Submitted to Journal of Inorganic Biochemistry April 7, 2022

The ArsI C-As lyase: Elucidating the catalytic mechanism of degradation of organoarsenicals

Venkadesh Sarkarai Nadar[1], Palani Kandavelu[2], Banumathi Sankaran[3], Barry P. Rosen[1], orcid.org/0000-0002-5230-4271 and Masafumi Yoshinaga[1]*, orcid.org/0000-0002-7243-1761

[1]*Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199, USA*

[2]*SER-CAT and the Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA*

[3]*Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley Laboratory, Berkeley Center for Structural Biology, Berkeley, CA 94720, USA*

***Corresponding author**: Masafumi Yoshinaga (myoshina@fiu.edu), Department of Cellular Biology and Pharmacology, Florida International University, Herbert Wertheim College of Medicine, 11200 SW 8th Street, AHC1 419G, Miami, Florida, U. S. A. 33199, Tel: (+1) 305-348-1489, Fax: (+1) 305-348-0651

Running title: Catalytic mechanism of the ArsI C-As lyase

Abbreviations: 2,3-HPCD, homoprotocatechuate 2,3-dioxygenase; As(III), arsenite; βME, β-mercaptoethanol; BphC, 2,3-dihydroxybiphenyl 1,2-dioxygenase; Fe(II), ferrous iron; ITC, isothermal titration calorimetry; MAs(III), methylarsenite; MAs(V), methylarsenate; MSMA, monosodium methylarsenate; Ni(II), nickel; NTA, nitrilotriacetic acid; Rox(V), roxarsone (4–

hydroxyl-3-nitrophenylarsenate)

Abstract

Organoarsenicals such as monosodium methylarsenate (MSMA or MAs(V)) and roxarsone (4–hydroxyl-3-nitrophenylarsenate or Rox(V)) have been extensively used as herbicides and growth enhancers for poultry, respectively. Degradation of organoarsenicals to inorganic arsenite (As(III)) contaminates crops and drinking water. One such process is catalyzed by the bacterial enzyme ArsI, whose gene is found in many soil bacteria. ArsI is a non-heme ferrous iron (Fe(II))-dependent dioxygenase that catalyzes oxygen-dependent cleavage of the carbon-arsenic (C-As) bond in trivalent organoarsenicals, degrading them to inorganic As(III). From previous crystal structures of ArsI, we predicted that a loop-gating mechanism controls the catalytic reaction. Understanding the catalytic mechanism of ArsI requires knowledge of the mechanisms of substrate binding and activation of dioxygen. Here we report new ArsI structures with bound Rox(III) and mutant enzymes with alteration of active site residues. Our results elucidate steps in the catalytic cycle of this novel dioxygenase and enhance understanding of the recycling of environmental organoarsenicals.

Introduction

Arsenic is a ubiquitous toxic and carcinogenic metalloid, wide spread in the earth's crust that enters the biosphere from natural geothermal sources such as volcanic emissions. Anthropogenic sources of arsenic include organoarsenical herbicides such as monosodium methylarsenate (MSMA or MAs(V)), which has been used for selective post-emergent weed control in golf courses (1), cotton fields, highway rights-of-way and sod farms (2). In addition, synthetic aromatic arsenicals such as roxarsone (4-nitro-3-hydroxyphenylarsate or Rox(V)) were used since the 1940s as antimicrobial growth promoter in animal husbandry (3). In the arsenic biogeochemical cycle, microorganisms play a major role in the biotransformation of inorganic arsenic to organoarsenicals and degradation of organoarsenicals back to inorganic arsenic (4). For example, both prokaryotic and eukaryotic microbes methylate trivalent inorganic arsenite (As(III)) to highly toxic trivalent organoarsenicals such as methylarsenite (MAs(III)) and dimethylarsenite . Microbes in both soil and water detoxify the trivalent methylated species and trivalent aromatic arsenicals to inorganic arsenic by demethylation and dearylation (5-8). These reactions increase arsenic mobility, creating environmental hazards (4, 9).

We previously identified a two-step pathway for MAs(V) degradation from bacteria isolated from Florida golf course soil (9). In the first step, some members of the microbial community reduce MAs(V) to MAs(III), and, in the second step, other community members demethylate MAs(III) to inorganic As(III) (9). From a *Bacillus* species isolated from that soil, we cloned the *arsI* gene responsible for MAs(III) demethylation. The product of *arsI* gene is ArsI, a non-heme ferrous iron (Fe(II))-dependent extradiol dioxygenase that cleaves the C-As bond (10). In Fe(II)-dependent ring-cleaving extradiol dioxygenase enzymes, the role of Fe(II) is to activate O_2, which in turn is incorporated into the target C-C bond, forming a C-O bond that results in cleavage of the C-C bond (11). In the case of ArsI, the C-As bond would be cleaved rather than a C-C bond. ArsI not only degrades MAs(III) but also the trivalent form of aromatic organoarsenicals such as roxarsone,

nitarsone (*p*-nitrophenyl arsenate) and atoxyl (*p*-aminophenyl arsenate or *p*-arsanilic acid) (12). To date, ArsI is the only known enzyme that cleaves the C-As bond, thus the enzymatic pathway of this novel enzyme is considerable interest.

The overall structure and active site of TcArsI, the thermophilic C-As lyase ortholog from *Thermomonospora curvata* has been reported (13). However, the catalytic mechanism of the enzyme remains unclear because of the lack of structures with bound substrates. Our previous crystallographic and biochemical results demonstrate that the Fe(II)-binding site is composed of the triad of residues Gln8, His65 and Glu117, and substrate-binding site is composed of the vicinal cysteine residues Cys98 and Cys99, and single mutations of any of those residues resulted in loss of C-As lyase activity (12,13). Our previous structural data show that Lys105 and Gln103 are in the second coordination sphere, and Tyr38 interacts with the second coordination sphere via a water molecule (13). In ring-cleaving extradiol dioxygenases, residues in the second coordination sphere are involved in activation of dioxygen and stabilization of reaction intermediates, playing essential roles in catalysis (11). We propose that Lys105, Gln103 and Tyr38 in ArsI are similarly involved in activation of dioxygen for C-As bond cleavage and stabilization of the reaction intermediates. Here we report the metal-binding and substrate-binding affinity and crystal structure of mutants with substitutions of those residues. For example, K105A, an inactive mutant, was crystallized with bound substrate trivalent roxarsone (Rox(III)) and catalytic metal Fe(II). The results of mutant structures contribute to our understanding of the binding mode of the substrate and the catalytic mechanism.

Materials and Methods

Site direct mutagenesis

Mutants were generated by site-directed mutagenesis using a quick-change mutagenesis kit

(Stratagene, La Jolla, CA). The forward and reverse oligonucleotide primers used for mutagenesis are shown in Supplementary Table 1. The codon of the Lys105 was mutated into those of alanine, arginine and glutamic acid, generating the K105A, K105R and K105E derivatives. The codon of Tyr38 was mutated to that of phenylalanine, generating the Y38F derivative. The codon of Qln103 was changed to codons of alanine and histidine to obtain derivatives Q103A and Q103H.

Protein purification and crystallization

All TcArsI constructs were generated without the final twenty-five C-terminal residues, which are not required for activity, and were expressed and purified as described previously (14). Crystallization conditions were as previously reported (14, 13). Crystallization trials were conducted in an anaerobic tent filled with argon gas to avoid oxidation of protein cysteines, Fe(II) and Rox(III). Proteins were incubated with Rox(III) for 30 min on ice, following which the ArsI-Rox(III) complex was buffer-exchanged using 10 kDa cutoff Amicon Ultra centrifugal filters (MilliporeSigma, Burlington, MA) to remove unbound Rox(III). Mutant proteins were crystallized with 15-20% polyethylene glycol 4000 and 0.1 M sodium acetate, pH 4.5, at room temperature. The crystals were frozen in liquid nitrogen at 100 K for X-ray diffraction analysis. The X-ray data of K105A were collected on beamline 5.0.2 at Advanced Light Source (ALS), Lawrence Berkeley National Laboratory with an ADSC Quantum 315 detector. The data of K105R and K105E were collected on beamline 22-ID at the Advanced Photon Source (APS), Argonne National Laboratory using a Mar300 detector. The K105A data were indexed and scaled using iMosflm and Scala software (15). The other two data sets were indexed and scaled using HKL2000 software (16). The data collection and index details are shown in Table 1. The structures were solved by molecular replacement with the structure of wild type TcArsI (PDB ID: 5CB9) used as the model. In K105A, a strong positive peak (Fo-Fc) at the 11 σ level was observed near the active site cysteine residues that resembled the aromatic ring of Rox(III) (Supplementary Fig. 1A (green)). The arsenic atom of Rox(III) was confirmed by the anomalous difference map (Supplementary

Fig. 1A (red)). After refinement, Rox(III) satisfactorily fitted in the electron density map (Supplementary Fig. 1B (blue)). In the K105R and K105E structures, the substrate-binding loop was not observed due to disorder. The K105R structure, a divalent nickel (Ni(II)) atom was identified and refined. Since no Ni(II) was added during crystallization, we assume that it came from the Ni- nitrilotriacetic acid (NTA) resin used for protein purification. The refinement statistics were shown in Table 1. Phaser (17) in the CCP4 suit (18) was used for structure solution, and Refmac5 (19) and Phenix (20) were used for structure refinement. Coot (21) was used for model and electron density visualization and manual fitting of the side chains. Molecular models were rendered with Pymol (22). All data sets and coordination were deposited in the National Center for Biotechnology Information Protein Data Bank.

Assay of ArsI activity

ArsI activity was assayed by demethylation of MAs(III) both *in vivo* by heterologous expression of wild type and mutant genes in *E. coli* and *in vitro* with purified proteins, as described previously (9, 10). Cells of *E. coli* BL21(DE3) carrying plasmids with wild type or mutant *TcarsI* genes were grown in lysogeny broth (LB) medium (23) containing 50 μg ml^{-1} kanamycin at 37 °C. Protein expression was induced by adding 0.1 mM isopropyl β-d-1-thiogalactopyranoside at when the culture reached an absorbance of 0.6 at 600 nm. After 2 h, cells were harvested and washed one time with ST 10^{-1} medium (24). Demethylation of MAs(III) by cells was conducted in ST^{-1} medium containing 1 μM MAs(III) overnight at 30 °C. Demethylation activity by purified enzyme (1 μM) was assayed in a buffer consisting of 0.1 M (3-(*N*-morpholino)propanesulfonic acid), 0.15 M KCl, pH 7.0, containing 3 mM tris(2-carboxyethyl)phosphine, 1 mM cysteine and 0.1 mM $FeSO_4$. The reaction was initiated by addition of 2 μM MAs(III), incubated for 2 h with shaking at 30 °C and terminated by addition of 5 mM ethylenediaminetetraacetic acid. Arsenic was speciated by high pressure liquid chromatography (HPLC) (NexSAR HPLC system, PerkinElmer, Waltham, MA) with BioBasic-18 5 μm C18 300 Å reverse-phase column (250 mm × 4.6 mm; Thermo Fisher

Scientific, Waltham, MA), and the amount of arsenic was quantified by inductively coupled plasma mass spectroscopy (ICP-MS) using an NexION 1000 (PerkinElmar) (10).

Assay of Fe(II) binding by isothermal titration calorimetry (ITC)

The binding properties of Fe(II) with wild type and mutant proteins were determined by ITC. To prevent oxidation of Fe(II), protein solutions were prepared with degassed water, and Fe(II) solutions were prepared by dissolving $(NH_4)_2Fe(SO_4)_2(H_2O)_6$ in degassed water. Binding assays were carried out using a MicroCal iTC200 (GE Healthcare Bio Sciences, Piscataway, NJ) with protein concentrations between 50 to 100 μM and Fe(II) concentrations from 0.5 to 1 mM. Data were collected at 37 °C with 20 injections at 10 min intervals with a stirring speed of 1000 rpm. Spectra were analyzed using Origin 7.0 software (TA Instruments, New Castle, DE), and stoichiometry ($\underline{n}$), binding constants (K_d) and enthalpy (ΔH (kcal/mol)) were calculated with a one site binding model.

Assay of substrate binding by quenching of intrinsic tryptophan fluorescence

Intrinsic tryptophan fluorescence quenching assays were performed using a Quanta-Master UV-vis QM-4 steady state spectroflurometer (PTI, Birmingham, NJ) at 23 °C. Fluorescence of 1 μM protein was assayed in a buffer consisting of 50 mM (3-(*N*-morpholino)propanesulfonic acid) , pH 7.5, containing 0.5 M NaCl, and 1 mM tris(2-carboxyethyl)phosphine at excitation and emission wavelengths of 295 and 340 nm, respectively. The fluorescence intensities without (I_o) and with (I) addition of the trivalent organoarsenical substrate Rox(III) as quenching reagent were acquired, and K_d values were calculated from *Io-I/Io* differences at various concentrations of Rox(III) using the Stern Volmer relationship (25)**.** All experiments were performed in triplicate with independent samples.

Results

Comparison of the active site structures of ArsI and a ring-cleaving extradiol dioxygenase

Our previous structural analysis identified three residues, Tyr38, Gln103 and Lys105 in TcArsI that form an inner shell of residues that interacts with the catalytic metal via water molecules (Fig .1A). Tyr38 and Lys105 are conserved among ArsI orthologs, whereas Gln103 is replaced by serine in some orthologs (Supplementary Fig. 2) Interestingly, the Type I extradiol ring-cleaving homoprotocatechuate 2,3-dioxygenase (2,3-HPCD) from *Brevibacterium fuscum* also has three amino acid residues that interact with the catalytic Fe(II) through water molecules: Asn157, His200, and Tyr257 (PDB ID: 2IG9) (26) (Fig 1B). His200 and Tyr257 but not Asn157 are highly conserved with other ring-cleaving extradiol dioxygenases such as AkbC (methylcatechol 2,3-dioxygenase from *Rhodococcus* sp. strain DK17) (27) and BphC (2,3-dihydroxybiphenyl 1,2-dioxygenase from *Pseudomonas* sp. strain KKS102) (28) (Supplementary Fig. 3). During ring cleavage by 2,3-HPCD, these three residues interact with dioxygen bound to Fe^{2+}, the substrate-alkylperoxo-Fe^{2+} intermediate and the open ring product in the oxygen activation and insertion steps, suggesting that the three residues are necessary for catalysis (26). Although these water-interacting residues are not conserved between 2,3-HPCD and TcArsI, their properties are similar to each other. Both proteins commonly utilize one polar residue (Asn157 in 2,3-HPCD and Gln103 in TcArsI), one aromatic residue (Tyr257 and Tyr38) and one positive residue (His200 and Lys105). By analogy, it is reasonable to assume thatTcArsI residues Tyr38, Gln103 and Lys105 that interact with Fe(II) via water molecules also play a critical role in catalysis.

The effect of ArsI mutations on catalytic activity

To test our hypothesis, seven single mutants of TcArsI residues Try38, Gln103 and Lys105 were constructed. The ability of *E. coli* cells expressing genes for wild type TcArsI and derivatives Y38F, Q103A/H and K105A/E/R to demethylate MAs(III) was determined by HPLC-ICP-MS analysis of the culture media (Fig. 2A). The catalytic activity of wild type TcArsI and derivatives was assayed directly with purified enzymes (Fig. 2B). The results from both *in vivo* and *in vitro* assays show

that the Q103H derivative retains demethylation activity similar to the wild type, the Y38F and Q103A derivatives retain less than 5% of wild-type activity, and the K105A/E/R derivatives are nearly inactive.

The effect of ArsI mutations on metal binding

The effect of amino acid substitutions on Fe(II) binding affinity was determined by ITC (Table 2, Supplementary Fig. 4). The K_d of wild type TcArsI was determined to be 0.4 μM, which is similar to those of BmArsI, the ArsI ortholog from *Bacillus* sp. MD1 (12) and other non-heme Fe(II) dependent ring-cleaving extradiol dioxygenases (29). The K_d of the K105R derivative was determined to be 0.5 μM, similar to that of the wild type enzyme. In contrast, the affinity of the Y38F and Q103A/H derivatives was an order of magnitude lower than the wild type, and the K105A/E derivatives bound Fe(II) with nearly 100-fold lower affinity compared to wild type TcArsI.

The effect of ArsI mutations on substrate binding

Binding of substrate by wild type TcArsI and derivatives was estimated from the quenching of intrinsic protein fluorescence by addition of increasing concentrations of Rox(III) (Fig. 3). There were no significant differences between the apparent affinity of wild type TcArsI and derivatives. These results demonstrate that substitution of the residues forming or interacting with the second coordination sphere have essentially no effect on substrate binding.

Structural analysis of binding of substrate and metal in K105A

Co-crystallization of TcArsI with Fe(II) and Rox(III) was conducted using wild type TcArsI and derivatives. To reduce oxidation of protein, substrate and metal, crystallization trials were performed in an anaerobic tent filled with argon gas. We successfully crystallized and solved the structures of the K105A derivative with bound Fe(II) and Rox(III). The difference map and anomalous map at the 11 and 7.5 σ levels confirm the presence of Rox(III) in the K105A structure

(PDB ID 5V0F) (Supplementary Fig. 1). The arsenic atom of Rox(III) is directly coordinated with the thiolates of Cys98 and Cys99 (Fig. 4A). The distance between the sulfur and arsenic atom is approximately 2.0 Å. The aromatic ring of Rox(III) stacked with the five-membered ring of His65 on one side and the six-membered ring of Trp107 on the other. The hydroxyl group of Rox(III) forms a hydrogen bond with the side chain of Asp64. In addition, hydrophobic interactions between the side chain of Ile53 and the aromatic ring of Rox(III) help to stabilize binding of the substrate. The arsenic atom of Rox(III) has a trigonal pyramidal geometry with the ring C1 and the two cysteine sulfur atoms (Fig. 4B). Tyr38, Gln103 and Glu117 make water-mediated interactions with the Fe atom (Fig. 4A).

In the previously reported TcArsI structure (PDB ID 5CB9), the sulfur atoms of the exogenously added reductant β-mercaptoethanol (βME) and Cys98 coordinated with the metal center (13). When the structure of the K105A derivative was superimposed on the TcArsI-βME structure, the substrate binding loop of the mutant protein was observed to move approximately 2.8 Å towards metal center compared with the loop in the in the TcArsI-βME structure (Fig. 5A). The loop movement towards metal center brings the C-As bond of Rox(III) to the first coordination sphere of metal (Fig. 4A). In the structure of TcArsI-Fe co-crystallized with Fe(II) under aerobic conditions (13), the Fe atom makes a five coordinate complex and forms a square pyramidal geometry with three coordinates from the metal-binding triad of residues Gln8, His65 and Glu117 and two coordinations from water molecules (PDB ID: 5D4F) (Fig. 6, brown sphere (a)). The Fe binding mode in the K105A derivative differs from the TcArsI-Fe structure (Fig. 6, brown sphere (b)). Compared with the structure of TcArsI-Fe, the position of the Fe atom in TcArsI-K105A shifts 1.9 Å from that in TcArsI-Fe, which flips the side chain of Glu117 away from the metal center. Due to this shift, Glu117, which directly coordinates with the Fe atom in TcArsI-Fe, forms an indirect coordination with Fe through a water molecule. The other two coordinations are to the arsenic atom of the Rox(III) and a water molecule. The average distance between the Fe atom and the

coordinated protein residues is 3.1 Å, which is larger than that in the TcArsI-Fe structure (2.2 Å). The ITC results demonstrate that the K105A protein binds Fe(II) with two orders of magnitude lower affinity compared to wild type (Table 2, Supplementary Fig. 2B). The structural results are consistent with poorer binding of the Fe atom binds in the K105A protein compared with the wild type protein. In addition, the lower affinity results in a lower Fe occupancy of only 34% in structure of the mutant protein. The structural results provide insight into the reason for the reduction in metal affinity.

The previously reported crystal structure of TcArsI showed that the substrate binding loop is flexible, and the loop was captured in three different confirmations. Conformers 1 and 2 are distant from the active site and considered open states. Conformer 3 is close to the active site and considered to be a closed state. We proposed a loop gating mechanism in which the substrate binding loop serves as a gate to entry of the substrate to the active site (13). In the wild-type TcArsI-βME structure the substrate binding loop in the closed state is stabilized by two interactions with the metal center: one with bound βME and the other with Gln103 via a water molecule (Fig. 5B). In contrast, in the K105A there are three interactions between the metal center and the substrate binding loop in the closed state: the substrate-binding loop is stabilized by interaction of Rox(III) with the metal center and Gln103 through water molecule, as seen in the TcArsI-βME structure. In addition, in the K105A structure, there is an additional stacking interaction of Tyr100 with the β2 strand. This additional interaction may increase the stability of the substrate binding loop in the closed state by binding of substrate and make available the substrate for C-As binding by the metal center.

Structure of the K105R and K105E derivatives

The K105R derivative crystallized with two molecules in the asymmetric unit and Ni(II), which likely came from Ni-NTA resin used for protein purification, is in both metal binding sites (PDB ID

6XA0) (Fig. 7). In both molecules the substrate-binding loop is disordered. In the K105R structure, binding of Ni(II) is similar to the previously reported wild type structure with bound Ni(II), but Arg105 loses the water-mediated interaction with Ni(II) because the amino and guanidine groups of the side chain of Arg105 face away from the metal center in the K105R structure. Arg105 does not affect metal binding because it maintains a hydrogen bond between Glu117 and Arg105 by the guanidinium group of arginine (Fig. 7). The K105E derivative only crystallized without metal, and the active site loop is missing in the electron density (PDB ID 6XCK) (Fig. 8). In the previously reported wild type apo- and Ni(II)-bound structures (13), Lys105 forms a hydrogen bond with Glu117. In the K105E structure the side chain of Glu105 is oriented away from the metal center, and the protein loses the hydrogen bond between Glu117 and Glu105 (Fig. 8). This conformational change is predicted to reduce metal binding affinity, consistent with the ITC results that show that the K105E derivative has two orders lower affinity for Fe(II).

Discussion:

In catechol dioxygenase, amino acid side chains in the second coordination sphere of the metal binding site play several roles in catalysis. They serve as an active site base to encourage formation of the monoanionic substrate complex, stabilize the oxygen complex by hydrogen bonding and orient the substrates through steric interactions to promote oxygen attack (23). The structure-function analysis of TcArsI reported here contributes to our understanding of the catalytic roles of conserved second-sphere residues in the TcArsI C-As lyase.

Role of Cys98 and Cys99

ArsI utilizes a substate-binding mode that is unique in the dioxygenase family. Other members of that family such as the ring-cleaving extradiol catechol dioxygenase utilize direct interaction of the substrate with the metal center (30). In contrast, ArsI brings the substrate into contact with the metal center using an independent substrate binding site that is composed of two vicinal cysteine

residues in the flexible loop that are conserved in ArsI orthologs. We previously demonstrated that these conserved cysteine residues are critical for binding arsenical substrates as well as for catalytic activity (13). The previous crystallographic studies suggested a loop-gating mechanism of movement of the substrate binding site toward the metal center. However, understanding how the substrate interacts with the metal center has been limited due to lack of a structure with bound substrate. In this study, the K105A derivative crystallized with both Rox(III) and the Fe atom. This may have been successful because the K105A mutant is catalytically inactive. This first structure of ArsI with an arsenical substrate clearly depicts the interaction of the As atom of Rox(III) bound to the vicinal cysteine pair in a planar trigonal geometry with the Fe atom. In addition, the flexible binding loop orients the arsenic atom for dioxygen attack of the C-As bond, placing the metal and As at a distance of 3.6 Å from each other.

Role of Lys105

Conserved residue Lys105, which is near the metal center, is not involved directly in Fe(II) binding. However, the K105A and K105E derivatives have lower affinity for metal, suggesting an indirect role in metal binding. In the crystal structure of wild type TcArsI, Lys105 is observed to form a hydrogen bond with Glu117, one of triad of metal binding residues. In the K105A and K105E structures, the interaction with Glu117 is lost. This results in movement of Glu117 away from the metal center, which is predicted to result in lower affinity for metal, and consequently leading to loss of catalytic activity. This is confirmed by the results of ITC metal binding assays that directly demonstrate that both mutant proteins have two orders of magnitude lower affinity for metal than the wild type of enzyme. In contrast, the K105R derivative has wild-type affinity for metal, and the structure of K105R with metal shows that the guanidinium group of arginine maintains a hydrogen bond with Glu117.

Since the K105R mutation has little effects on substrate binding and metal binding, the question

arises why this mutation leads to loss of catalytic activity. Form the structure of other dioxygen-bound ring cleaving extradiol dioxygenases, a conserved water-interacting histidine residue was suggested to be involved in acid catalysis as a proton donor to the metal bound oxygen atom even though it is not directly involved in binding to substrate or metal (24). Based on this reasoning, we hypothesize that the conserved residue Lys105 in ArsI activates the bound dioxygen molecule, promoting attack of the initial adduct on the substrate. The structure of K105R clearly shows that the arginine amino and guanidinium groups face away from the metal center, so this residue should be unable to activate the dioxygen molecule. This is a reasonable explanation for the lack of catalytic activity in the K105R derivative.

Role of Tyr38

From the results of mutagenesis of conserved residue Tyr38 to a phenylalanine, it appears that the tyrosine hydroxyl group is not essential for catalysis. In catechol dioxygenase, the equivalent conserved tyrosine residue stabilizes the alkyl peroxo intermediate by hydrogen bonding (25). In K105A-Rox(III) structure,Tyr38 interacts through a water molecule with the metal, and this water molecule is also near the sulfur atom of Cys99 (Fig. 4A, Supplementary Fig. 5A). We predict that Tyr38 similarly stabilizes the reaction intermediate during C-As bond cleavage. The loss of the hydroxyl group in the phenylalanine derivative may remove a hydrogen bond interaction with the intermediate, resulting in slower catalysis.

Role of Gln103

Residue Gln103, which is conserved in most but not all orthologs (Supplementary Fig. 2), interacts with the metal through a water molecule. When Gln103 was changed to an alanine residue, catalytic activity was largely reduced, while the Q103H derivative retained full activity. In the wild-type TcArsI-βME and K105A-Rox(III) structures, Gln103 stabilizes the active site loop when the loop enters the catalytic site. In addition, Gln103 interacts with the metal through a water molecule

that is close to sulfur atom of Cys98 and to the arsenic atom of Rox(III). We predict that Gln103 also stabilize the reaction intermediate (Fig. 4A, Supplementary Fig. 5B). In the Q103A mutant, the hydrogen bond is lost, and so the active site loop and reaction intermediate are not stabilized.

A proposed catalytic mechanism

We previously proposed a loop-gating mechanism for ArsI substrate binding (13). Here we extend that proposal to a more complete description of the catalytic mechanism of C-As bond cleavage (Fig. 9). In the ArsI enzyme, Fe(II) is bound in an octahedral geometry with six coordinations, three to the facial conserved amino acid triad and three with solvent (Fig. 9A). In the first step, the vicinal cysteines Cys98 and Cys99 in the flexible substrate-binding loop capture the trivalent arsenical substrate and ferry it from solvent to the catalytic metal center, where the vicinal cysteine residues orient the bound substrate such that the C-As bond faces the metal center. An As-Fe coordination results when one water molecule is replaced by the As atom of Rox(III) (Fig. 9B). We predict that the dioxygen molecule binds the metal and replaces the third water molecule, although we have been unable to obtain structures with bound dioxygen. The negative charge of the dioxygen molecule is stabilized by the side chain of Lys105 (Fig. 9C). In extradiol catechol dioxygeneses, the active site histidine residue involves acid-base catalysis during C-C bond cleaving (31). We predict a similar mechanism in ArsI, where Lys105 is protonated by the side chain of Gln103, which prepares it to donate a proton to the dioxygen molecule. Protonation of dioxygen would then convert it into a superoxide anion, which makes a nucleophile attack on the C-As bond of Rox(III), forming an alkylperoxo-Fe(II) intermediate and cleaving the C-As bond. In summary, our structural and biochemical analyses demonstrate that ArsI has a catalytic mechanism similar to other extradiol dioxygenases while utilizing a novel mechanism of substrate binding and activation.

Accession numbers

5V0F, 6XA0, 6XCK

CRediT authorship contribution statement

Venkadesh S. Nadar: Funding acquisition, Project administration, Investigation, Methodology, Writing - original draft. **Palani Kandavelu**: Investigation. **Banumathi Sankaran**: Investigation. **Barry P. Rosen**: Funding acquisition, Project administration, Supervision, Writing - review & editing. **Masafumi Yoshinaga**: Funding acquisition, Project administration, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by NSF BIO/MCB grant 1817962 to M.Y., NIH grants R35GM136211, R01GM55425 and R01 ES023779 to B.P.R. and a pilot project grant from the Research Center in Minority Institutions, Florida International University (1U54MD012393-03) to V.S.N. This project utilized the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beam line of the Advanced Photon Source, Argonne National Laboratory. SER-CAT is supported by its member institutions, and equipment grants (S10_RR25528, S10_RR028976 and S10_OD027000) from the National Institutes of Health. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science and Office of Basic Energy Sciences under contract no. W-31-109-Eng-38. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The ALS-ENABLE beamlines are supported in part by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01 and the Howard Hughes Medical Institute. The Advanced Light Source is a Department of Energy Office of Science

406 User Facility under Contract No. DE-AC02-05CH11231.

407

408

409 **References:**

410 1. Chen Z, Cai Y, Liu G, Solo-Gabriele H, Snyder GH, Cisar JL. Role of soil-derived dissolved
411 substances in arsenic transport and transformation in laboratory experiments. Sci. Total Environ.
412 2008;406:180–189.
413 2. Bednar AJ, Garbarino JR, Ranville JF, Wildeman TR. Preserving the distribution of inorganic
414 arsenic species in groundwater and acid mine drainage samples. Environ. Sci. Technol.
415 2002;36:2213–2218.
416 3. Garbarino JR, Bednar AJ, Rutherford DW, Beyer RS, Wershaw RL. Environmental fate of
417 roxarsone in poultry litter. I. Degradation of roxarsone during composting. Environ. Sci. Technol.
418 2003;37:1509–1514.
419 4. Zhu YG, Yoshinaga M, Zhao FJ, Rosen BP. Earth abides arsenic biotransformations. Annu.
420 Rev. Earth Planet. Sci. 2014;42:443–467.
421 5. Sanders JG. Microbial role in the demethylation and oxidation of methylated arsenicals in
422 seawater. Chemosphere. 1979:135–7.
423 6. Lehr C, Polishchuk E, Radoja U, Cullen WR. Demethylation of methylarsenic species by
424 Mycobacterium neoaurum. Appl. Organomet. Chem. 2003;17:831–834.
425 7. Huang JH, Scherr F, Matzner E. Demethylation of dimethylarsinic acid and arsenobetaine in
426 different organic soils. Water Air Soil Pollut. 2007;182:31–41.
427 8. Yoshinaga M, Cai Y, Rosen BP. Demethylation of methylarsonic acid by a microbial community.
428 Environ. Microbiol. 2011;13:1205–1215.
429 9. Thomas DJ, Rosen BP. Arsenic methyltransferases. In: Kretsinger RH, Uversky VN,
430 Permyakov EA, editors. Encyclopedia of Metalloproteins. Springer New York; New York: 2013.
431 pp. 138–43.

10. Yoshinaga M, Rosen BP. A C–As lyase for degradation of environmental organoarsenical herbicides and animal husbandry growth promoters. Proc. Natl. Acad. Sci. U. S. A. 2014;111:7701–7706.

11. Bugg TD, Ramaswamy S. Non-heme iron-dependent dioxygenases: unravelling catalytic mechanisms for complex enzymatic oxidations. Curr. Opin. Chem. Biol. 2008;12:134–140.

12. Pawitwar SS, Nadar VS, Kandegedara A, Stemmler TL, Rosen BP, Yoshinaga M. Biochemical characterization of ArsI: a novel C–As Lyase for degradation of environmental organoarsenicals. Environmental science & technology. 2017; 51:11115-25.

13. Nadar VS, Yoshinaga M, Pawitwar SS, Kandavelu P, Sankaran B, Rosen BP. Structure of the ArsI C–As Lyase: insights into the mechanism of degradation of organoarsenical herbicides and growth promoters. Journal of molecular biology. 2016; 428:2462-73.

14. Nadar SV, Yoshinaga M, Kandavelu P, Sankaran B, Rosen BP. Crystallization and preliminary X-ray crystallographic studies of the ArsI C–As lyase from Thermomonospora curvata. Acta Crystallographica Section F: Structural Biology Communications. 2014;70:761-4.

15. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr. 2011; 67:271-81.

16. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997;276:307‑326.

17. McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. 2007;63:32‑41.

18. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 2011;67:235‑242.

19. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. D Biol. Crystallogr. 2004;60:2184‑2195.

20. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 2010;66:213‑221.

21. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 2004;60:2126‑2132.

22. DeLano WL. The PyMOL user's Manual. DeLano Scientific; San Carlos, CA: 2001. https://www.pymol.org/citing.

23. Sambrook J, Fritsch E.F, and Maniatis T. (1989) Molecular Cloning, a Laboratory Manual. Cold Spring Harbor Laboratory, New York.

24. Maki T, Hasegawa H, Watarai H, and Ueda K. (2004) Classification of dimethylarsenate-decomposing bacteria using a restrict fragment length polymorphism analysis of 16S rRNA genes. *Anal. Sci.* 20, 61-68.

25. Stern O, and Volmer M (1919) Uber die abklingzeit der fluoreszenz. *phys. Z* 20, 183-188.

26. Kovaleva EG, Lipscomb JD. Crystal structure of Fe2+ dioxygenase superoxo, alkylperoxo, and bound product intermediates. Science. 2007; 316: 453-457.

27. Cho HJ, Kim K, Sohn SY, Cho HY, Kim KJ, Kim MH, Kim D, Kim E, and Kang BS. (2010) Substrate-binding mechanism of type 1 extradiol dioxygenase. J. Biol. Chem. 285, 34643-34652.

28. Uragami Y, Senda T, Sugimoto K, Sato N, Nagarajan V, Masai E, Fukuda M, and Mitsui Y. (2001) Crystal structures of substrate free and complex forms of reactivated BphC, an extradiol type ring-cleavage dioxygenase. J. Inorg. Biochem. 83, 269-279.29. Leitgeb S, Straganz GD, Nidetzky B. Biochemical characterization and mutational analysis of the mononuclear non-haem

479 Fe^{2+} site in Dke1, a cupin-type dioxygenase from *Acinetobacter johnsonii*. Biochem
480 J. 2009;418(2):403–411.
481 30. Kovaleva EG, Lipscomb JD. Versatility of biological non-heme Fe(II) centers in oxygen
482 activation reactions. Nat. Chem. Biol. 2008; 4: 186-193.
483 31. Sato N, Uragami Y, Nishizaki T, Takahashi Y, Sazaki G, Sugimoto K, Nonaka T, Masai E,
484 Fukuda M, and Senda T. (2002) Crystal structures of the reaction intermediate and its homologue
485 of an extradiol-cleaving catecholic dioxygenase. J. Mol. Biol. 321, 621-636.

Figure 1. Comparison of the second coordination sphere of TcArsI (PDB ID: 5CB9) with extradiol ring-cleaving homoprotocatechuate 2,3-dioxygenase (2,3-HPCD) from *Brevibacterium fuscum* (PDB ID: 2IG9). (**A)** In TcArsI, Tyr38, Gln103 and Lys105 form an inner shell of residues. (**B)** In 2,3-HPCD, Asn157, His200, and Tyr257 form an inner shell of residues. Ni(II) and Fe(II) are shown as green and brown spheres respectively. Water molecules are shown as red spheres.

Figure 2. Effect of site direct mutagenesis on TcArsI activity. MAs(III) demethylation by *E. coli* cells heterologously expressing wild type and mutant TcArsI (A) and purified wild type and mutants TcArsI (B) were determined as described in *Materials and Methods*.

Figure 3. Rox(III) binding affinity of wild type mutant ArsI proteins. The relative affinity Rox(III) was estimated from the quenching of tryptophan fluorescence at the indicated concentrations of Rox(III), as described under Materials and Methods. The proteins were (●) wild-type TcArsI (apparent K_d = 2.5 μM); (■) Y38F (apparent K_d = 2.1 μM); (□) Q103A (apparent K_d = 2.5 μM); (♦) Q103H, (apparent K_d = 1.8 μM); (o) K105A (apparent K_d = 0.9 μM); (▼) K105E (apparent K_d = 1.1 μM); (Δ) K105R (apparent K_d = 0.7 μM)

Figure 4. Binding mode of Rox (III). (A) The arsenic atom of Rox(III) is directly coordinated with the vicinal cysteine pair Cys98 and Cys99. The distance between the sulfur and arsenic atom is approximately 2.0 Å. The aromatic ring of Rox(III) stacks with the five-membered ring of His65 on one side and the six-membered ring of Trp107 on the other side. The hydroxyl group of Rox(III) forms a hydrogen bond with the side chain of Asp64. Other than these interactions, the hydrophobic interaction between Ile53 and the aromatic ring of Rox(III) helps to stabilize substrate binding. **(B**) The arsenic atom of Rox(III) has a trigonal pyramidal geometry.

Figure 5. Superimposition of the structure of TcArsI-βME (gray) with TcArsI-Rox(III) (green).

The active site loop moves approximately 3.0 Å in the direction of the metal center, bringing the C-As bond of Rox(III) into closer to the metal than C-S bond of βME. The conformation of the loop is stabilized by the stacking interaction of Tyr100 with the β2 strand (**A**) and the water-mediated interaction of Gln103 with the metal center **(B**). In the TcArsI-βME structure, Tyr100 is flanked by the solvent channel **(A)**.

Figure 6. Superimposition of metal binding in Wild type TcArsI_Fe (PDB ID: 5D4F) (cyan) with K105A_Rox(III) (PDB ID: 5V0F) (green). The position of Fe atom of K105A_Rox(III) (brown sphere (b)) shifts 1.9 Å away from that of the wild-type protein (brown sphere (a)). This results in the Fe atom making only two coordinations (Gln8 and His65) with the protein molecule. The side chain of Glu117 moves away from the metal center but interacts with Fe through a water molecule. The other two coordinations are with the arsenic of the Rox(III) and a water molecule. The average Fe-ligand distance in the K105A structure is 3.2 Å, compared with 2.2 Å in the wild-type TcArsI structure. The position and occupancy of the Fe atom shows that it binds to K105A with lower affinity compared with the wild-type enzyme. Fe and water molecules are shown in brown and red spheres, respectively. The hydrogen bond in TcArsI_Fe shows in black and K105A_Rox(III) shows in gray. Fe atoms are in brown sphere. Water molecule in TcArsI_Fe in light purple and K105A_Rox(III) in red spheres.

Figure 7: Structure of K105R. (A) and (B) The K105R crystal has two molecules in the crystallographic asymmetric unit. In (**A**) binding of Ni(II) is similar to the previously reported structure, but in (**B**) Ni(II) has a perturbed geometry. In both conformations, Arg105 loses the water-mediated interaction with metal because the amino and guanidino groups of Arg105 face away from the metal center. The substitution of arginine for lysine in residue 105 does not affect metal binding because it maintains the hydrogen bond between Glu117 and Arg105 by the –C-NH-C- group of the arginine residue.

Figure 8. K105E in two conformations with Glu105. In the glutamate substitution of Lys105, **t**he side chain of moves away from the metal center and loses the hydrogen bond between Glu117 and Glu105. This is comparable to TcArsI-apo structure, in which Lys105 hydrogen bonds with Glu117 even in the absence of metal.

Figure 9. Proposed mechanism of C-As bond cleavage by ArsI. (A) Fe(II) is bound in an octahedral geometry with six coordination, three to the facial triad and three with solvent (blue). **(B)** The vicinal cysteine residues in the substrate-binding loop orient the C-As bond with the metal center. One water molecule is replaced by the As atom of Rox(III) by As-Fe coordination. **(C)** The dioxygen molecule binds the metal and replaces the third water molecule. In (**A**) and (**B)**, the position of Fe(II) and water molecules are derived from the wild type TcArsI-Fe and K105A structures. In (**C**), the Lys105 residue derived by superimposition of the K105A and TcArsI-βME structures and binding of dioxygen is modeled using the dioxygen-bound catechol dioxygenase structure (26).

Highlights

- ArsI is an Fe(II)-dependent dioxygenase that cleaves the C-As bond in organoarsenicals
- We report new structures of ArsI, the only known enzyme that cleaves the C-As bond
- Our iron- and substrate-bound ArsI structures shed light on the catalytic mechanism
- The residues in the second coordination sphere play roles in the ArsI catalysis
- We propose that a loop-gating mechanism governs the catalytic reaction in ArsI

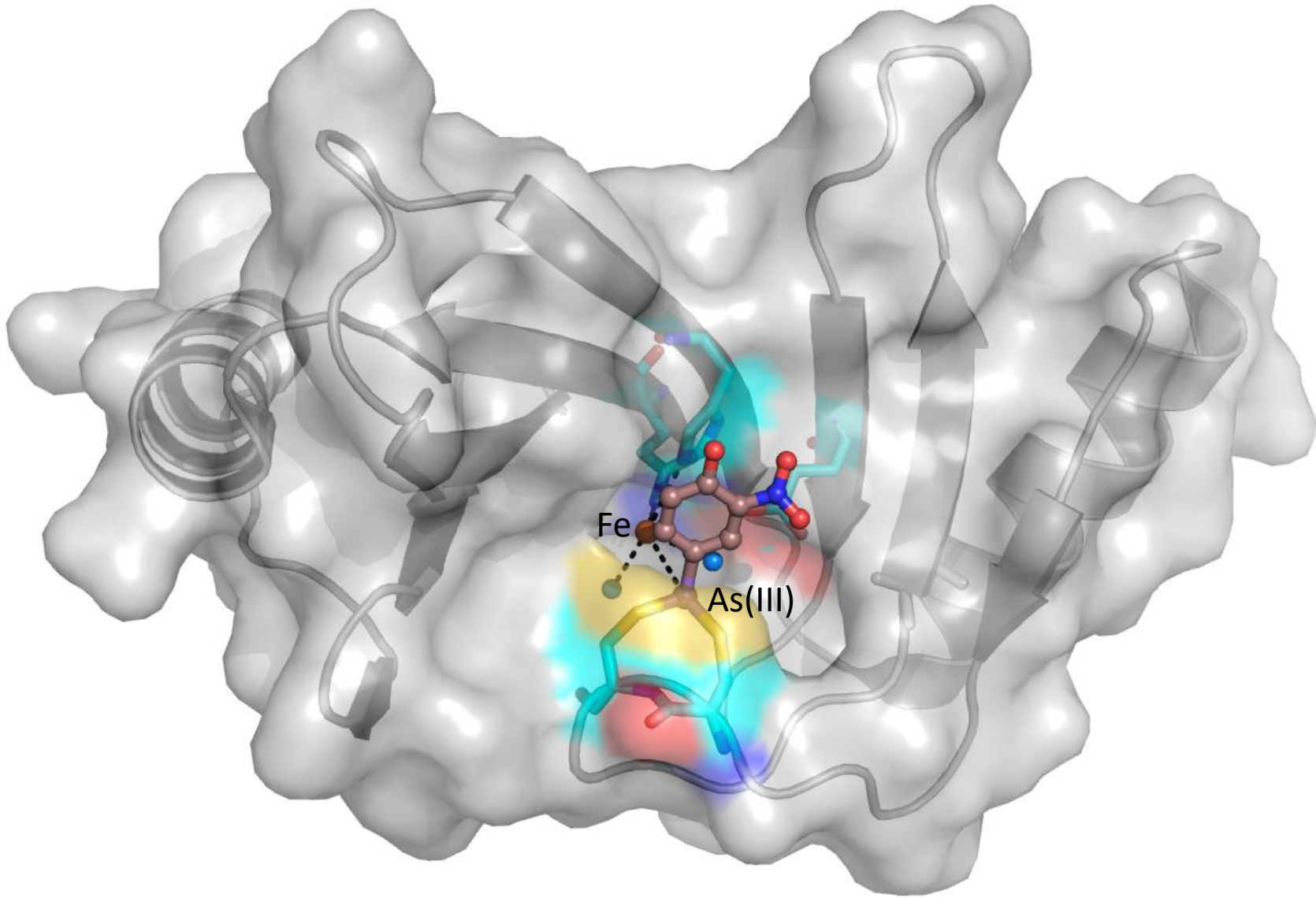
CRediT authorship contribution statement

Venkadesh S. Nadar: Funding acquisition, Project administration, Investigation, Methodology, Writing - original draft. **Palani Kandavelu**: Investigation. **Banumathi Sankaran**: Investigation. **Barry P. Rosen**: Funding acquisition, Project administration, Supervision, Writing - review & editing. **Masafumi Yoshinaga**: Funding acquisition, Project administration, Supervision, Writing - review & editing.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:


Fe
As(III)

Synopsis for the Graphical Abstract

ArsI is an Fe(II)-dependent dioxygenase that catalyzes oxygen-dependent cleavage of the carbon-arsenic (C-As) bond in trivalent organoarsenicals, degrading them to inorganic arsenite (As(III)). Here we report new ArsI structures with Fe(II) and substrate bound, elucidating steps in the catalytic cycle of this novel and unique dioxygenase.

Table 1. Data collection, indexing and refinement statistics.

	TcArsI-K105A	TcArsI-K105R	TcArsI-K105E
Source	ALS-5.0.2	APS 22-ID	APS 22-ID
Wavelength (Å)	0.9999	1.0000	1.0000
Resolution (Å)	41.77-2.23 (2.35-2.23)	50.00-2.15 (2.23-2.15)	50.00-1.62 (1.68-1.62
Space group	P43212	P212121	P212121
Unit cell	a=b=41.73 and c= 119.1 Å	a=42.08, b=43.17 and c=118.94	a=42.25, b=43.20 and c=119.43
Unique reflections	5438 (782)	12438 (1194)	28571 (2845)
Multiplicity	6.7 (6.9)	13.6 (13.9)	24.0 (13.8)
Completeness (%)	97.6 (98.8)	99.8 (100.0)	99.5 (100.0)
Mean I/sigma(I)	9.5 (2.5)	40.5 (8.3)	36.3 (5.4)
R-merge (%)	12.8 (86.3)	7.3 (43.5)	6.5 (55.6)
Wilson B-factor ($Å^2$)	31.2	23.2	12.5
R-work (%)	19.3	19.4	21.5
R-free (%)	26.7	26.8	25.4
No. of atoms			
Macromolecules	860	1691	1724
Ligands	43	2	0
Water	13	118	200
RMS (bonds) (Å)	0.016	0.008	0.007
RMS (angles) (°)	1.7	1.2	1.2
Ramachandran favored (%)	97.27	97.67	99.10
Ramachandran allowed (%)	1.82	1.86	0.90
Ramachandran outliers (%)	0.91	0.47	0.00
Average B-factor ($Å^2$)			
Macromolecules	35.22	22.6	15.4
Ligands	35.04	58.9	-
Water	42.68	31.5	30.3
PDB ID	5V0F	6XA0	6XCK

Figure 1

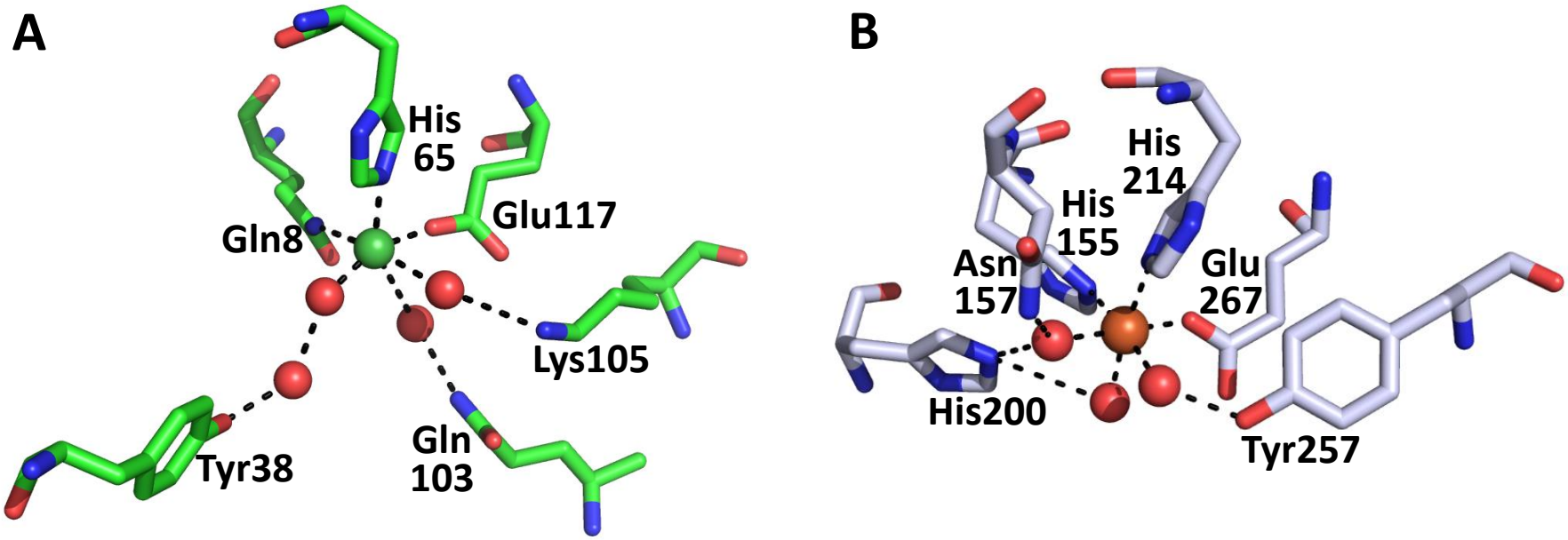

A
His
65
Glu117
Gln8
Lys105
Tyr38
Gln
103
B
His
214
His
155
Asn
157
Glu
267
His200
Tyr257

Figure 2

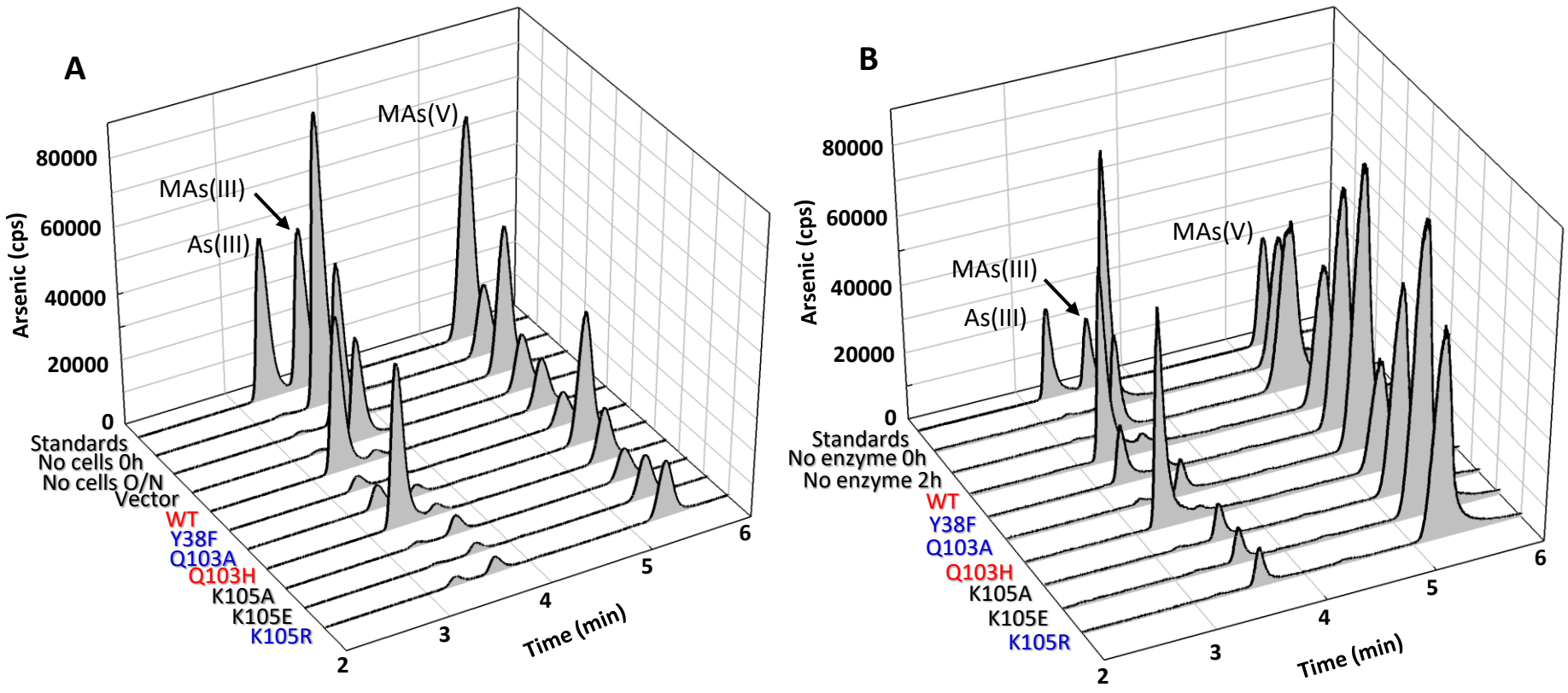

A
MAs(V)
MAs(III)
As(III)
80000
60000
40000
20000
0
Arsenic (cps)
Standards
No cells 0h
No cells O/N
Vector
WT
Y38F
Q103A
Q103H
K105A
K105E
K105R
2
3
4
5
6
Time (min)
B
MAs(V)
MAs(III)
As(III)
80000
60000
40000
20000
0
Arsenic (cps)
Standards
No enzyme 0h
No enzyme 2h
WT
Y38F
Q103A
Q103H
K105A
K105E
K105R
2
3
4
5
6
Time (min)

Figure 3

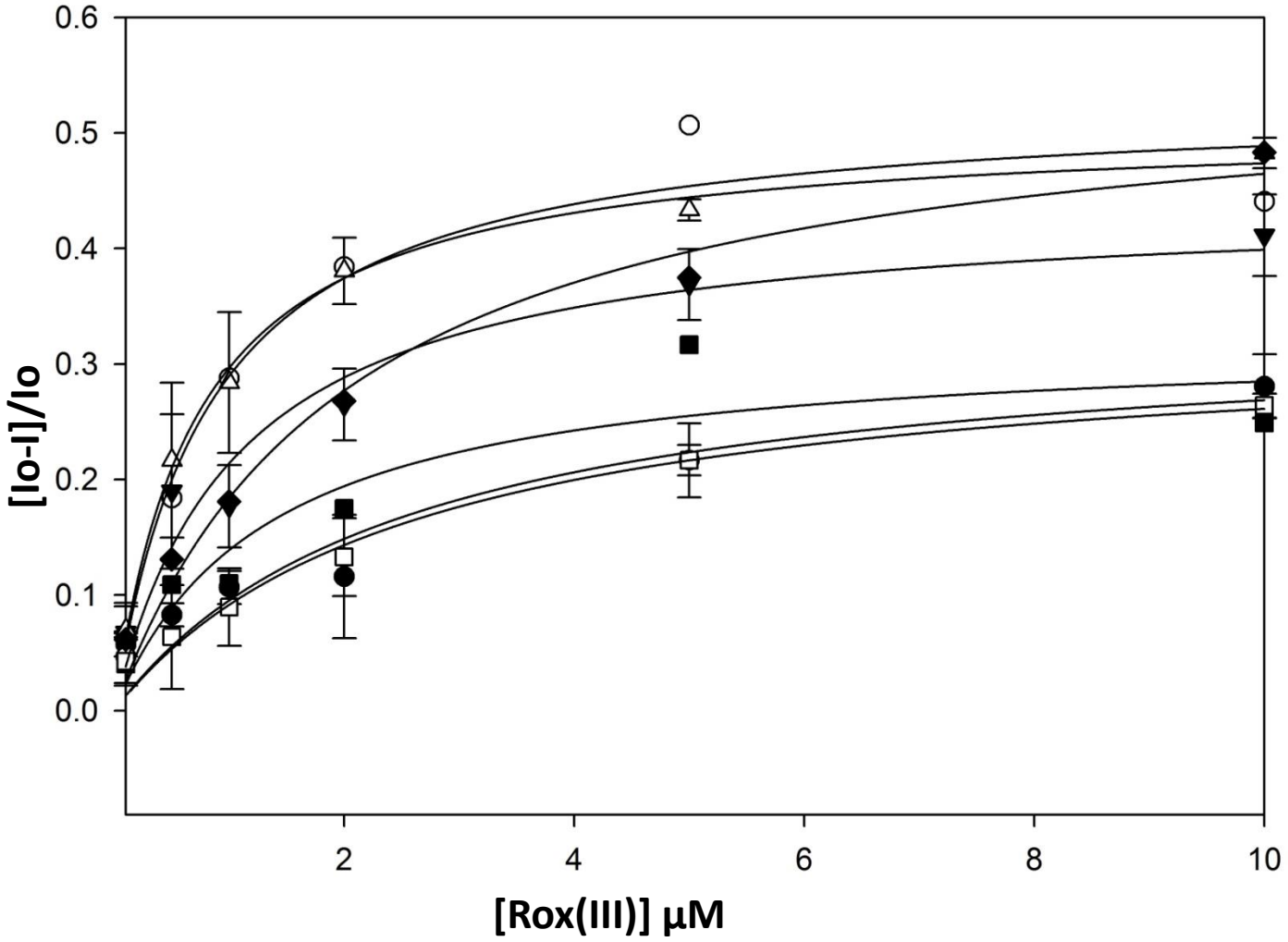


Figure 4

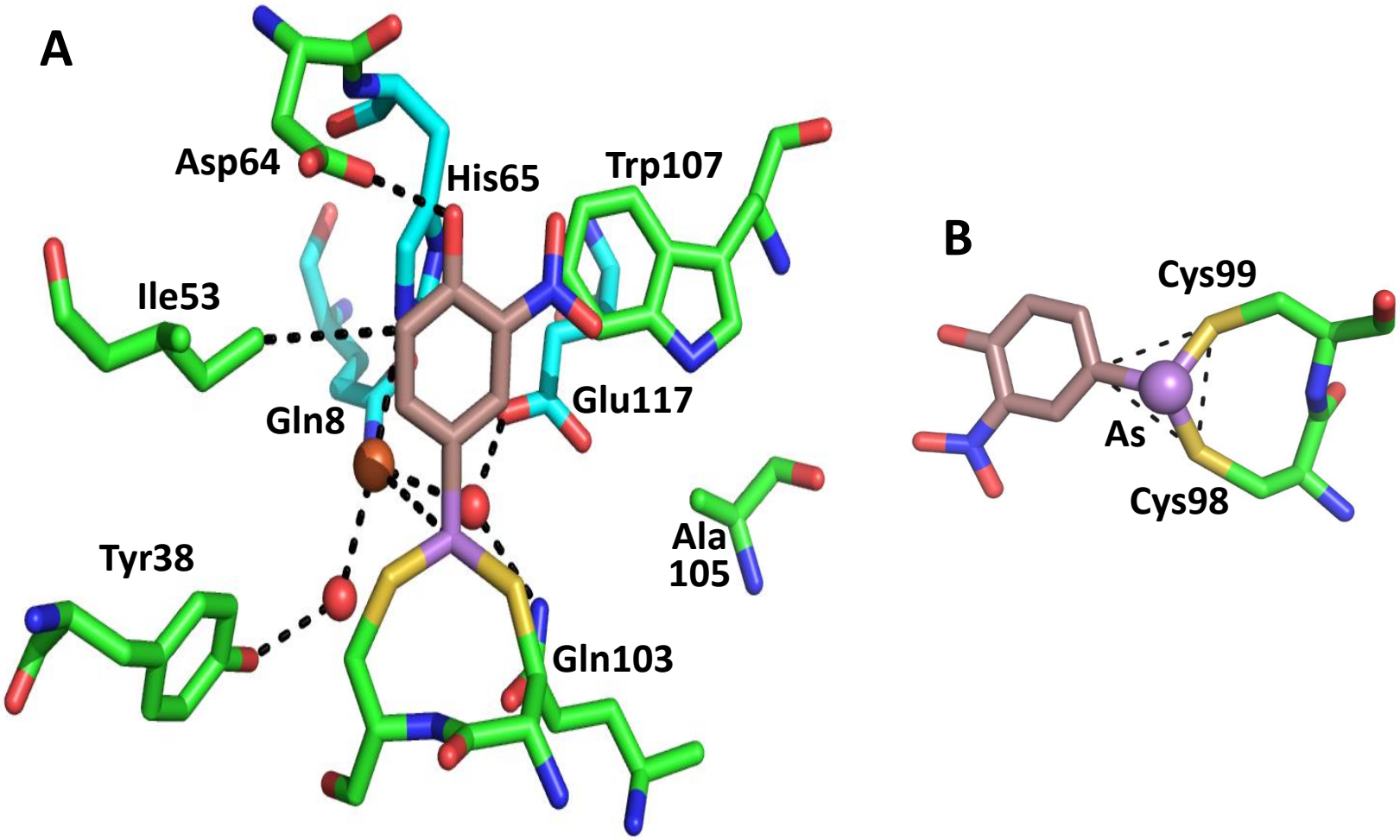

A
Asp64
His65
Trp107
Ile53
Gln8
Glu117
Tyr38
Ala
105
Gln103
B
Cys99
As
Cys98

Figure 5

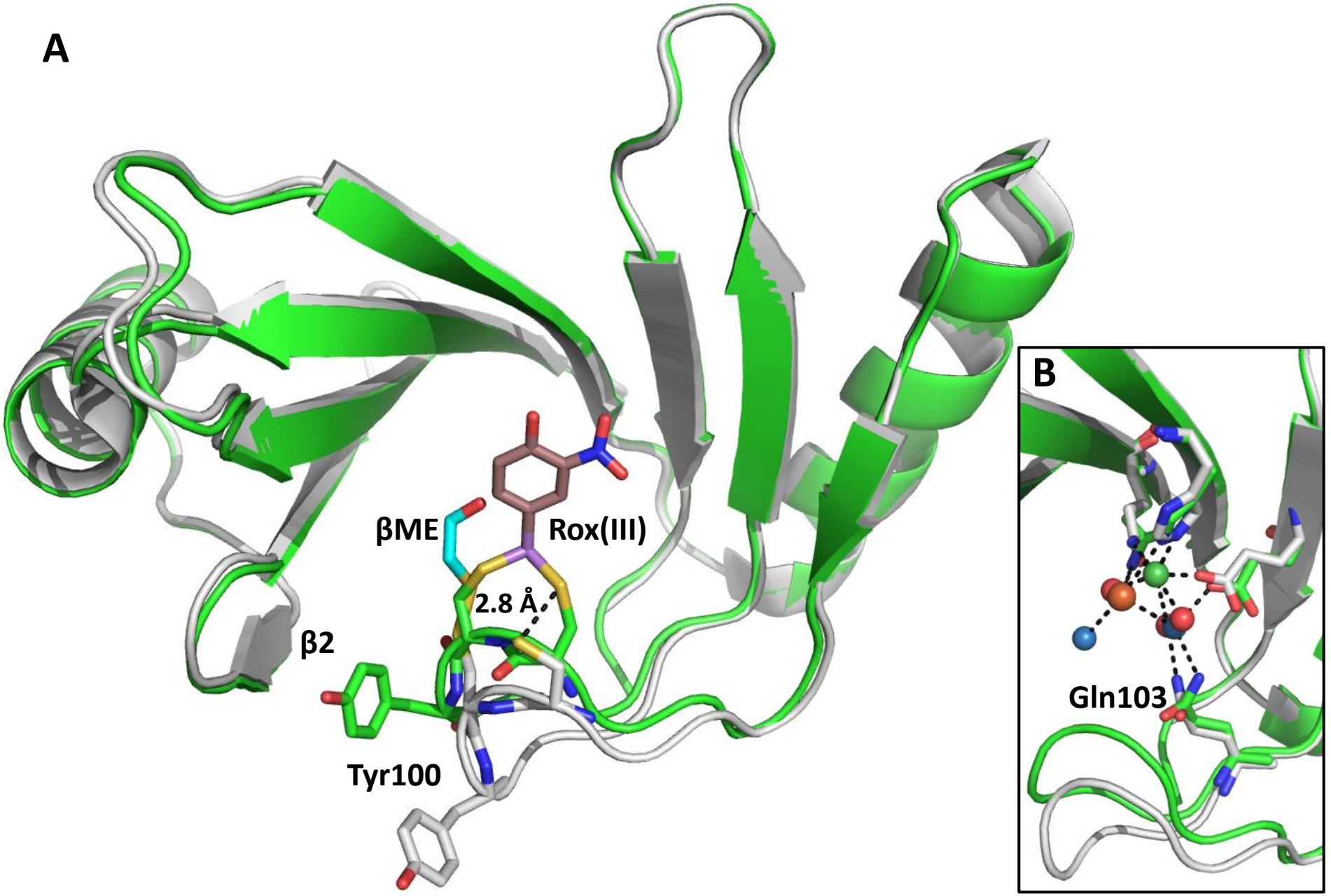


Figure 6

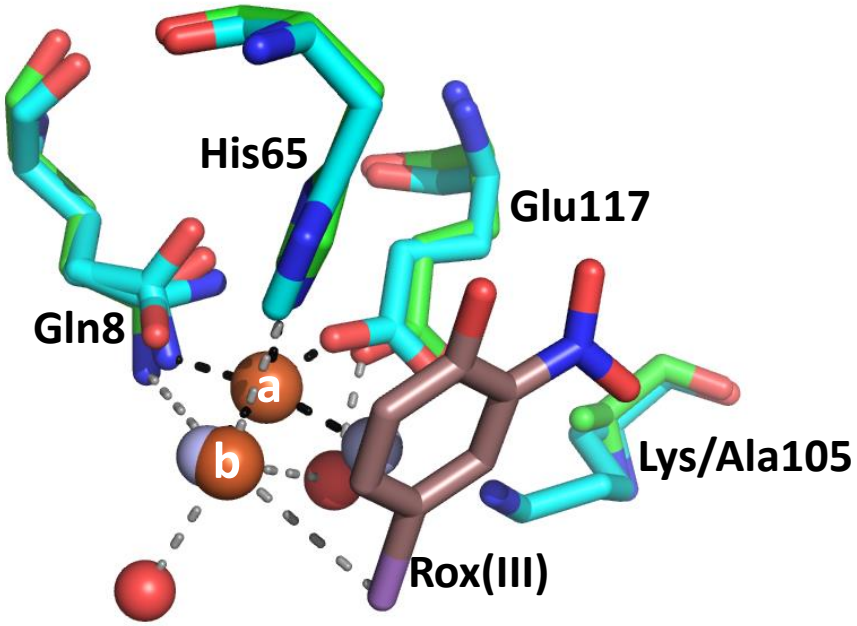


Figure 7

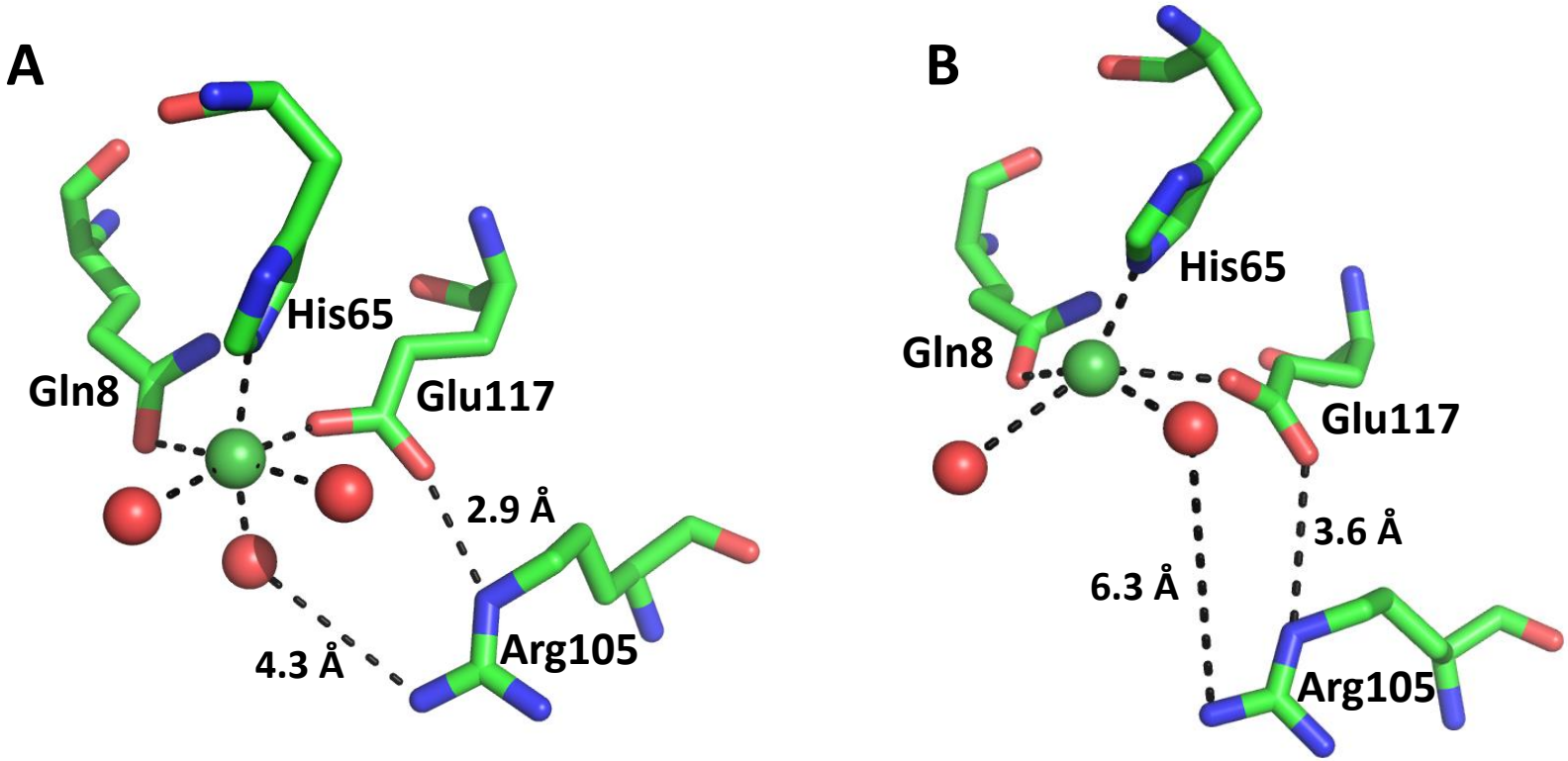

A
His65
Gln8
Glu117
2.9 Å
4.3 Å
Arg105
B
His65
Gln8
Glu117
3.6 Å
6.3 Å
Arg105

Figure 8

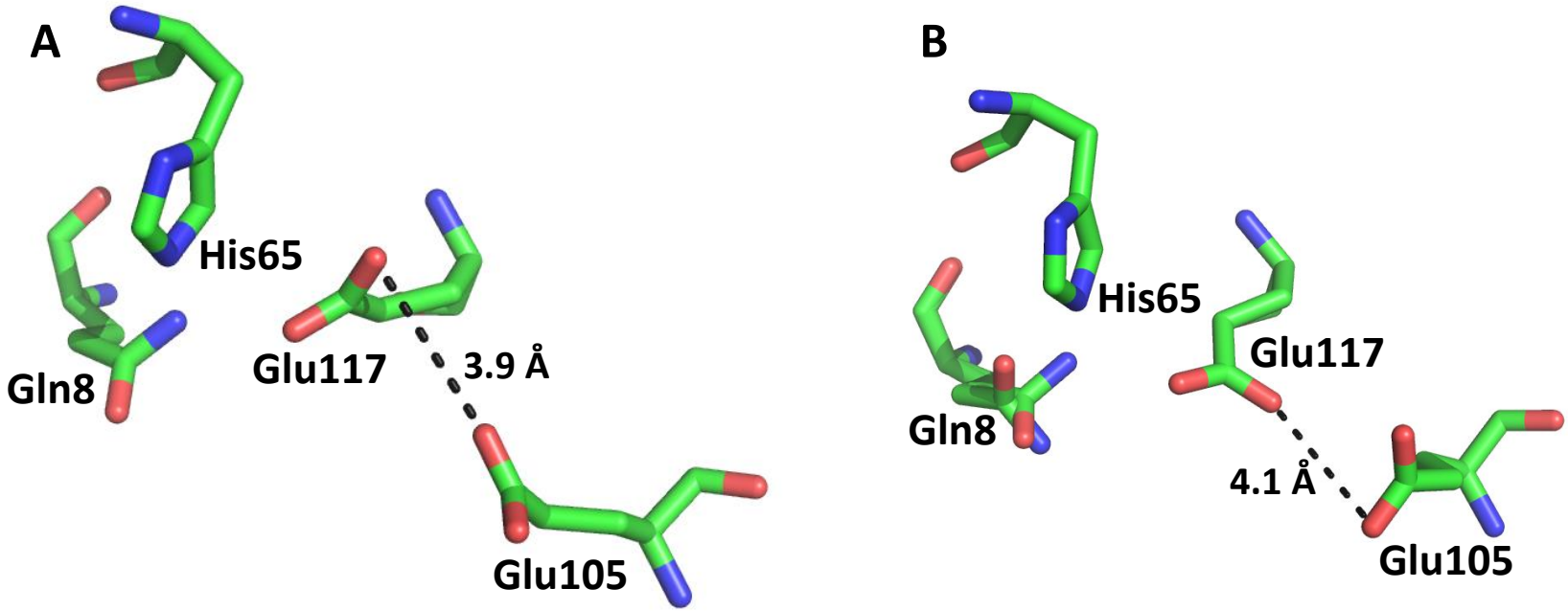
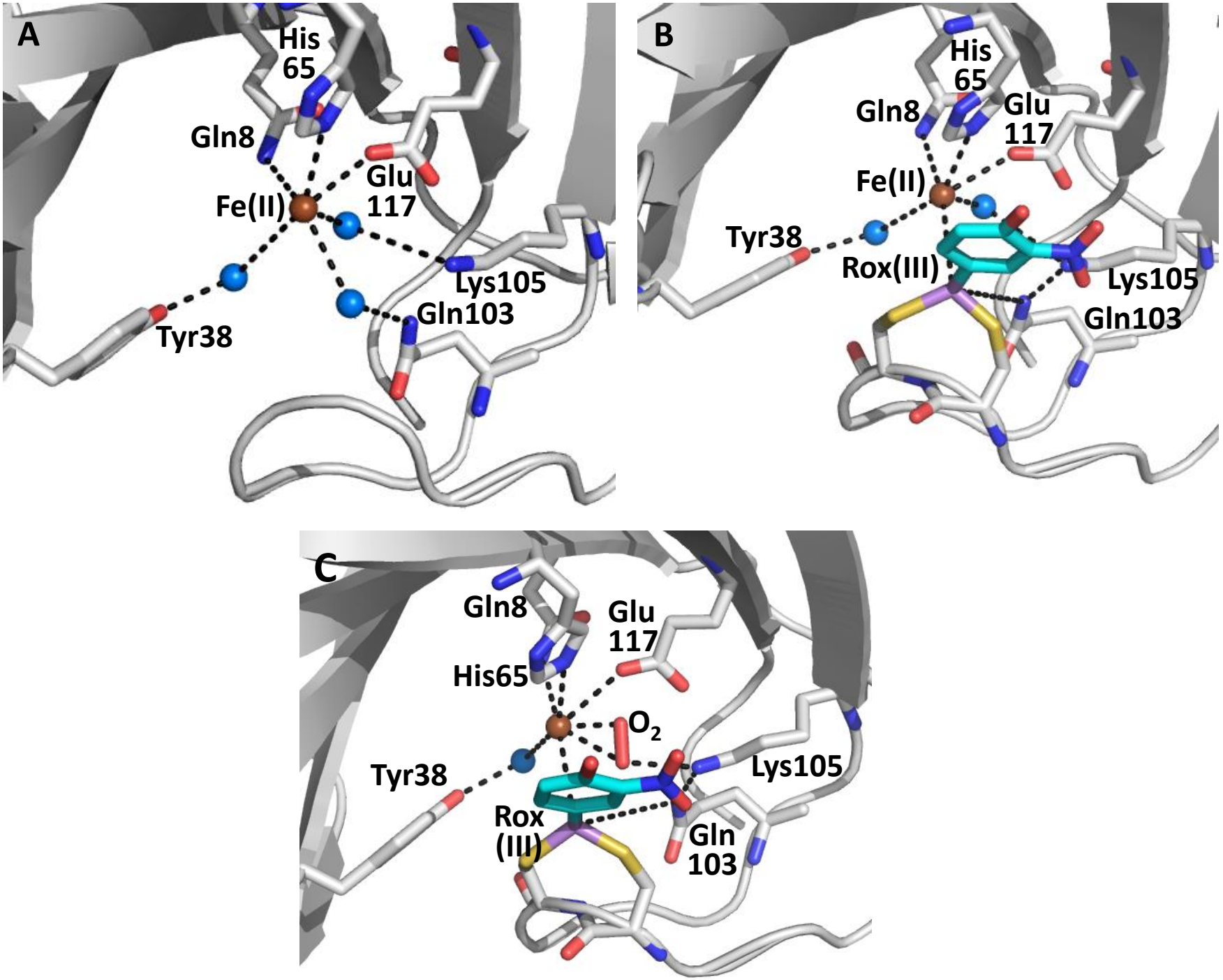

A
His65
Gln8
Glu117
3.9 Å
Glu105
B
His65
Glu117
Gln8
4.1 Å
Glu105

Figure 9


A
His 65
Gln8
Glu 117
Fe(II)
Lys105
Gln103
Tyr38
B
His 65
Gln8
Glu 117
Fe(II)
Tyr38
Rox(III)
Lys105
Gln103
C
Gln8
Glu 117
His65
O2
Tyr38
Lys105
Rox (III)
Gln 103

Submitted to Journal of Inorganic Biochemistry March 7, 2022

The ArsI C-As lyase: Elucidating the catalytic mechanism of degradation of organoarsenicals

Venkadesh Sarkarai Nadar[1], Palani Kandavelu[2], Banumathi Sankaran[3], Barry P. Rosen[1] and Masafumi Yoshinaga[1]*

[1]*Department of Cellular Biology and Pharmacology, Herbert Wertheim College of Medicine, Florida International University, Miami, FL 33199, USA*

[2]*SER-CT and the Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA*

[3]*Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley Laboratory, Berkeley Center for Structural Biology, Berkeley, CA 94720, USA*

***Corresponding author**: Masafumi Yoshinaga (myoshina@fiu.edu), Department of Cellular Biology and Pharmacology, Florida International University, Herbert Wertheim College of Medicine, 11200 SW 8th Street, AHC1 419G, Miami, Florida, U. S. A. 33199, Tel: (+1) 305-348-1489, Fax: (+1) 305-348-0651

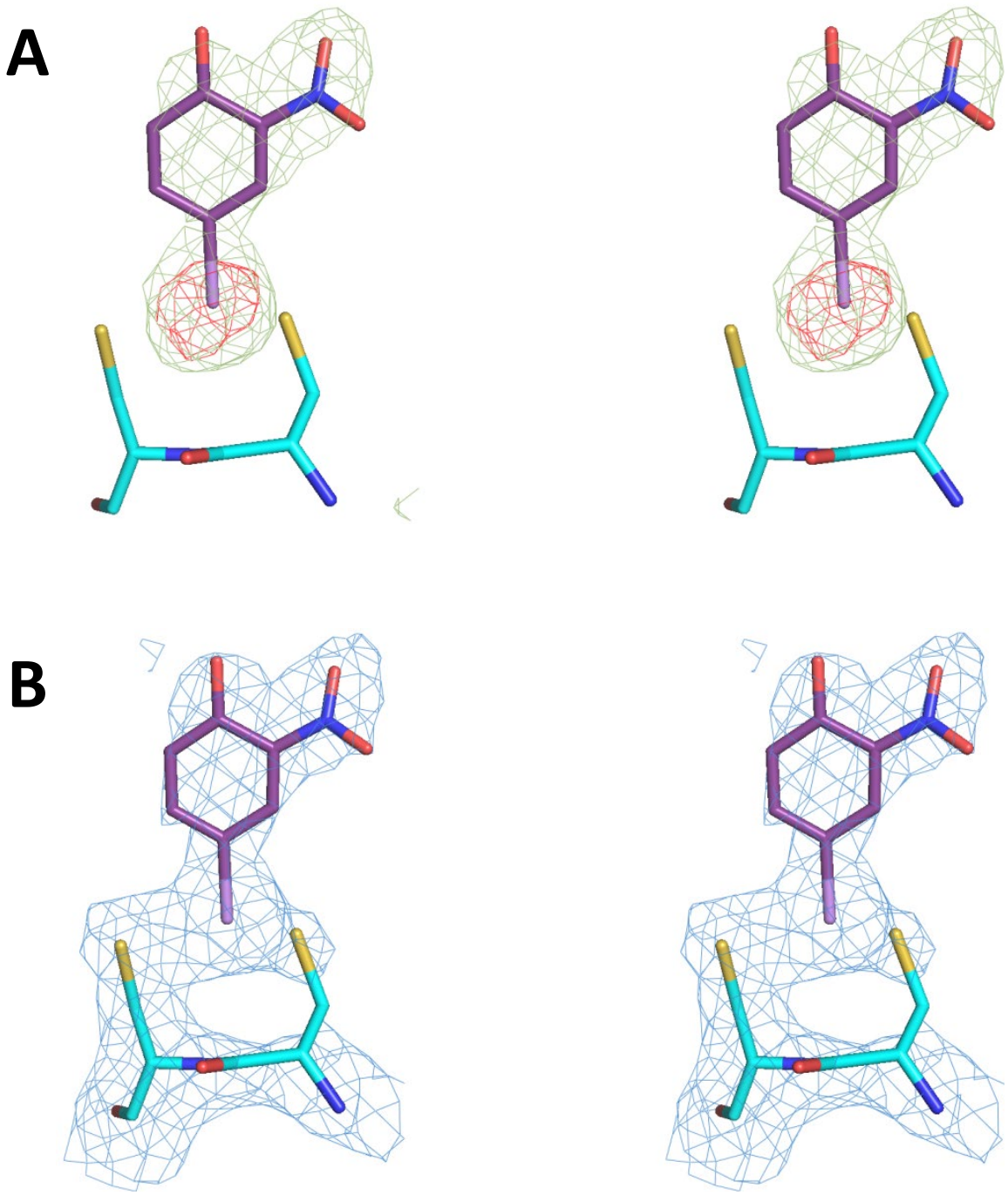
Running title: Catalytic mechanism of the ArsI C-As lyase

Abbreviations: As(III), arsenite; DMAs(III), dimethylarsenite; Fe(II), ferrous iron; ITC, isothermal titration calorimetry; MAs(III), methylarsenite; MAs(V), methylarsenate; MSMA, monosodium methylarsenate; Ni(II), nickel; Rox(V), roxarsone (4–hydroxyl-3-nitrophenylarsenate)

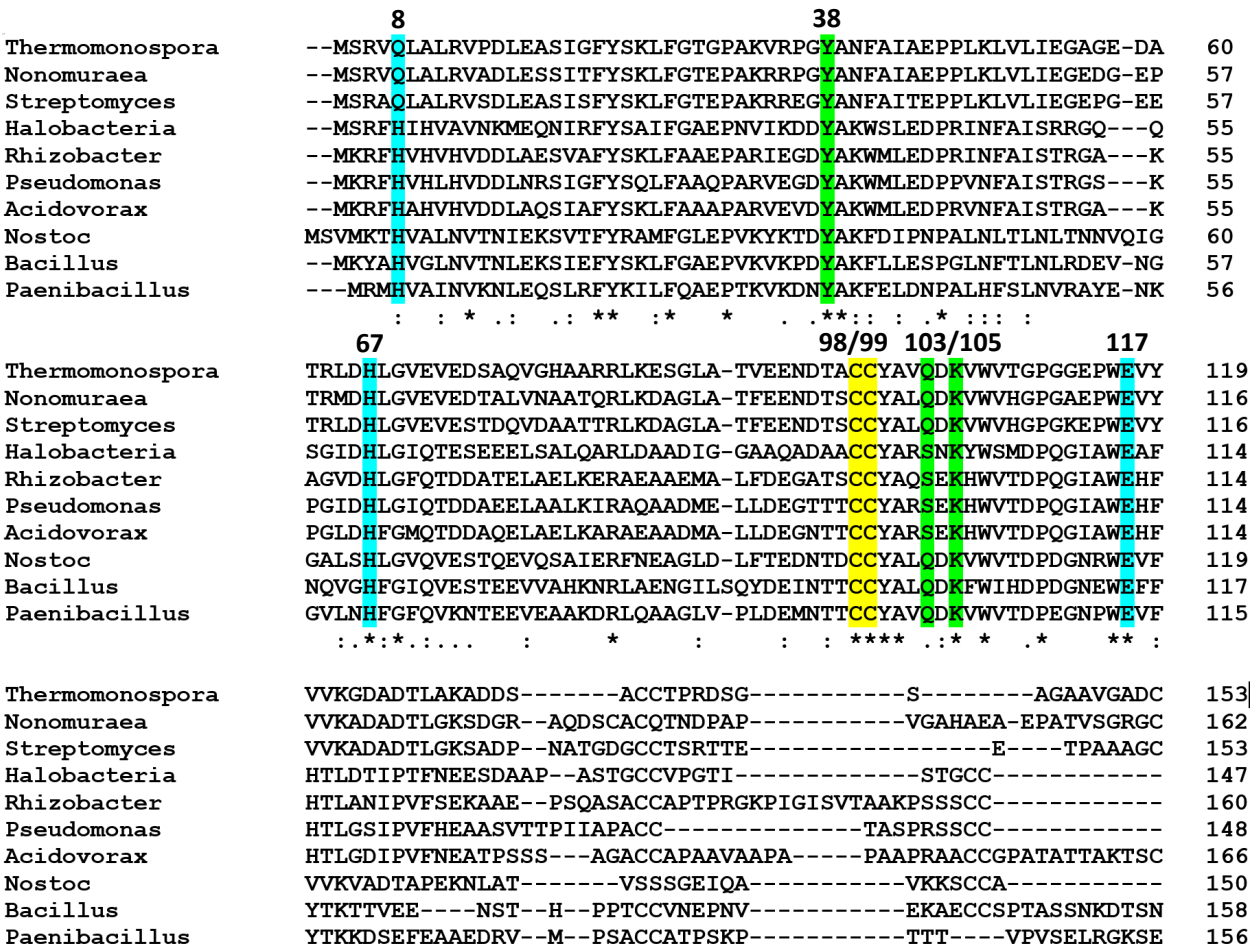
Supplemental information

Supplementary Table 1. Oligonucleotide primers for mutagenesis

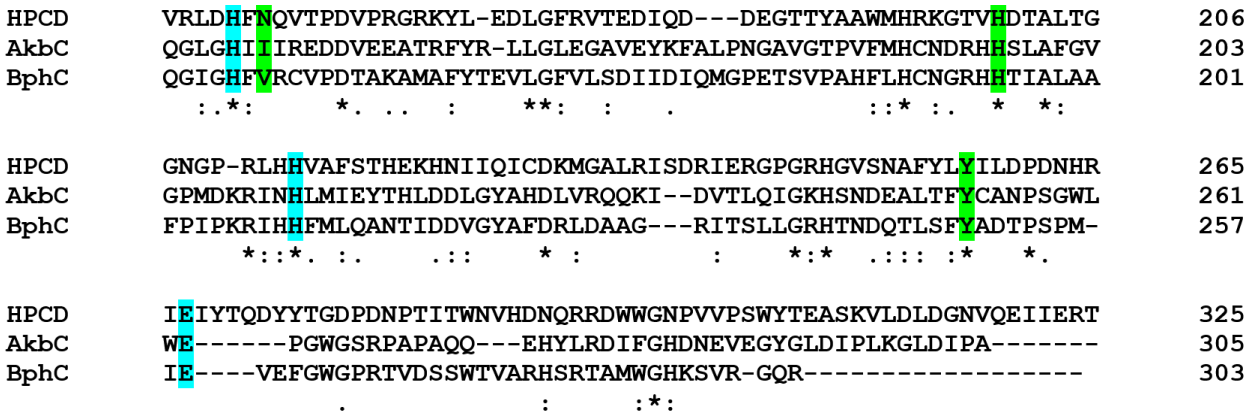
Primer	**Sequence (5'-3')**	**Result**
TcArsI_K105A Fw	AC GCC GTC CAG GAC GCG GTG TGG GTC ACC GG	Lys105 changed to Ala in wild type background
TcArsI_K105A Rv	CC GGT GAC CCA CAC CGC GTC CTG GAC GGC GT	
TcArsI_K105E Fw	AC GCC GTC CAG GAC GAG GTG TGG GTC ACC GG	Lys105 changed to Glu in wild type background
TcArsI_K105E Rv	CC GGT GAC CCA CAC CTC GTC CTG GAC GGC GT	
TcArsI_K105R Fw	AC GCC GTC CAG GAC CGG GTG TGG GTC ACC GG	Lys105 changed to Arg in wild type background
TcArsI_K105R Rv	CC GGT GAC CCA CAC CCG GTC CTG GAC GGC GT	
TcArsI_Q103A Fw	GC TGC TAC GCC GTC GCG GAC AAG GTG TGG GT	Gln103 changed to Ala in wild type background
TcArsI_Q105A Rv	AC CCA CAC CTT GTC CGC GAC GGC G TA GCA GC	
TcArsI_Q103H Fw	GC TGC TAC GCC GTC CAC GAC AAG GTG TGG GT	Gln103 changed to His in wild type background
TcArsI_Q105H Rv	AC CCA CAC CTT GTC GTG GAC GGC GTA GCA GC	
TcArsI_Y38F Fw	AG GTC CGG CCC GGC TTC GCC AAC TTC GCC AT	Tyr38 changed to Phe in wild type background
TcArsI_Y38F Rv	AT GGC GAA GTT GGC GAA GCC GGG CCG GAC CT	



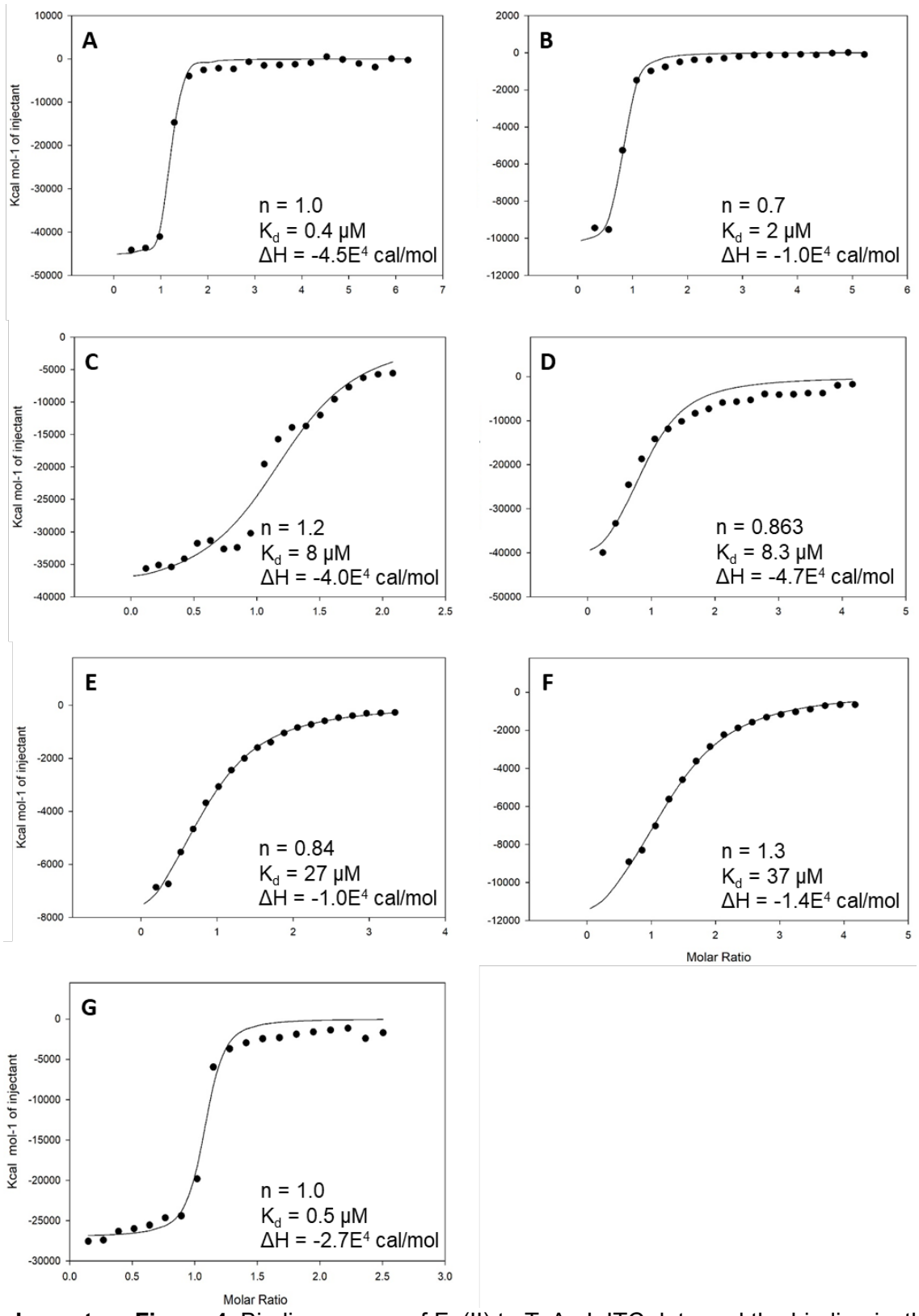
Supplementary Figure 1. Stereo view of the electron density of Rox(III) in the TcArsI-K105A structure. A. The Fo-Fc map (green) and anomalous map (red) contoured at 2.5σ and 3.5σ level, respectively, confirming the presence of Rox(III). **B.** 2Fo-Fc map (blue) contoured at 1.0σ, showing that Rox(III) well fitted in electron density.



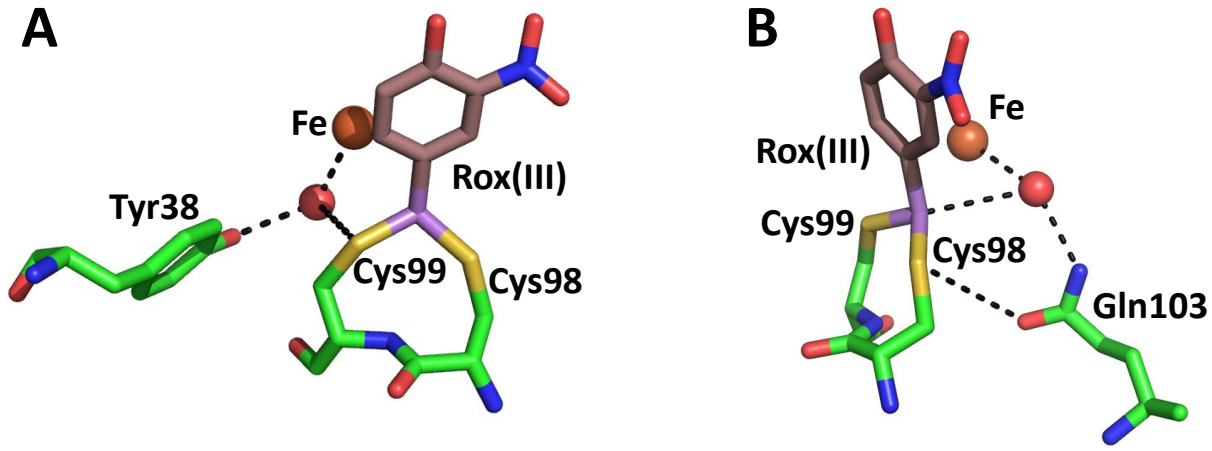
Supplementary Figure 2. Multiple alignment of ArsI orthologs. The protein sequence of ArsI from *Thermomonospora curvata* DSM43183 (ACY99683.1) is compared with ArsI orthologs from four Gram-positive bacteria [*Nonomuraea jiangxiensis* (WP_090937645.1), *Streptomyces* sp. SID8499 (WP_164387993.1), *Paenibacillus* sp. VKM B-2647 (WP_041052430.1) and *Bacillus* sp. MD1 (AIA09488), three Gram-negative bacteria, *Pseudomonas aeruginosa* (ACD39059.1), *Acidovorax* sp. ROOT402 (WP_056066914.1), *Rhizobacter* sp. Root1221 (WP_056655130.1), one cyanobacterium [*Nostoc* sp. PCC 7120 (BAB73061.1)] and one archaeon [*Halobacteria archaeon* (NNG12772.1)]. GenBank accession numbers in parentheses. The Fe(II)-binding residues are shaded in cyan. The organoarsenical binding residues are shaded in yellow. The residues mutated in this study are shaded in light green.



Supplementary Figure 3. Multiple alignment of type I extradiol ring-cleaving dioxygenases (C-terminal domains only). The protein sequences of three representative type I extradiol ring-cleaving dioxygenases [homoprotocatechuate 2,3-dioxygenase encoded by *hpcd* from *Brevibacterium fuscum* (3OJT_A), methylcatechol 2,3-dioxygenase encoded by *akbC* from *Rhodococcus* sp. strain DK17 (AAR90133.1) and 2,3-dihydroxybiphenyl 1,2-dioxygenase isozyme encoded by *bphC* from *Pseudomonas* sp. strain KKS102 (AAA25750.1) are aligned and compared to each other. The Fe(II)-binding residues and the second coordination sphere residues are highlighted in cyan and light green, respectively. GenBank accession numbers in parentheses.



Supplementary Figure 4. Binding assays of Fe(II) to TcArsI. ITC data and the binding isotherm of TcArsI with Fe(II) binding. Data were collected at 37 °C with 20 injections at 10 min intervals: **A**, wild type TcArsI; **B**, Y38F; **C**, Q103A; **D**, Q103H; **E** K105A; **F**, K105E; **G**, K105R.



Supplementary Figure 5. Water mediated interaction of Tyr38 and Gln103 with metal. Tyr38 (**A**) and Gln103 (**B**) interact with the Fe atom through the water molecules (red) near to Cys99 and Cys98, respectively.