture and poultry industry, many reports indicate a health risk of these arsenicals released into the environment (Feng, Schrlau et al. 2005). The European Union considered the threat of these arsenicals use and after release into the environment and banned its use in 1999. In 2001, they banned chicken manure in cow feed. In the U.S., the animal drug producer Alpharma, a subsidiary of Pfizer, Inc. and Zoetis Inc. voluntarily suspended animal drugs 3-Nitro® (roxarsone), arsanilic acid, nitarsone and carbarsone by the end of 2015 [FDA, (Huang, Yao et al. 2014)]. However, these drugs are still produced and regularly used in China, India and many developing countries (Bednar, Garbarino et al. 2003; Gandhi 2008).

The standard practice is the use of 50 ppm roxarsone for eight weeks as an antimicrobial, followed by withdrawal for five days before slaughter with intent to decrease roxarsone residuals below 0.5 ppm. According to the Code of Federal Regulations Section 121.1138, the residual arsenic limit in muscle meat and edible byproducts such as liver are 0.5 p.p.m and 1 p.p.m, respectively (Morrison 1969).

Analysis of chicken liver extract by ICP-MS and ESI-MS confirmed the presence of As(III), DMAs, MAs, As(V), 3-AHPAA species in the extracts. Similar compounds were identified in breast meat sold in US supermarket (Nachman, Baron et al. 2013; Peng, Hu et al. 2014). A direct arsenical metabolite in chicken meat is a cause of concern, but so is biotransformation of roxarsone into toxic arsenic in chicken excretions (Nachman, Baron et al. 2013; Peng, Hu et al. 2014). The major concern is because 90% of poultry litter is applied to agricultural lands (Moore, Daniel et al. 1995). Different factors including photodegradation, composting and biotransformation shown to facilitate degradation of roxarsone into carcinogenic arsenic (Bednar, Garbarino et al. 2002; Bednar, Garbarino et al. 2003; Cortinas, Field et al. 2006). Poultry litter containing roxarsone excreted by chickens was shown to be degraded into As(V) after composting a manure mixed with water at 40 °C in 30 days. Zhan et al. studied the comparative cytotoxicity and arsenic accumulation after *Tetrahymenathermophile thermophile* cell exposed to roxarsoen and its photodegradates. They reported that the degraded arsenicals were 10-fold more

toxic than roxarsone in consideration with IC_{50} in *T. thermophile* single cell parasite A (Zhang, Xu et al. 2015). *T. thermophile* is a ubiquitous freshwater unicellular protozoan, which offers both cellular structural and functional complexity in comparison to metazoan cells. They also reported *T. thermophile* arsenic uptake in both the cell membrane and cytoplasm after exposure to *p*-arsanilic acid. In addition, the products of biotransformation are more water soluble than roxarsone, which increases its leaching process, contamination of nearby water bodies and uptake by plants.

Morrison et al. studied the effect of the use of arsenic contaminated poultry litter and reported no effect on arsenic soil content or crops are grown using such litter. However, the assay method used to detect arsenic was a less sensitive colorimetric method (Morrison 1969), so the significance of this study is questionable.

Moody and Williams studied roxarsone metabolism in hens. The study compares oral and intramuscular (IM) dose effect on the fate of roxarsone. They reported that the oral dose was excreted drug slowly compare to IM. The oral dose took 9 to 11 days for complete excretion of roxarsone, compared to 3 days when injected by IM route. Moreover, the main metabolite of roxarsone observed when administered under both oral and IM route was 3-amino-4-hydroxy phenylarsonic acid [3A4HPAA]. Interestingly, reduction of the roxarsone nitro group is related to diet, as starved hens showed less reduction compare to well-fed hens, and the reduction process was reported to occur in the crop organ

(Moody and Williams 1964). Frensemeier et al. recently reported oxidative transformation of roxarsone by electrochemistry [EC]. They used an electrochemical cell consisting of a boron-doped diamond electrode, a titanium auxillary electrode, and a Pd/H2 reference electrode, and applied a potential to oxidize roxarsone. Foe identification, EC effluent was coupled to HILIC and ESI-MS. Additionally, complementary HILIC-ICP-MS was used to confirm the resulting arsenic species. They identified many arsenic-containing products, including hydrogenated and dehydrogenated roxarsone. They also reported few non-arsenic products, suggesting C-As bond cleavage. The ICP-MS study showed toxic As(V) was the main metabolite (Frensemeier, Büter et al. 2017). It should be emphasized, however, that this method is nonenzymatic and nonphysiological. These findings are like those of Oshea et al., who also reported similar products of phenylarsonic acid degradation. They performed ϒ radiolysis under conditions to generate hydroxyl radicals and degrade phenylarsonic acid [PA]. The PA degradation products identified using HPLC-ICP-MS were mono- and di-hydroxy PA, inorganic arsenite, and arsenate. Additionally, they also proposed ipso-addition of the hydroxyl radical, resulting in hemolytic cleavage of the C-As bond and reported the presence of phenol by HPLC analysis (Xu, Kamat et al. 2007).

Although numerous studies have demonstrated roxarsone degradation and identified the products, few identified the non-arsenic products (Frensemeier, Büter et al. 2017). The non-arsenic products of the metabolism and

biotransformation of roxarsone have not been fully elucidated yet. To our knowledge, we reported the first molecular mechanism of an enzyme involved in degradation of organoarsenicals and identified the products by a combination of ICP-MS and ESI-MS.

5.2 Materials and Methods

5.2.1 Separation of Enzymatically Degraded Roxarsone Product

Enzymatically degraded roxarsone products separate by integrated HPLC system (Thermo Scientific, USA). Arsenic species after enzymatic degradation analyzed by Inertsil C4 HPLC column (150 mm X 2.1 MM; 5µ; GL Sciences).

Sample injection (10 µL) and LC separation were performed by a Thermo Surveyor Plus HPLC system equipped with a hydrophilic interaction liquid chromatography purchased from Merck (ZIC-pHILIC 150 X 4.6, 5um) protected by a guard column. Mobile phase gradient was performed between 10 mM ammonium formate dissolved in Optima LC-MS grade water with pH adjusted to 9 with ammonium hydroxide (aqueous) and Optima LC-MS grade acetonitrile (organic). (LC Method: 20170222_HILIC_Rosen_V2). MS Detection was performed by an AB Sciex QTRAP 5500 Triple-Quadrupole mass spectrometer, equipped with a Turbo V™ ESI ionization source. The instrument was operated under full scan, negative mode (Acquisition method: 20170222_neg_Q3MS_100-600_15MIN.dam).

The reaction mixture contained MOPS buffer, Fe(II) and reducing agents (TCEP, cysteine) in a. 1:1 protein to substrate ratio were used in the reaction

mixture, and the reaction was carried out at 30° C for 3 hours at 200 rpm in a shaking incubator. The reaction mixture was filtered through a 3-kDa cut-off Amicon Ultrafilter to remove protein after the reaction was completed. The enzyme-free filtrate was lyophilized, dissolved with water:Isopropanol [50:50 v/v], and analyzed by liquid chromatography with gradient elution on a hydrophilic interaction liquid chromatography [HILIC] protected by a guard column. MS detection was performed by Triple-Quadrupole mass spectrometer, equipped with a ESI ionization source.

5.2.2 Rox(III) Degradation Spectral Change

Purified 100μM ArsI were incubated in a buffer consist of 0.1 M MOPS, 0.15 M NaCl (pH 7.2) containing 1 mM cysteine, 3 mM TCEP and 0.1 mM Fe(II). The reaction was initiated by addition of 200 μM of Rox (III), incubated at 30°C with a medium shaking for 1 h in 96 well plate in Synergy Hybrid H4 plate reader.

5.3 Results

Determination of Roxarsone Degradation Product of ArsI

ArsI, a C-As lyase has catalysis ability on different trivalent organoarsenicals including methyl arsenite, roxarsone, nitarsone and *p*-arsanilic acid (Yoshinaga and Rosen 2014). HPLC-ICP-MS is useful to determine arsenic species after degradation of these trivalent organoarsenicals by ArsI. However, the C-product [non-arsenic part] after ArsI catalysis was unidentified.

The antimicrobial growth promoter Rox(III) has a characteristic absorption spectrum when catalyzing *in vitro* by ArsI as a function of time (Fig. 17).

Roxarsone has absorption at 420nm that is blue shifted when incubated with purified ArsI, and shows a change in λ_{max} to 350nm. It has a potential to be a real-time assay of trivalent arsenical by ArsI.

The product of Rox(III) cleavage is uncertain. If the ring is cleaved, the product might be 4-hydroxy-5-nitro-hexa-2,4-dienal or a similar open chain compound. If the ring is not cleaved, it might be 4-hydroxy-2-nitrophenol or a related compound. To determine the product of roxarsone degradation, mass spectrometry of the roxarsone and roxarsone degradation product was performed.

A simple liquid chromatography coupled with electron spray ionization mass spectrometry [LC-ESI-MS] used for the determination of trivalent roxarsone [Rox(III)] degradation carbon product. Purified ArsI was used to cleave the Rox(III).

To achieve optimum separation of roxarsone and its degradation product, I examined the separation conditions, such as pH of the mobile phase [pH from 7.0 to 9.5], the concentration of ammonium formate [from 10 mM to 20 mM], the concentration of acetonitrile [0 to 20%], the flow rate of the mobile phase [0.5 to 1.5 mL/min]. Using 10 mM ammonium formate and 5% acetonitrile as the mobile phase, I could achieve separation of roxarsone and its degradation product. The total elution time was 15 min. I chose an ammonium formate aqueous solution with ammonium hydroxide (aq) and acetonitrile (organic) as the mobile phase

because it is compatible with ESI-MS detection. Typical chromatograms obtained from HILIC separation with low resolution [LR] ESI-MS (Fig 18).

5.4 Figures

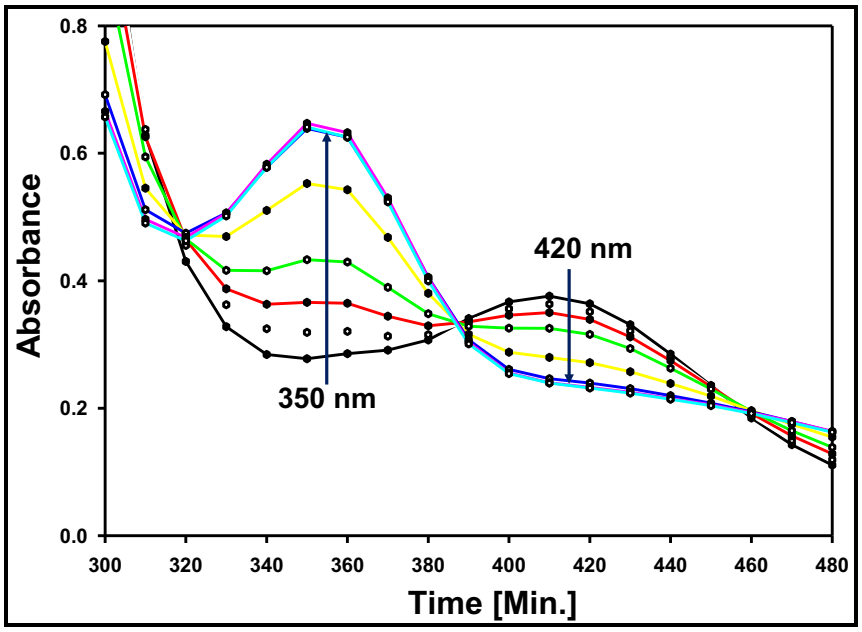
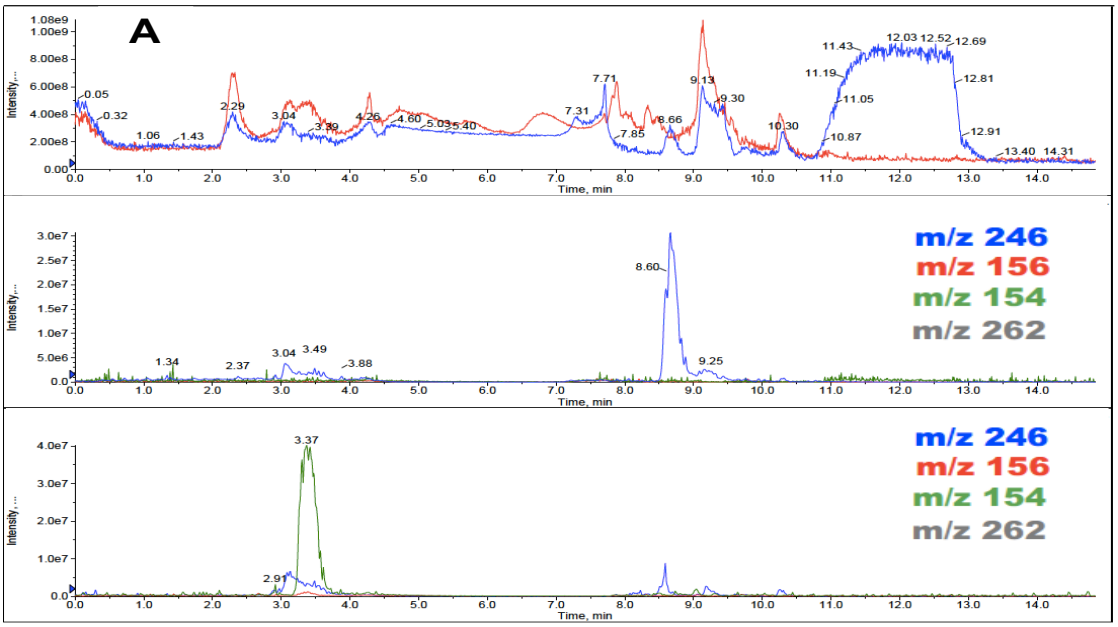


Figure 17. Rox(III) spectral changes: a potential real-time assay of catalysis.



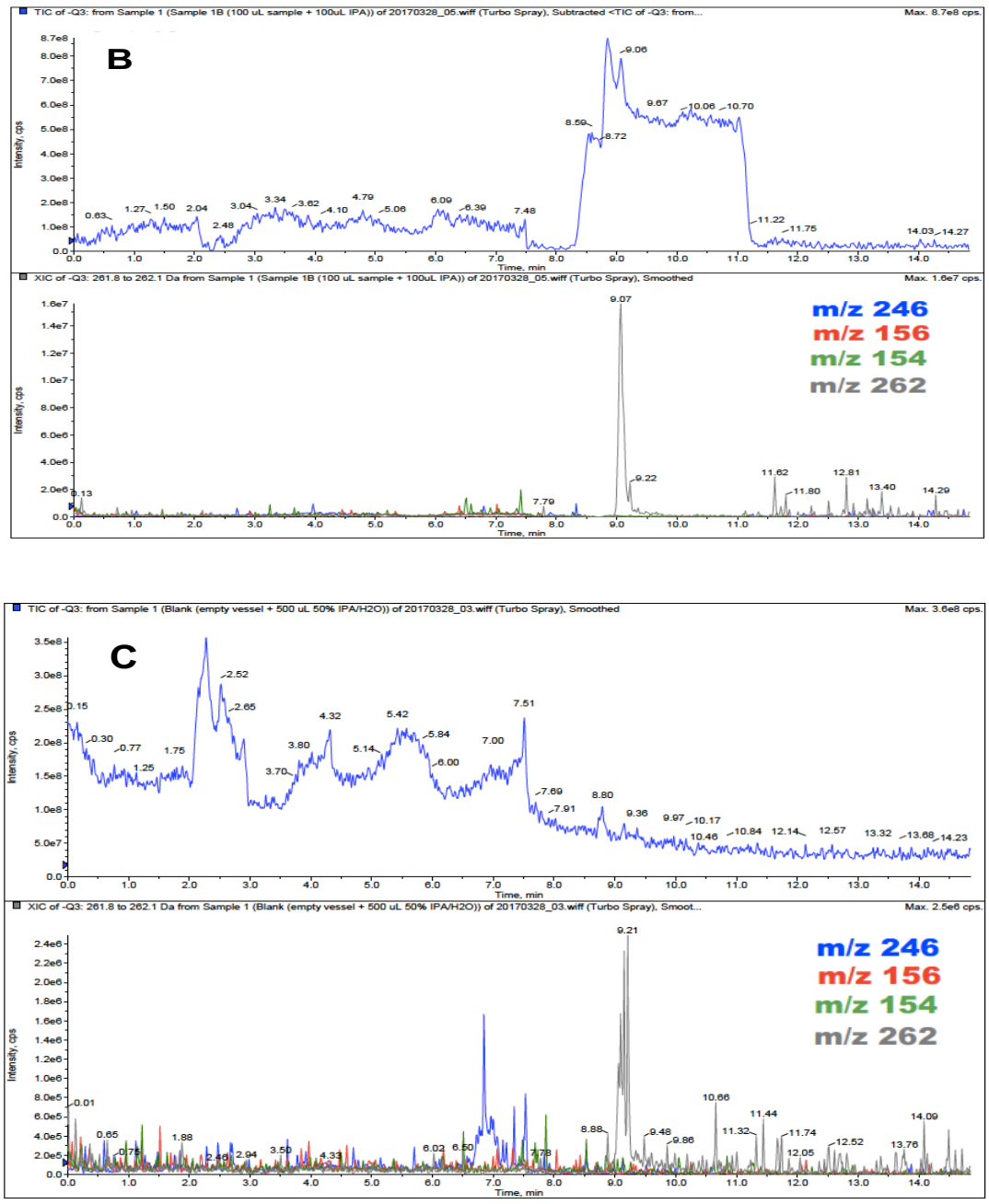
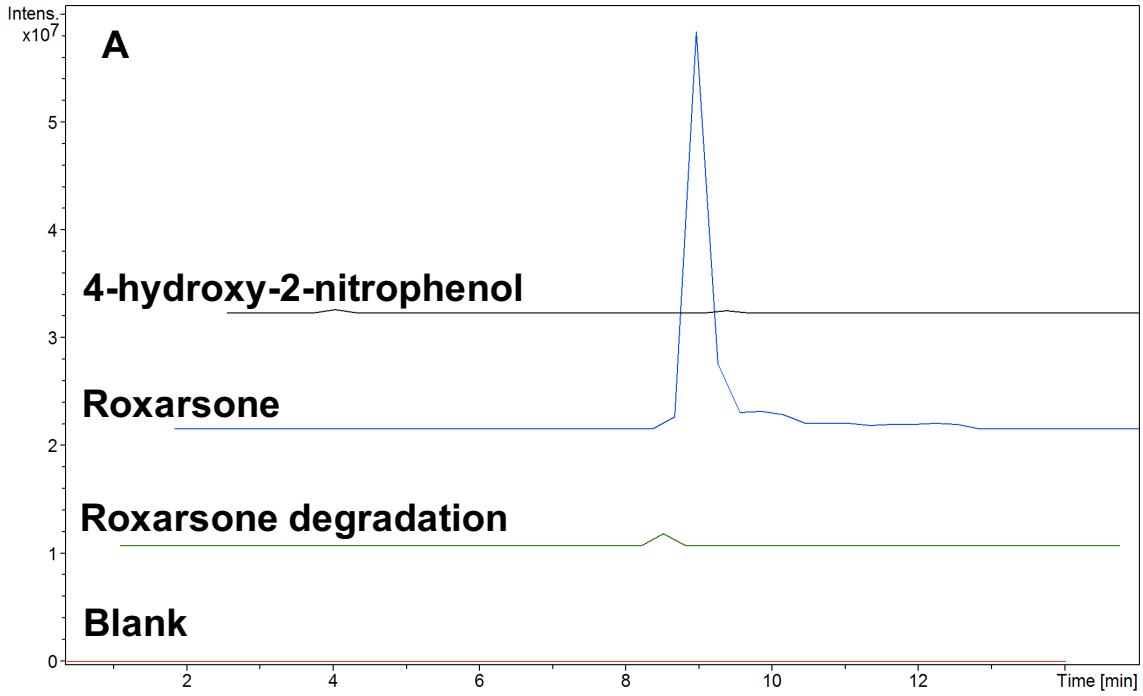
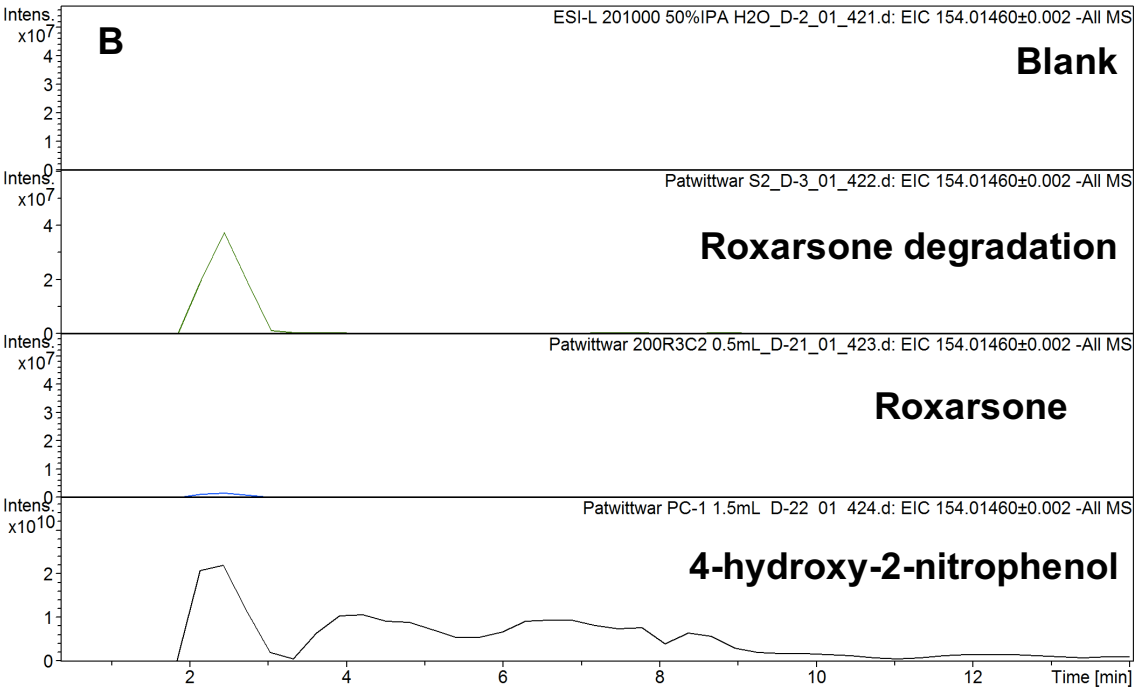


Figure 18. Chromatograms showing separation of Rox(III) [blue] from roxarsone degradation product [red] with LR-ESI-MS detection [A], Rox(V) [B] and water as blank [C]. Depicted are the extracted ion chromatograms of detectable products as well as their retention times. m/z at 156 correspond to 4-hydroxy-5-nitro-hexa-2,4-dienal; m/z at 154 correspond to 4-hydroxy-2-nirophenol, m/z at 246 correspond to Rox(III) and m/z at 262 correspond to Rox(V). Water blank show no target compounds contamination.


A
Intens.
x10^7
5
4
3
2
1
0
4-hydroxy-2-nitrophenol
Roxarsone
Roxarsone degradation
Blank
2
4
6
8
10
12
Time [min]


B
ESI-L 201000 50%IPA H2O_D-2_01_421.d: EIC 154.01460±0.002 -All MS
Blank
Patwittwar S2_D-3_01_422.d: EIC 154.01460±0.002 -All MS
Roxarsone degradation
Patwittwar 200R3C2 0.5mL_D-21_01_423.d: EIC 154.01460±0.002 -All MS
Roxarsone
Patwittwar PC-1 1.5mL_D-22_01_424.d: EIC 154.01460±0.002 -All MS
4-hydroxy-2-nitrophenol
Intens.
x10^7
x10^10
2
4
6
8
10
12
Time [min]

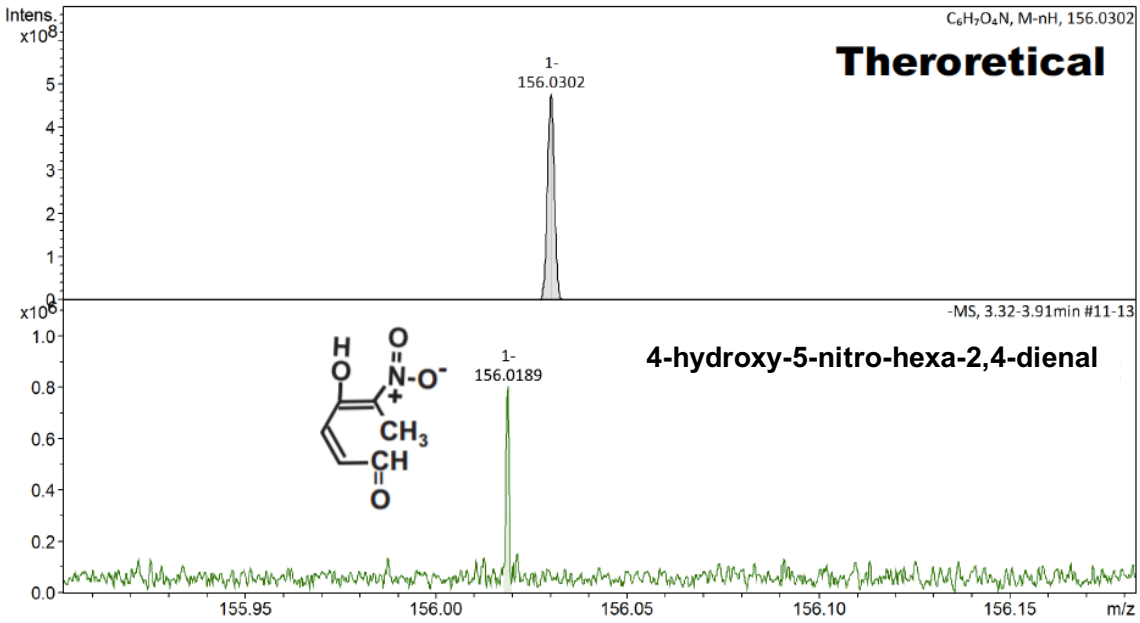


Figure 19a. HR-ESI-MS spectrum with extracted ion chromatograms (XICs) for the deprotonated pseudo molecular ion of C6H6O5NAs (m/z 245.9389) [A], C6H7O4N (m/z 156.0302), C6H5O4N (m/z 154.0146) [B] compared to theoretical mass.

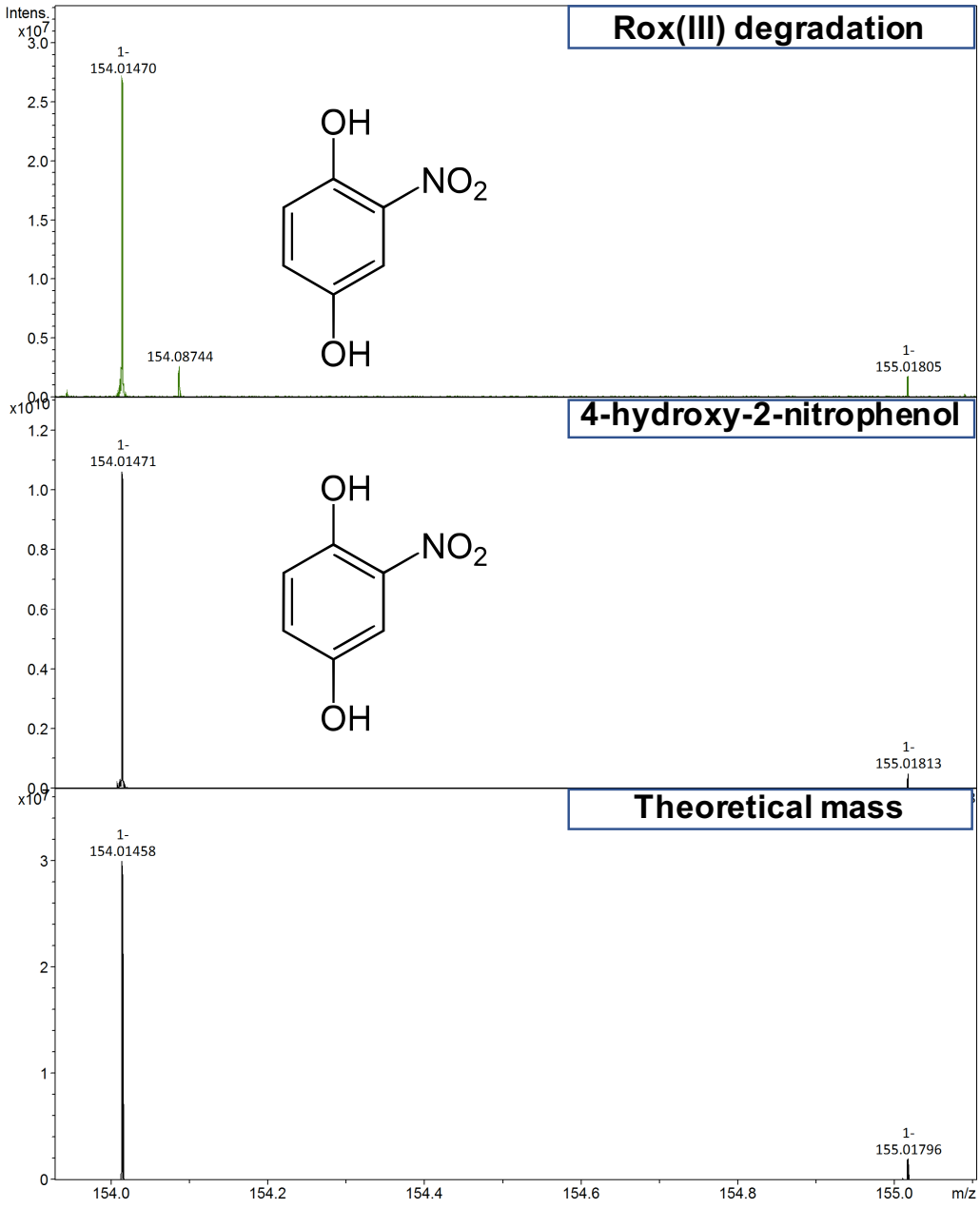


Figure 19b. Roxarsone degradation product compared with the 4-hydroxy-2-nitrophenol (authentic compound) and theoretical mass for further confirmation [C].

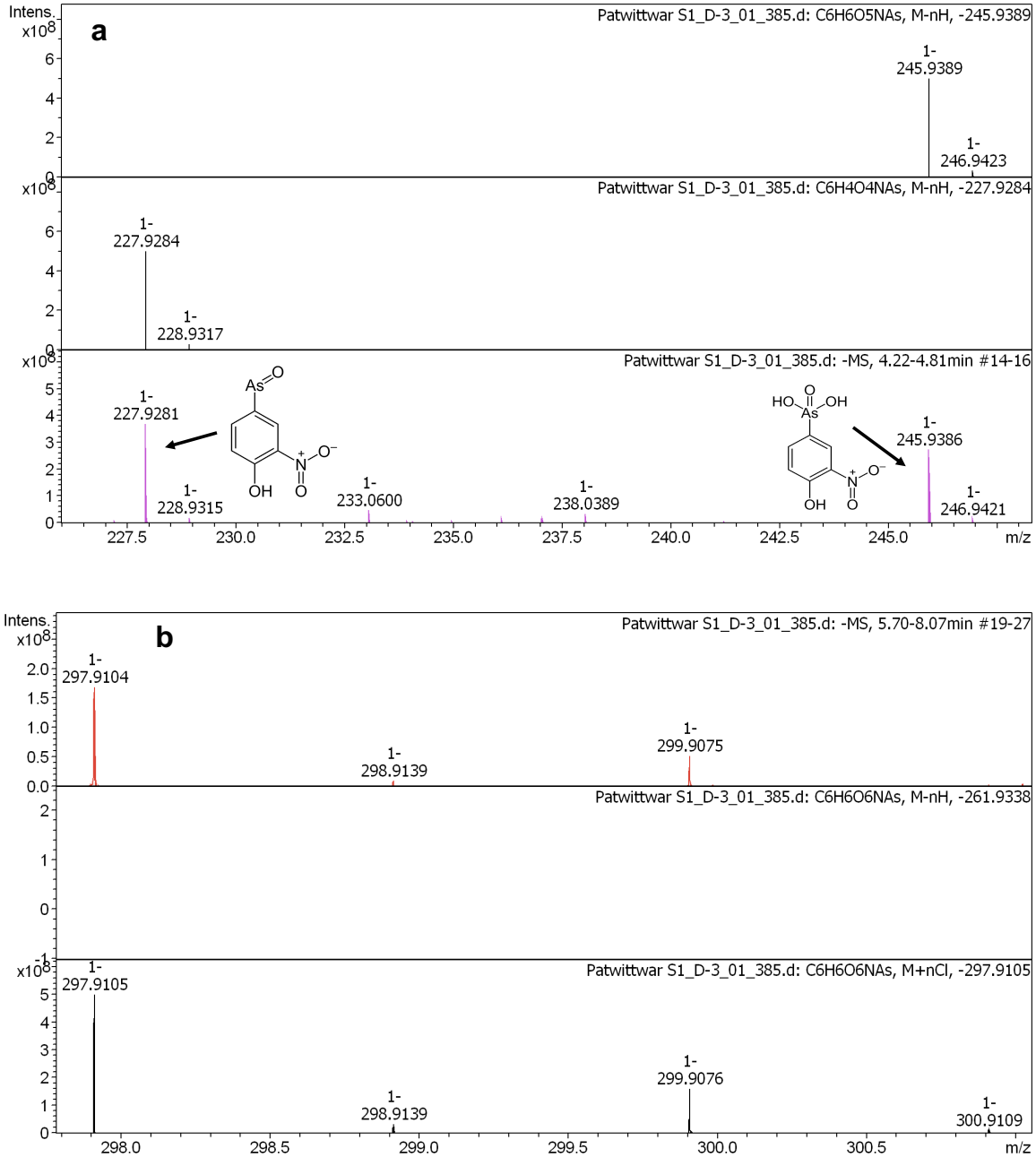


Figure 20. HR-ESI-MS spectrum showing Rox(III) dehydrogenated fragmentation [a] and Rox(V) chlorine adduct [b] with their respective isotopes.

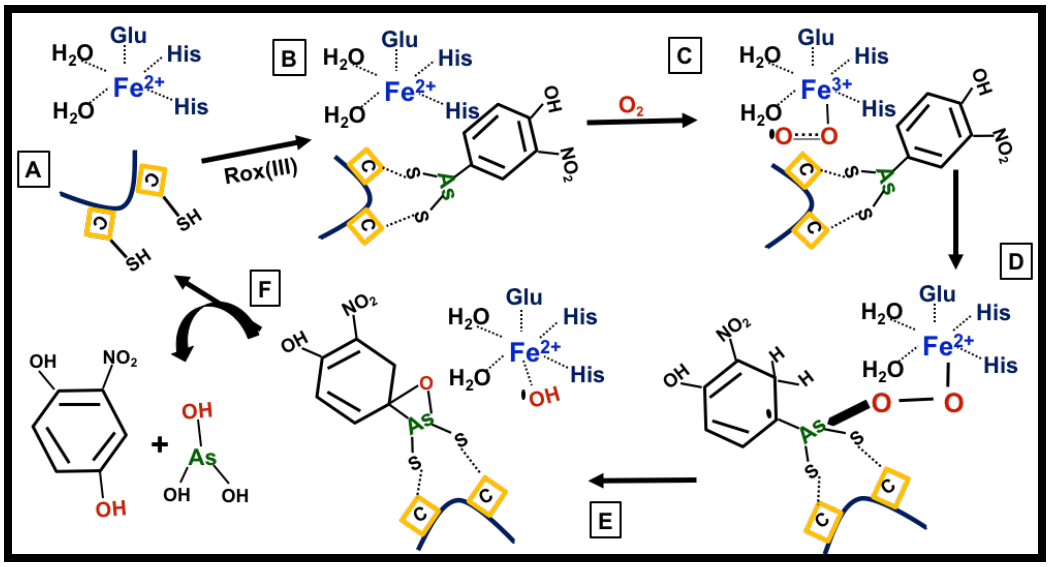


Figure 21. Modified proposed mechanism and products for ArsI catalyzed degradation of MAs(III) and Rox(III).

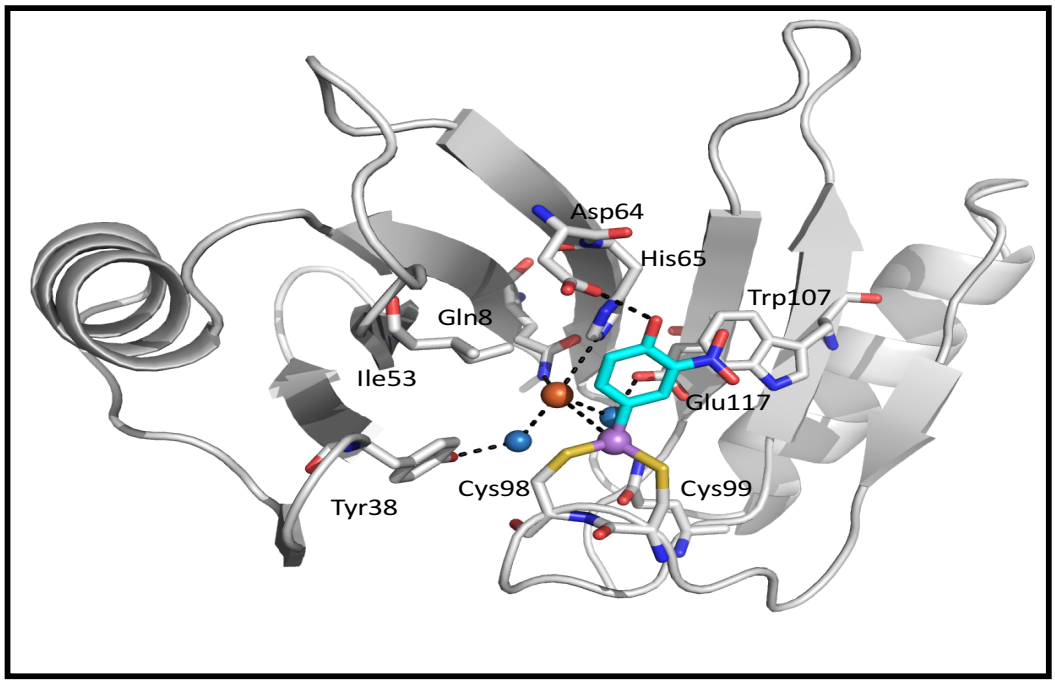


Figure 22. TcArsI bound Rox(III) crystal structure. Rox(III) interaction with TcArsI. Fe is in brown sphere and water molecules are in blue sphere. The As atom of Rox(III) is covalently attached with thiol group of Cys98, Cys99 and interact directly with iron molecule. The aromatic group of Rox(III) is stabilized by aromatic ring of Trp107, His65 and the hydrophobic residue Ile53. The hydroxyl group of Rox(III) interact with Asp64 (Unpublished data).

Table 4. Calculated mass error of target product analyzed by HR-ESI-MS. In general error of ± 5 ppm is accepted.

Target product and m/z	Mass error [ppm]
Rox(III) # 246	0.08
Rox(V) # 263	0.8
4-hydroxy-2-nitrophenol # 155	0.76
4-hydroxy-2-nitrophenol [Purchased compound] # 155	0.84
4-hydroxy-5-nitro-hexa-2,4-dienal # 157	72.4

After confirmation of substrate and product in LR-ESI-MS with optimum separation, I did an ESI-high resolution MS [ESI-HR-MS] spectrum to identify the monoisotopic mass of carbon product after Rox(III) enzymatic degradation. 4-hydroxy-2-nitrophenol shows tailing because of high concentration of the compound (Fig 19). High resolution (HR) ensures mass accuracy over sample range. The ions (in negative ion mode) at m/z 246, 262, and 154 represents Rox(III), Rox(V) and 4-hydroxy-2-nitrophenol respectively. The small peak at m/z 247, 263 and 155 correspond to isotopes respectively. The presence of a substantial product peak at m/z 154 indicates that the carbon product of the ArsI reaction is 4-hydroxy-2-nirophenol. The lack of a significant peak at m/z 156 suggests that 4-hydroxy-5-nitro-hexa-2,4-dienal was not produced (Fig 19). Table 4 provides mass error information. Generally, an error of ± 5 ppm is accepted. All

target products are within this range except 4-hydroxy-5-nitro-hexa-2,4-dienal which is 72 ppm.

In the same spectrum of Rox(III), the peak at m/z 228 could formally correspond to dehydrated Rox(III) ions. The difference between 246 and 228, which is 18, is characteristic of the loss of a neutral mass of 18 by the molecular ion and is typical of a water molecule. In the same spectrum, the peak at 298 could formally relates to chlorinated Rox(V) ions. The difference between 263 and 298, is 35 and is typical addition of a chlorine ion. The presence of dehydrated Rox(III) and chlorinated Rox(V) do not necessarily implies enzymatically induced loss of water or addition of chlorine group to Rox(V) but rather is an ionization fragment of roxarsone formed in the HR-MS source (Fig 20). In brief, the proposed mechanism of ArsI dioxygenase (Fig. 21), C-As bond cleavage of the substrate methyl arsenite [MAs(III)] [upper] and trivalent roxarsone [Rox(III)] [lower] shown with Fe(II) in the catalytic center ligated with 2-his-1-carboxylate triad and three water molecules. After catalysis of this substrate by ArsI dioxygenase, the products are inorganic arsenite [As(III)] and predicted carbon compounds as formaldehyde for MAs(III). The product of Rox(III) cleavage was uncertain. If the ring is cleaved, it may be 4-hydroxy-5-nitro-hexa-2,4-dienal or a similar open chain compound. If the ring is not cleaved, it may be 4-hydroxy-2-nitrophenol. ESI-MS data solved this uncertainty and provided the information about Rox(III) degradation product and suggests the ring is not cleaved and it is 4-hydroxy-2-nitrophenol. It also suggests, ArsI catalyzes C-As

bond cleavage by incorporating an atom of molecular oxygen into each carbon and arsenic of the target bond.

5.5 Discussion

All known type I extradiol dioxygenases break a C-C bond. ArsI is a new enzyme in biology. It also belongs to the family of type I extradiol dioxygenases. ArsI is different from other type-I extradiol dioxygenase. Instead of C-C bond cleavage, ArsI cleaves a C-As bond and apparently does not cleave the phenyl ring. It is involved in the second step of two-step organoarsenicals degradation pathway (Yoshinaga and Rosen 2014).

It is necessary to determine ArsI catalysis mechanism. Yoshinaga and Rosen demonstrated that ArsI could degrade different trivalent organoarsenicals including methyl arsenite, roxarsone, *p*-arsanillic acid and nitarsone. The inorganic arsenic presence after ArsI catalysis in determined by HPLC-ICP-MS. It suggests ArsI breaks C-As bond and releases inorganic arsenic. Additionally, an ITC study shows ArsI has 100-fold lower affinity for inorganic arsenite compare to methyl arsenite and 1000-fold less affinity compare to aromatic arsenicals. The lower affinity for As(III) indicates that the product is easily released from the enzyme after the reaction. However, the identity of the carbon product after ArsI catalysis was unknown. Yoshinaga and Rosen had predicted the product of roxarsone degradation is 4-hydroxy-5-nitro-hexa-2,4-dienal or a similar open chain compound if the ring is opened. If the ring is not cleaved, it may be 4-hydroxy-2-nitrophenol or a similar aromatic product. To determine the roxarsone

carbon product, I systematically carried out an MS study. I first optimized conditions to differentiate between roxarsone and the product by LR-ESI-MS. For identification of the monoisotopic mass of roxarsone and its degradation product, I used HR-ESI-MS. I determined the monoisotopic mass to four decimal places with a mass error within ± 1 ppm. The MS data suggest that ArsI cleaves the C-As bond, followed by addition of one oxygen to inorganic arsenic product and another to C-atom, resulting in the formation of 4-hydroxy-2-nitrophenol or a similar aromatic compound.

Our recently solved TcArsI structure with Rox(III) bound suggests substrate binding to ArsI is different from other dioxygenases. In case of AkbC and other dioxygenase, in the active C-domain an iron is coordinated by a 2-His-1-carboxylate facial triad motif. It occupies one face of the sphere, and the other face is occupied by water molecule. Substrate binding causes the release of a water molecule, followed by coordination of an oxygen molecule. Molecular oxygen activation takes place by an electron transferred from the substrate via the iron. From the structure and function studies shows, additional conserved residues, including a tyrosine and histidine, play a role in substrate deprotonation and in oxygen activation (Bugg 2003; Lipscomb 2008; Cho, Kim et al. 2010). In AkbC, an additional conserved Tyr253 is directly involved in iron binding and in deprotonation of the hydroxyl group of 2,3-dihydroxybiphenyl. The newly-solved TcArsI structure with bound Rox(III) shows that substrate binding is different compared to other type I extradiol dioxygenases like AkbC and DHBD. Unlike in

AkbC, in ArsI the substrate Rox(III) covalently binds to the vicinal cysteine pair Cys98-Cys99 present in the flexible loop, and interacts with the iron at 3.5 Å. The aromatic group of Rox(III) is stabilized by interactions with the aromatic ring of Trp107, His65, and the hydrophilic residue Ile53. Additionally, the hydroxyl group of Rox(III) forms an H-bond with Asp64 (Fig. 22). It may be possible that, like AkbC, Rox(III) binding triggers oxygen activation, but this is not clear yet. Further studies are necessary to determine the role of other conserved amino acids in substrate deprotonation and oxygen activation. Overall, further studies will help to understand the underlying mechanism of ArsI catalysis.

5.6 References

Bednar, A. J., J. R. Garbarino, et al. (2003). "Photodegradation of roxarsone in poultry litter leachates." Science of the Total Environment **302**(1–3): 237-245.

Bednar, A. J., J. R. Garbarino, et al. (2002). "Presence of Organoarsenicals Used in Cotton Production in Agricultural Water and Soil of the Southern United States." Journal of Agricultural and Food Chemistry **50**(25): 7340-7344.

Bugg, T. D. (2003). "Dioxygenase enzymes: catalytic mechanisms and chemical models." Tetrahedron **59**(36): 7075-7101.

Chapman, H. and Z. Johnson (2002). "Use of antibiotics and roxarsone in broiler chickens in the USA: analysis for the years 1995 to 2000." Poultry Science **81**(3): 356-364.

Cho, H. J., K. Kim, et al. (2010). "Substrate binding mechanism of a type I extradiol dioxygenase." Journal of Biological Chemistry **285**(45): 34643-34652.

Cortinas, I., J. A. Field, et al. (2006). "Anaerobic biotransformation of roxarsone and related N-substituted phenylarsonic acids." Environmental Science & Technology **40**(9): 2951-2957.

Feng, M., J. E. Schrlau, et al. (2005). "Arsenic transport and transformation associated with MSMA application on a golf course green." Journal of Agricultural and Food Chemistry **53**(9): 3556-3562.

Frensemeier, L. M., L. Büter, et al. (2017). "Investigation of the oxidative transformation of roxarsone by electrochemistry coupled to hydrophilic interaction liquid chromatography/mass spectrometry." Journal of Analytical Atomic Spectrometry **32**(1): 153-161.

Gandhi, M. (2008) "Arsenic poison in chicken feed ".

Huang, L., L. Yao, et al. (2014). "Roxarsone and its metabolites in chicken manure significantly enhance the uptake of As species by vegetables." Chemosphere **100**: 57-62.

Lipscomb, J. D. (2008). "Mechanism of extradiol aromatic ring-cleaving dioxygenases." Current Opinion in Structural Biology **18**(6): 644-649.

Moody, J. and R. Williams (1964). "The metabolism of 4-hydroxy-3-nitrophenylarsonic acid in hens." Food and Cosmetics Toxicology **2**: 707-715.

Moore, P., T. Daniel, et al. (1995). "Poultry manure management: Environmentally sound options." Journal of Soil and Water Conservation **50**(3): 321-327.

Morrison, J. L. (1969). "Distribution of arsenic from poultry litter in broiler chickens, soil, and crops." Journal of Agricultural and Food Chemistry **17**: 1288-1293.

Nachman, K. E., P. A. Baron, et al. (2013). "Roxarsone, inorganic arsenic, and other arsenic species in chicken: a US-based market basket sample." Environmental Health Perspectives **121**(7): 818.

Peng, H., B. Hu, et al. (2014). "Liquid chromatography combined with atomic and molecular mass spectrometry for speciation of arsenic in chicken liver." Journal of Chromatography A **1370**: 40-49.

Xu, T., P. V. Kamat, et al. (2007). "Hydroxyl radical mediated degradation of phenylarsonic acid." The Journal of Physical Chemistry A **111**(32): 7819-7824.

Yoshinaga, M. and B. P. Rosen (2014). "AC· As lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters." Proceedings of the National Academy of Sciences **111**(21): 7701-7706.

Zhang, W., F. Xu, et al. (2015). "Comparative cytotoxicity and accumulation of Roxarsone and its photodegradates in freshwater Protozoan Tetrahymenathermophila." Journal of Hazardous Materials **286**: 171-178.

CHAPTER 6. DRAFT GENOME SEQUENCE OF *BURKHOLDERIA* SP. MR1, A METHYLARSENATE-REDUCING BACTERIAL ISOLATE FROM FLORIDA GOLF COURSE SOIL

6.1 Introduction

Reductases play a key role in cellular metabolism (Nelson 2008).. For example, reductases from anaerobic bacteria play critical roles in aromatic metabolism (Boll and Fuchs 1995). Reductases are clinically important as well. For example, male pseudohermaphroditism, a condition in which complete male sexual development is affected, results from a deficiency of 5-alpha reductase, and cardiovascular morbidity is reduced by HMG Co-A reductase inhibitors (Laufs, Custodis et al. 2006).

In the context of the arsenic biogeocycle, reductases also play important roles (Ahmann, Roberts et al. 1994). Bacteria use arsenate reductase (ArrAB) to generate energy through oxidative phosphorylation (Silver and Phung 2005). Methylarsenicals such as MAs(V) and dimethylarsinic acid [DMAs(V)] have been used as pesticides in agriculture such as cotton and other crops. In addition, they have been used on golf course, lawns and storage yards to maintain turf. These pentavalent methylated arsenicals are readily soluble and considered less toxic than inorganic arsenic. However, these less toxic methylarsenicals are

biotransformed into more toxic inorganic arsenicals, which results in increased environmental pollution (Yoshinaga, Cai et al. 2011).

In the US, in a few states such as California, Texas, Montana, Oregon, Nevada and Minnesota arsenic contamination in some water supplies is 50 µg/ml, five-fold more than the EPA's maximum containment level. Human exposure to inorganic arsenical forms is implicated in various diseases such as diabetes, cardiovascular conditions, and cancers (Tchounwou, Patlolla et al. 2003). The long-term effects of inorganic arsenic exposure are hyperkeratosis on the palms and soles of the feet, skin cancer, bladder cancer, lung cancer and Blackfoot disease, the result of peripheral vascular necrosis. The International Agency for Research on Cancer (IARC) has affirmed inorganic arsenic as a carcinogenic to humans. The EPA banned the use of chromate copper arsenate (CCA)-treated wood for household and playgrounds use.

CCA-treatment protects wood from rotting and increase its total life period. However, the risks outweigh their benefits to humans. Because of extreme toxicity and carcinogenicity of arsenicals in the environment, proper disposal of CCA treated wood is necessary (EPA).

Ubiquitous toxic arsenic presence has resulted in the evolution of detoxifying systems in nearly every organism. Inorganic arsenic methylation is a common mechanism for detoxification with catalysis by As(III) S-adenosylmethionine methyltransferase enzymes (ArsM in microbes and AS3MT in animals) (Qin, Lehr et al. 2009). Methylated arsenic demethylating microbes have been isolated from soil (Yoshinaga, Cai et al. 2011) and lake water. In only

a few cases, specific organisms were associated with demethylation such as Montrachet wine yeast (Crecelius 1977), *Candida humicola* (Lehr, Polishchuk et al. 2003), and *Mycobacterium neoaurum* (Lehr, Polishchuk et al. 2003). DMAs(V) and MAs(V) are demethylated to less methylated compounds in sediment, and DMAs(V) was demonstrated to be demethylated to MAs(V) in sludge (Sierra-Alvarez, Yenal et al. 2006). However, there are only a few reports on reduction of methylarsenicals in the environment (Yoshinaga, Cai et al. 2011). ArsC arsenate reductase, a member of arsenic resistance (*ars*) operon genes, reduces neither MAs(V) nor DMAs(V) (Yoshinaga, Cai et al. 2011), implying that there are specific reductases for methylarsenicals.

To elucidate the environmental organoarsenical biocycle, we isolated a soil organism, *Burkholderia* sp. MR1, which reduces relatively nontoxic pentavalent methylarsenate to the more toxic trivalent methylarsenite, with the goal of identifying the gene for the reductase. Here, we report the draft genome sequence of *Burkholderia* sp. MR1.

6.2 Results

6.2.1 Genome Announcement

Arsenic is the most prevalent environmental toxin by the U.S. Environmental Protection Agency (EPA). Prolonged exposure to arsenic in drinking water results in diseases such as skin, bladder, and prostate cancer (Abernathy, Thomas et al. 2003; Tchounwou, Centeno et al. 2004). Pentavalent organoarsenical herbicides such as monosodium methylarsenate (MSMA) [MAs(V)] undergo environmental reduction to the more toxic trivalent species and

eventual C-As bond cleavage to produce arsenite [As(III)] (Yoshinaga, Cai et al. 2011), which contaminates our drinking water supplies (Feng, Schrlau et al. 2005). The poultry growth promoter roxarsone (4-hydroxy-3-nitrophenylarsonic acid) has been proposed to undergo a similar pathway of reduction and C-As bond cleavage (Yoshinaga and Rosen 2014). The arsI gene of soil organisms encodes a C-As bond lyase that catalyzes the second reaction, but the pathway for reduction of MAs(V) to MAs(III) is unknown. The objective of this study was to identify the gene(s) and enzyme(s) responsible for MAs(V) reduction.

MSMA is applied as an herbicide to golf courses in Florida. We isolated a MAs(V)-reducing organism from simulated golf course soil. From the 16S rRNA gene sequence, it is a Gram-negative bacterial strain closely related to *Burkholderia glathei* (Yoshinaga, Cai et al. 2011). The environmental isolate has been named *Burkholderia* sp. MR1. Genome sequencing was performed using an Illumina HiSeq platform, with quality-based trimming, as described previously (Woo, Utturkar et al. 2014). After trimming, 6,727,114 paired-end reads remained, with an average read length of 90 bp, comprising genome coverage of 100×. After evaluation of several approaches (Utturkar, Klingeman et al. 2014), optimal assembly was obtained through SPAdes software (version 3.1.1). The assembly consisted of 58 large contigs (≥500 bp), with a total genome size of 6.01 Mb, an N50 contig size of 244 kb, and the largest contig of 432 kb. Gene prediction and annotation were performed at Oak Ridge National Laboratory, as described previously (Brown, Utturkar et al. 2012). The draft genome sequence has 5,554 candidate protein-coding genes and a G+C content of 63.2%.

The genomic sequence of *Burkholderia* sp. MR1 has a single ars operon comprising the genes arsH [encoding an NADPH-dependent flavin mononucleotide (FMN) MAs(III) oxidase] (Chen, Bhattacharjee et al. 2015), acr3 [encoding an As(III) efflux permease] (Yang, Fu et al. 2012), and arsR [encoding an As(III)-responsive repressor protein] (Xu and Rosen 1999). The operon is on the complementary strand, so the putative operon structure is arsR-acr3-arsH. Note that ArsH detoxifies MAs(III) (Chen, Bhattacharjee et al. 2015) and so may protect this organism from the toxicity of the product of the reductase. None of these gene products catalyze MAs(V) reduction, so the gene or genes for reductase activity must be elsewhere in the chromosome. In the genome, there are 195 genes annotated as reductases that are candidates for the MAs(V) reductase. However, further biochemical or genetic analysis is required to identify the responsible gene(s).

6.2.2 Nucleotide Sequence Accession Number

The draft genome sequence of strain AE038-8 has been deposited at DDBJ/EMBL/GenBank (accession number JWHM00000000). The version described in this paper is the first version.

6.3 Discussion

To study the molecular mechanism of MSMA reductase, there is need to identify a gene responsible for MSMA reduction. Here we sequenced soil isolate *Burkholderia* sp. MR1, an MSMA reducing bacteria. Our objective is to avail genome sequence of *Burkholderia* to identify the gene responsible for the methylarsenical reduction. Our lab is developing a screening method to identify a

protein from *Burkholderia* capable of reducing MSMA, a reverse engineering method. Protein fractions which reduces MSMA can be sequenced and from that sequence, a DNA sequence can be obtained. Once a gene sequence identified, the genome sequence of *Burkholderia* reported here will provide the information necessary to identify the gene responsible for MSMA reduction, its location, whether it under the regulation of any *ars* regulatory protein. Thus, the genome sequence of *Burkholderia* would contribute to identifying an MSMA reductase and once it is identified, the molecular mechanism of MSMA reduction can be studied in the future.

6.4 References

Abdul Ajees, A., J. Yang, et al. (2011). "The ArsD As(III) metallochaperone." Biometals **24**(3): 391-399.

Abernathy, C. O., D. J. Thomas, et al. (2003). "Health effects and risk assessment of arsenic." The Journal of Nutrition **133**(5 Suppl 1): 1536S-1538S.

Ahmann, D., A. L. Roberts, et al. (1994). "Microbe grows by reducing arsenic." Nature **371**(6500): 750-750.

Bednar, A. J., J. R. Garbarino, et al. (2003). "Photodegradation of roxarsone in poultry litter leachates." Science of the Total Environment **302**(1–3): 237-245.

Bednar, A. J., J. R. Garbarino, et al. (2002). "Presence of Organoarsenicals Used in Cotton Production in Agricultural Water and Soil of the Southern United States." Journal of Agricultural and Food Chemistry **50**(25): 7340-7344.

Boll, M. and G. Fuchs (1995). "Benzoyl-Coenzyme A Reductase (Dearomatizing), a Key Enzyme of Anaerobic Aromatic Metabolism." European Journal of Biochemistry **234**(3): 921-933.

Brown, S. D., S. M. Utturkar, et al. (2012). "Twenty-one genome sequences from Pseudomonas species and 19 genome sequences from diverse bacteria

isolated from the rhizosphere and endosphere of Populus deltoides." Journal of Bacteriology **194**(21): 5991-5993.

Chapman, H. and Z. Johnson (2002). "Use of antibiotics and roxarsone in broiler chickens in the USA: analysis for the years 1995 to 2000." Poultry Science **81**(3): 356-364.

Chen, J., H. Bhattacharjee, et al. (2015). "ArsH is an organoarsenical oxidase that confers resistance to trivalent forms of the herbicide monosodium methylarsenate and the poultry growth promoter roxarsone." Molecular Microbiology **96**(5): 1042-1052.

Chen, J., M. Madegowda, et al. (2015). "ArsP: a methylarsenite efflux permease." Molecular Microbiology **98**(4): 625-635.

Cortinas, I., J. A. Field, et al. (2006). "Anaerobic biotransformation of roxarsone and related N-substituted phenylarsonic acids." Environmental science & technology **40**(9): 2951-2957.

Crecelius, E. A. (1977). "Arsenite and arsenate levels in wine." Bulletin of Environmental Contamination and Toxicology **18**(2): 227-230.

Dheeman, D. S., C. Packianathan, et al. (2014). "Pathway of human AS3MT arsenic methylation." Chemical Research in Toxicology **27**(11): 1979-1989.

Feng, M., J. E. Schrlau, et al. (2005). "Arsenic transport and transformation associated with MSMA application on a golf course green." Journal of Agricultural and Food Chemistry **53**(9): 3556-3562.

Frensemeier, L. M., L. Büter, et al. (2017). "Investigation of the oxidative transformation of roxarsone by electrochemistry coupled to hydrophilic interaction liquid chromatography/mass spectrometry." Journal of Analytical Atomic Spectrometry **32**(1): 153-161.

Gandhi, M. (2008) "Arsenic poison in chicken feed ".

Green M. R, S. J., and Sambrrok J. (2012). Molecular cloning–A laboratory manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.

Groce, S. L., M. A. Miller-Rodeberg, et al. (2004). "Single-turnover kinetics of homoprotocatechuate 2, 3-dioxygenase." Biochemistry **43**(48): 15141-15153.

Huang, L., L. Yao, et al. (2014). "Roxarsone and its metabolites in chicken manure significantly enhance the uptake of As species by vegetables." Chemosphere **100**: 57-62.

Laufs, U., F. Custodis, et al. (2006). "HMG-CoA reductase inhibitors in chronic heart failure." Drugs **66**(2): 145-154.

Lehr, C., E. Polishchuk, et al. (2003). "Demethylation of methylarsenic species by Mycobacterium neoaurum." Applied Organometallic Chemistry **17**(11): 831-834.

Leitgeb, S., G. D. Straganz, et al. (2009). "Biochemical characterization and mutational analysis of the mononuclear non-haem Fe^{2+} site in Dke1, a cupin-type dioxygenase from Acinetobacter johnsonii." Biochemical Journal **418**(2): 403-411.

Lin, Y. F., J. Yang, et al. (2007). "ArsD: an As(III) metallochaperone for the ArsAB As(III)-translocating ATPase." Journal of Bioenergetics and Biomembranes **39**(5-6): 453-458.

Maki, T., W. Hirota, et al. (2009). "Seasonal dynamics of biodegradation activities for dimethylarsinic acid (DMA) in Lake Kahokugata." Chemosphere **77**(1): 36-42.

Marapakala, K., J. Qin, et al. (2012). "Identification of catalytic residues in the As(III) S-adenosylmethionine methyltransferase." Biochemistry **51**(5): 944-951.

Moody, J. and R. Williams (1964). "The metabolism of 4-hydroxy-3-nitrophenylarsonic acid in hens." Food and Cosmetics Toxicology **2**: 707-715.

Moore, P., T. Daniel, et al. (1995). "Poultry manure management: Environmentally sound options." Journal of Soil and Water Conservation **50**(3): 321-327.

Morrison, J. L. (1969). "Distribution of arsenic from poultry litter in broiler chickens, soil, and crops." Journal of Agricultural and Food Chemistry **17**: 1288-1293.

Nachman, K. E., P. A. Baron, et al. (2013). "Roxarsone, inorganic arsenic, and other arsenic species in chicken: a US-based market basket sample." Environmental Health Perspectives (Online) **121**(7): 818.

Nelson, C. (2008). Lehninger Principles of Biochemistry, W.H. Freeman.

Nogales, J., Á. Canales, et al. (2011). "Unravelling the gallic acid degradation pathway in bacteria: the gal cluster from Pseudomonas putida." Molecular Microbiology **79**(2): 359-374.

Nozaki, M. (1979). "Oxygenases and dioxygenases." Biochemistry: 145-186.

Peng, H., B. Hu, et al. (2014). "Liquid chromatography combined with atomic and molecular mass spectrometry for speciation of arsenic in chicken liver." Journal of Chromatography A **1370**: 40-49.

Qin, J., C. R. Lehr, et al. (2009). "Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga." Proceedings of the National Academy of Sciences **106**(13): 5213-5217.

Qin, J., B. P. Rosen, et al. (2006). "Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase." Proceedings of the National Academy of Sciences **103**(7): 2075-2080.

Sierra-Alvarez, R., U. Yenal, et al. (2006). "Anaerobic Biotransformation of Organoarsenical Pesticides Monomethylarsonic Acid and Dimethylarsinic Acid." Journal of Agricultural and Food Chemistry **54**(11): 3959-3966.

Silver, S. and L. T. Phung (2005). "Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic." Applied and Environmental Microbiology **71**(2): 599-608.

Tchounwou, P. B., J. A. Centeno, et al. (2004). "Arsenic toxicity, mutagenesis, and carcinogenesis--a health risk assessment and management approach." Molecular and Cellular Biochemistry **255**(1-2): 47-55.

Tchounwou, P. B., A. K. Patlolla, et al. (2003). "Carcinogenic and systemic health effects associated with arsenic exposure--a critical review." Toxicologic Pathology **31**(6): 575-588.

Utturkar, S. M., D. M. Klingeman, et al. (2014). "Evaluation and validation of de novo and hybrid assembly techniques to derive high-quality genome sequences." Bioinformatics **30**(19): 2709-2716.

Woo, H. L., S. Utturkar, et al. (2014). "Draft genome sequence of the lignin-degrading Burkholderia sp. strain LIG30, isolated from wet tropical forest soil." Genome Announcements **2**(3): e00637-00614.

Xu, C. and B. P. Rosen (1997). "Dimerization is essential for DNA binding and repression by the ArsR metalloregulatory protein of Escherichia coli." Journal of Biological Chemistry **272**(25): 15734-15738.

Xu, C. and B. P. Rosen (1999). Metalloregulation of soft metal resistance pumps. Metals and Genetics. B. Sarkar. New York, Plenum Press**:** 5-19.

Xu, T., P. V. Kamat, et al. (2007). "Hydroxyl radical mediated degradation of phenylarsonic acid." The Journal of Physical Chemistry A **111**(32): 7819-7824.

Yang, H. C., H. L. Fu, et al. (2012). "Pathways of arsenic uptake and efflux." Current Topics in Membranes **69**: 325-358.

Yang, J., S. Rawat, et al. (2010). "Arsenic binding and transfer by the ArsD As(III) metallochaperone." Biochemistry **49**(17): 3658-3666.

Yoshinaga, M., Y. Cai, et al. (2011). "Demethylation of methylarsonic acid by a microbial community." Environmental microbiology **13**(5): 1205-1215.

Yoshinaga, M. and B. P. Rosen (2014). "A CAs lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters." Proceedings of the National Academy of Sciences **111**(21): 7701-7706.

Zhang, W., F. Xu, et al. (2015). "Comparative cytotoxicity and accumulation of Roxarsone and its photodegradates in freshwater Protozoan Tetrahymenathermophila." Journal of Hazardous Materials **286**: 171-178.

Zhou, T. and B. P. Rosen (1997). "Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase." Journal of Biological Chemistry **272**(32): 19731-19737.

Zhou, T., J. Shen, et al. (2002). "Unisite and multisite catalysis in the ArsA ATPase." Journal of Biological Chemistry **8**(6): 23815-23820.

CHAPTER 7. CONCLUSIONS, SIGNIFICANCE AND FUTURE RESEARCH PROSPECTIVE

7.1 Conclusion

The primary theme of my research has been to elucidate the catalytic mechanisms of ArsI, ArsI affinity with different substrates, role of predicted catalytic residues in MAs(III) demethylation and determination of carbon product after C-As bond cleavage by ArsI. In my dissertation research, I investigated two prominent ArsI catalytic mechanisms – structure function analysis of ArsI and determination of carbon product after ArsI catalysis. ArsI cleaves C-As bond and adds one oxygen to arsenic species and other to carbon product resulting in inorganic arsenite and 4-hydroxy-2-nitrophenol after Rox(III) degradation. However, there are contradicting views about the role of arsenic demethylation in detoxification. Several of the key events of arsenite demethylation are not clear, and understanding them will help elucidate the role of this enzyme.

Arsenic is the omnipresent environmental toxin. In addition to natural geothermal sources, various human activities such as mining, smelting, and agriculture are routes of arsenic introduction into the environment. Inorganic arsenic compounds are a defined carcinogen that leads to various types of cancers (Falk H et al. 1981), The U.S. Environmental Protection Agency (EPA) and International Agency for Research on Cancer (IARC) classified inorganic arsenic as group A (a human carcinogen) and group 1 (carcinogenic to humans), respectively. Exposure to arsenic also causes many other health problems including cardiovascular disease (States J et al. 2009), diabetes (Drobná Z et al.

2013), hearing loss (Bencko and Symon 1977), peripheral neuropathy (Mukherjee S et al. 2003) and neurological development (Tyler and Allan 2004). Due to its environmental ubiquity and severe health effects, EPA and Agency for Toxic Substances and Disease Registry (ATSDR) have considered arsenic to be the most prevalent environmental toxin and the most significant potential threat to human health and ranked it at the top of the US Priority List of Hazardous Substances (https://www.atsdr.cdc.gov/SPL/index.html).

The persistent environmental toxin arsenic has provided selective pressure for the evolution of arsenic detoxification mechanisms which distribute in nearly every organism at present. Especially, microbes, which thrive in almost every terrestrial environment, have evolved a variety of unique arsenic detoxifying genes (Meng Y et al. 2004; Gladysheva T et al. 1994). Arsenic toxicity alters depending on the species and one of the strategies microbes have adopted to cope with arsenic is transformation, which forms a significant part of the arsenic biogeochemical cycle. One of the major cycle comprising the arsenic biogeochemical cycle is methylation cycle, wherein some microbes convert inorganic arsenic, the dominant form found in the environment, into mono-, di- and tri-methylated arsenic derivatives whereas some others demethylate them back to inorganic forms. Methylarsenicals are also anthropogenically introduced into the environment, the representative example is methylarsonic acid [MAs(V)] that is currently registered for herbicidal use in the United States [https://www.epa.gov/ingredients-used-pesticide-products/monosodium-

methanearsonate-msma-organic-arsenical]. Moreover, manmade aromatic arsenicals such as roxarsone [(4-hydroxy-3-nitrophenyl) arsonic acid, Rox(V)], p-arsinilic acid [(4-aminophenyl) arsnoic acid, p-ASA(V)] or nitarsone [(4-nitrophenyl) arsonic acid, Nit(V)] have been widely used in animal husbandry as feed antimicrobial growth enhancers (Nachman K et al. 2013). The US Food and Drug Administration (FDA) withdrew the approval of all applications for use of Rox(V) and p-ASA(V) in 2011 and for Nit(V) in 2015 [https://www.fda.gov/animalveterinary/newsevents/cvmupdates/ucm440668.htm], they are still in production and use in many other countries. These aromatic arsenicals are largely excreted unchanged from animal bodies and introduced into the environment when livestock manure is applied to fields as fertilizer. Anthropogenic organoarsenical sources are incorporated into the methylation cycle, where they largely degrade into more toxic inorganic arsenicals and contaminate crops and drinking water supplies (Yoshinaga et al. 2011; Yoshinaga, M. and Rosen B 2014).

Pentavalent organoarsenicals are benign and much less toxic compared to an inorganic form, which has facilitated their agricultural use in the last few decades. This also implies that methylation of inorganic arsenic is a detoxification mechanism. In fact, the responsible gene arsM is widely found in bacterial arsenic resistance (ars) operons and the arsenic hypersensitive *Escherichia coli* strain AW3110 recovers its sensitivity to inorganic arsenite [As(III)] by heterologous expression of ArsM (QinJ et al. 2006; 2009). However, it has been

indicated that the real products of ArsM are trivalent forms such as methylarsonous acid [MAs(III)] and dimethylarsinous acid [DMAs(III)] that are even more toxic than inorganic species, although they get oxidized into benign pentavalent forms soon after the reaction under oxic condition (Marapakala K 2012; Dheeman D et al. 2014). This has raised an opposing theory that arsenic methylation by ArsM could be considered as a bioactivation process, especially in suboxic and anoxic environments.

As contrasted to methylation, despite many reports on microorganisms capable of demethylation of methylarsenicals, the molecular mechanisms had not been known. We previously identified a novel two-step MAs(V) degradation pathway catalyzed by a bacterial community, where some bacteria reduce MAs(V) to MAs(III) and some others demethylate MAs(III) to inorganic As(III), from Florida golf course soil where herbicide MAs(V) was repeated applied (Yoshinaga M et al. 2011). We identified the gene, termed arsI, responsible for the second step MAs(III) demethylation from the environmental MAs(III)-demethylating bacterium *Bacillus* sp. MD1. The gene product, ArsI, is a non-heme iron-dependent type-I extradiol dioxygenase with C-As (carbon-arsenic) lyase activity, which confers MAs(III) detoxification. In addition, ArsI cleaves the C-As bond in a wide range of trivalent aromatic arsenicals, including Rox(III), Nit(III) and p-ASA(III), and dearylates them into As(III) (Yoshinaga M. and Rosen B 2014). As of April 25, 2017, there are nearly 1305 putative ArsI orthologs identified exclusively in 963 aerobic bacterial species by a Basic Local Alignment

Search Tool Link to Protein and Structures (BLAST Link or BLink) search, suggesting that arsI genes are widely distributed in the oxic environment. It is interesting to note that the unicellular cyanobacterium *Nostoc* sp. PCC 7120 has both arsM and arsI genes and separately utilize each gene product for arsenic detoxification depending on arsenic concentration in the environment, which is considered to confer a broader spectrum of resistance to both As(III) and MAs(III) (Yan Y et al. 2015; 2017).

Currently, ArsI is the first and only enzyme known to be involved in organoarsenic degradation. We recently solved the crystal structure of the thermophilic ArsI ortholog from *Thermomonospora curvata* (TcArsI) in both the apo form and with metals (Nadar S et al. 2014), which identified the unique feature of ArsI proteins. Although the overall structure is highly like those of type-I extradiol ring-cleaving dioxygenases, ArsI proteins possess a distinctive flexible loop, which we propose captures the substrate and conveys it to the active site for the catalysis by a loop-gating mechanism. In this study, for further elucidation of the enzymatic mechanism of this unprecedented C-As lyase catalysis, we characterized the properties of metal and substrate bindings and demonstrated that all the residues involved in the bindings are essential for the enzyme reaction. In addition, we identified the carbon products of ArsI for aromatic arsenicals, which supports that ArsI is a dioxygenase and catalyzes C-As bond cleavage by incorporating an atom of molecular oxygen into each carbon and

arsenic of the target bond. Our results will provide a novel insight into the biogeochemical cycles not only for arsenic but also carbon.

In the third chapter, I investigated the role of conserved residues necessary for ArsI catalysis. TcArsI structure solved in both the apo form and with Ni(II), Co(II), or Fe(III) shows that Gln8-His65-Glu116 (equivalent to His5-His62-Glu115 in *Bacillus*) residues are involved in the metal binding (Nadar S et al. 2016). The conserved cysteine pair (Cys98 and Cys99 in TcArsI) that are predicted to bind to MAs(III) and trivalent roxarsone [Rox (III)] was predicted to be substrate binding site. Additionally, TcArsI was reacted with βME to form a mixed disulfide adduct with the conserved cysteine pair (Nadar S et al. 2014). To demonstrate role of these metal and substrate binding site, residue specific chemical modifiers on demethylation activity of ArsI performed. Loss of demethylation activity in the presence of chemical modifiers supported role of conserved metal and substrate binding residues. The participation of the residues for Fe(II) and substrate binding in catalysis was further investigated by mutagenesis. Each of the Fe(II)-binding residues was changed to alanine, whereas the cysteine residues in the predicted substrate binding site were individually mutated to serine. All introduced single mutations in the Fe(II)-, and substrate-binding residues resulted in loss of activity *in vivo* (Fig 9A) and *in vitro* (Fig 9B), implying that all those residues are critical for the ArsI catalysis. Additionally, recently TcArsI bound with substrate Rox(III) structure solved. It shows roxarsone bound to conserved vicinal cysteine pair Cys98-Cys99

(equivalent to Cys96-Cys97) is binding site for trivalent organoarsenicals (Fig 21) (Unpublished data).

In the fourth chapter, I examined the thermodynamic properties and real-time binding properties of the ArsI. First, isothermal titration calorimetry (ITC) was performed to characterize the binding properties of wild-type $ArsI_{124}$ for Fe(II) (Fig 10) and arsenicals (Fig 11 A-C). To avoid Fe(II) oxidation, the binding property to Fe(II) was anaerobically analyzed, whereas the substrate binding was measured in aerobic condition. Stoichiometry's and dissociation constants calculated from the ITC data and summarized in table 3. These values are comparable with those for other non-heme Fe(II)-dependent dioxygenases (Leitgeb S et al. 2009). MAs(III) and PhAs(III) bound to $ArsI_{124}$ with a similar stoichiometry while K_d of PhAs(III) was lower than that of MAs(III) by three orders of magnitude suggesting aromatic arsenicals are preferable substrates. The inorganic arsenic As(III) showed lower affinity by several orders of magnitude compared to MAs(III) and PhAs(III), the ArsI substrates. The affinity difference between the substrate and product may control the steps for the substrate binding and the product release and catalytic turnover of ArsI.

Further I employed ligand-dependent changes in intrinsic tryptophan fluorescence, by which binding of substrates in various arsenic detoxification proteins has been extensively characterized (Marapakala K et al. 2012; Zhou T and Rosen B 1997; Zhou T et al. 2002; Yang J et al. 2010), and demonstrated

that these conserved metal and substrate binding residues indeed participate in bindings of Fe(II) and trivalent organoarsenical substrates (Fig 12-14).

In the fifth chapter, I determined the carbon product of Rox(III) degradation after ArsI catalysis. I employed HILIC-ESI-MS method to separate roxarsone degradation product from the substrate roxarsone and injected to mass spectrometry (MS). I used both High resolution and low resolution MS. First, the target compound separation and identification carried out at low resolution MS. High resolution MS then further used to determine mono-isotopic mass of roxarsone degradation product with mass error within accepted range (< ± 5 ppm). The presence of a substantial product peak at m/z 154 indicates that 4-hydroxy-2-nirophenol (closed ring) structure was formed after ArsI degradation of Rox(III). The lack of a significant peak at m/z 156 suggests that 4-hydroxy-5-nitro-hexa-2,4-dienal (open ring) structure was not produced after ArsI catalysis (Fig 19).

The sixth chapter describes the genome sequence of soil isolate *Burkholderia* sp. MD1 capable of reducing MSMA to MAs(III). Until now, the gene responsible to carry out this reduction has not identified. Our lab is developing a screening method to identify a protein from *Burkholderia* capable of reducing MSMA, a reverse engineering method. Once the gene sequence identified, the *Burkholderia* genome sequence reported here will provide the further information necessary to identify the gene responsible for MSMA reduction. Thus, the

genome sequence of *Burkholderia* would contribute to identifying an MSMA reductase.

7.2 Significance

Enzymes are becoming more popular as an alternative to environmental remediation processes as they offer clean and environment friendly solutions. Bioremediation refers to microbial facilitated transformation or degradation of toxic compounds into non-lethal or less toxic compounds (http://dx.doi.org/10.4061/2011/805187). It includes bioaugmentation (whole cell introduction), biocatalyst (enzyme based) or biostimulation methods (use of the suitable conducive environment by addition of nutrients or other additives to activate native microbial community). Many dioxygenases are involved in degradation of pollutants. For example, catechol dioxygenase carries out aerobic aromatic degradation of petroleum pollutants efficiently. To date, the EPA has listed 20 bioremediation agents but only one enzyme additive called Petroleum Spill eater II. It has a shelf life of 5 years and degrades 90% of aromatic and alkane compounds in a month. There is need to develop more such enzyme additives as a bioremediation agent. Thermophilic ArsI, a heat-stable enzyme that detoxifies toxic trivalent organoarsenical, has a potential to be a bioremediation agent. Before it can be used as an effective bioremediation agent, it must be fully characterized, which has been the goal of my studies.

7.3 Future Research Perspectives

Here, I have reported the biochemical characterization of a novel ArsI, a C-As lyase. Additionally, I determined the carbon product of roxarsone degradation. Thus, this study helped to proposed mechanism for the ArsI dioxygenase. It is interesting to see how ArsI mechanism is different compare to other extradiol dioxygenases in the way substrate bind to cysteines present on the loop and move to metal center and catalysis of wide range of trivalent organoarsenicals including both aromatic and non-aromatic arsenicals. Moreover, unlike most of the extradiol dioxygenases, ArsI break C-As bond without aromatic ring cleavage. In addition to biochemical characterization, I also reported soil isolate *Burkholderia* sp. MR1 genome sequence which would contribute in the identification of the MAs(V) reductase. The future study would include-

1> In detail study of factors involved in Rox(III) spectral assay and how it can utilize for real-time ArsI catalysis assay

2> Thermodynamic and structural aspects of MAs(III), Rox(III) and Fe(II) binding to ArsI protein using differential scanning calorimeter (DSC) and circular dichroism (CD) spectrometer under an anaerobic chamber

3> Structural characterization of ArsI protein with substrate bound using x-ray absorption spectroscopy (XAS) to characterize metals and metalloids bound to proteins.

4> Identification of protein fraction of *Burkholderia* sp. MR1 capable of MAs(V) reduction and determination of amino acid sequence, and DNA

sequence. The reported *Burkholderia* sp. MR1 genome sequence would be helpful to compare identified DNA sequence within the genome.

5> The identified DNA sequence MAs(V) reductase activity can be confirmed by making a clone and expressing in a suitable vector.

6> If the constructed gene clone could reduce MAs(V), it shows that the identified DNA sequence is the MAs(V) reductase.

7> Biochemical characterization of identified MAs(V) reductase.

7.4 References

Bencko V and Symon K (1977) "Test of environmental exposure to arsenic and hearing changes in exposed children." Environmental Health Perspectives **19**: p. 95.

Dheeman DS, Packianathan C et al. (2014) "Pathway of human AS3MT arsenic methylation." Chemical Research in Toxicology **27**(11): p. 1979-1989.

Drobná Z, Del Razo LM, García-Vargas GG et al. (2013). "Environmental exposure to arsenic, AS3MT polymorphism and prevalence of diabetes in Mexico." Journal of Exposure Science and Environmental Epidemiology **23**(2), 151-155.

Falk H, Caldwell GG et al. (1981) "Arsenic-related hepatic angiosarcoma." American Journal of Industrial Medicine. **2**(1): p. 43-50.

Gladysheva TB, Oden KL and Rosen BP (1994) "Properties of the arsenate reductase of plasmid R773." Biochemistry **33**(23): p. 7288-93.

Leitgeb S, Straganz GD and Nidetzky B (2009) "Biochemical characterization and mutational analysis of the mononuclear non-haem Fe^{2+} site in Dke1, a cupin-type dioxygenase from Acinetobacter johnsonii." Biochemical Journal **418**(2): p. 403-411.

Marapakala K, Qin J and Rosen BP (2012) "Identification of catalytic residues in the As(III) S-adenosylmethionine methyltransferase." Biochemistry **51**(5): p. 944-51.

Meng YL, Liu Z, and Rosen BP (2004) "As(III) and Sb(III) uptake by GlpF and efflux by ArsB in Escherichia coli." Journal of Biological Chemistry **279**: p. 18334-41.

Mukherjee SC, Rahman MM, Chowdhury UK et al. (2003) “Neuropathy in arsenic toxicity from groundwater arsenic contamination in West Bengal, India.” Journal of Environmental Science and Health, Part A **38**(1): p. 165-183.

Nachman KE, Baron PA, Raber G et al. (2013) “Roxarsone, inorganic arsenic, and other arsenic species in chicken: a US-based market basket sample.” Environmental Health Perspectives **121**(7): p. 818.

Nadar SV, Yoshinaga M et al. (2014) “Crystallization and preliminary X-ray crystallographic studies of the ArsI C-As lyase from Thermomonospora curvata.” Acta Crystallographica Section F: Structural Biology Communications **70**(Pt 6): p. 761-4.

Nadar SV, Yoshinaga M et al. (2016) “Structure of the ArsI C–As Lyase: Insights into the Mechanism of Degradation of Organoarsenical Herbicides and Growth Promoters.” Journal of Molecular Biology **428**(11): p. 2462-2473.

Qin J, Lehr CR, Yuan C et al. (2009) “Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga.” Proceedings of the National Academy of Sciences **106**(13): p. 5213-7.

Qin J, Rosen BP, Zhang Y, Wang G et al. (2006) “Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase.” Proceedings of the National Academy of Sciences **103**(7): p. 2075-80.

States JC, Srivastava S et al. (2008). “Arsenic and cardiovascular disease.” Toxicological Sciences **107**(2), 312-323.

Tyler, C.R. and A.M. Allan (2014) “The effects of arsenic exposure on neurological and cognitive dysfunction in human and rodent studies: a review.” Current Environmental Health Reports **1**(2): p. 132-147.

Yang J, Rawat S et al. (2010) “Arsenic binding and transfer by the ArsD As(III) metallochaperone.” Biochemistry **49**(17): p. 3658-66.

Yan Y, Xue XM, Guo et al. (2017) “Co-expression of cyanobacterial genes for arsenic methylation and demethylation in Escherichia coli offers insights into arsenic resistance.” Frontiers in Microbiology **8**.

Yan Y, Ye J et al. (2015) “Arsenic demethylation by a C· As lyase in cyanobacterium Nostoc sp. PCC 7120.” Environmental Science & Technology **49**(24): p. 14350-14358.

Yoshinaga M, Cai Y and Rosen BP (2011) Demethylation of methylarsonic acid by a microbial community. Environmental Microbiology **13**(5): p. 1205-15.

Yoshinaga M and Rosen BP (2014) “A CAs lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters.” Proceedings of the National Academy of Sciences **111**(21): p. 7701-6.

Zhou Tand Rosen BP(1997) “Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase.” Journal of Biological Chemistry **272**(32): p. 19731-7.

Zhou T, Shen J et al. (2002) “Unisite and multisite catalysis in the ArsA ATPase.” Journal of Biological Chemistry **8**(6): p. 23815-23820.

VITA

SHASHANK S PAWITWAR

EDUCATION

2005 - 2008	Bachelor of Science (BS) in Biotechnology Swami Ramanand Teerth Marathwada University Nanded, Maharashtra, India
2008 - 2010	Master of Science (MS) in Biotechnology Amity University Delhi (NCR), Uttar Pradesh, India
2012 - 2017	Doctoral of Philosophy (Ph.D.) in Biomedical Sciences Herbert Wertheim College of Medicine Florida International University Miami, Florida, USA

PUBLICATIONS AND PRESENTATIONS

Pawitwar SS, Yoshinaga M, Nadar VS, Kandegedara A, Stemmler and Rosen BP. Biochemical Characterization of ArsI, a novel C-As lyase for the degradation of organoarsenical herbicides. Environmental Science and Technology (2017) (Manuscript under review)

Pawitwar SS, Dhar S, Tiwari S, Ojha CR, Lapierre J, Martins K, Rodzinski A, Parira T, Paudel I, Li J, Dutta RK, Silva MR, Kaushik A and El-Haze Nazira. Overview on the Current Status of Zika Virus Pathogenesis and Animal Related Research. Journal of Neuroimmune Pharmacology, Springer (2017)

Li J, Pawitwar SS, and Rosen BP. The organoarsenical biocycle centers around the primordial antibiotic methylarsenite. Metallomics, Royal Society of Chemistry (2016) [Equal contribution authorship].

Nadar VS, Yoshinaga, M., Pawitwar SS, Kandavelu P, Sankaran B and Rosen BP. The structure of the ArsI C-As lyase: insights into the mechanism of degradation of organoarsenical herbicides and growth promoters. Journal of Molecular Biology (2016).

Pawitwar SS, Utturkar SM, Brown SD, Yoshinaga M, Rosen BP. Draft genome sequence of Burkholderia sp. MR1, a methylarsenate-reducing bacterial isolate from Florida golf course soil. Genome Announc. 3(3):e00608-15, American Society of Microbiology (2015)

Presented a poster titled “ArsI: a novel C-As lyase for degradation of environmental organoarsenicals in the 50th Miami Winter Symposium 2017 on 22-25 January 2017 at Hyatt Regency, Miami, USA.

Presented a poster titled “Biochemical characterization of ArsI: a novel C-As lyase for degradation of organoarsenicals” in Graduate Student Appreciation Week on 28th March 2016 at Florida International University, Miami, USA.