HHS Public Access

Author manuscript

J Mol Biol. Author manuscript; available in PMC 2017 June 05.

Published in final edited form as:

J Mol Biol. 2016 June 5; 428(11): 2462–2473. doi:10.1016/j.jmb.2016.04.022.

Structure of the ArsI C–As Lyase: Insights into the Mechanism of Degradation of Organoarsenical Herbicides and Growth Promoters

Venkadesh Sarkarai Nadar[1], **Masafumi Yoshinaga**[1,2], **Shashank S. Pawitwar**[1], **Palani Kandavelu**[3], **Banumathi Sankaran**[4], and **Barry P. Rosen**[1]

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

[2]Department of Occupational and Environmental Health, Nagoya University Graduate School of Medicine, Nagoya, Aichi 466-8550, Japan

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

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

Abstract

Arsenic is a ubiquitous and carcinogenic environmental element that enters the biosphere primarily from geochemical sources, but also through anthropogenic activities. Microorganisms play an important role in the arsenic biogeochemical cycle by biotransformation of inorganic arsenic into organic arsenicals and vice versa. ArsI is a microbial non-heme, ferrous-dependent dioxygenase that transforms toxic methylarsenite [MAs(III)] to less toxic and carcinogenic inorganic arsenite [As(III)] by C–As bond cleavage. An ArsI ortholog, TcArsI, from the thermophilic bacterium *Thermomonospora curvata* was expressed, purified, and crystallized. The structure was solved in both the apo form and with Ni(II), Co(II), or Fe(III). The MAs(III) binding site is a vicinal cysteine pair in a flexible loop. A structure with the loop occupied with β-mercaptoethanol mimics binding of MAs(III). The structure of a mutant protein (Y100H/V102F) was solved in two different crystal forms with two other orientations of the flexible loop. These results suggest that a loop-gating mechanism controls the catalytic reaction. In the ligand-free open state, the loop is exposed to solvent, where it can bind MAs(III). The loop moves toward the active site, where it forms a closed state that orients the C–As bond for dioxygen addition and cleavage. Elucidation of the enzymatic mechanism of this unprecedented C–As lyase reaction will enhance our understanding of recycling of environmental organoarsenicals.

Correspondence to Barry P. Rosen: *brosen@fiu.edu.*

Accession numbers

5C4P, 5C6X, 5D4F, 5C68, 5CB9, 5DRH, 5DFG, 5HCW.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2016.04.022.

Keywords

C–As lyase; MSMA; roxarsone; organoarsenical herbicides; organoarsenical growth promoters

Introduction

Arsenic ranks first on the Agency for Toxic Substances and Disease Registry's and U.S. Environmental Protection Agency's (EPA) Superfund List of Hazardous Substances[†]. Arsenic compounds are highly toxic and carcinogenic substances that enter the biosphere from geochemical sources such as volcanic emissions and from anthropogenic sources such as the organoarsenical MSMA [monosodium methylarsenate or MAs(V)], which is still in use as an herbicide for golf courses [1] and on agricultural crops such as cotton, selective post-emergent weed control in cotton, highway rights-of-way, and sod farms [2]. Aromatic organoarsenicals such as roxarsone [3-nitro-4-hydroxyphenylarsonic acid or Rox(V)] are used worldwide as antimicrobial growth promoters in animal husbandry [3]. Microorganisms play an important role in biogeochemical cycling of arsenic by biotransformation of inorganic arsenic into organic arsenic species and degradation of organoarsenicals back to inorganic arsenic. For example, prokaryotic and eukaryotic microbes methylate arsenite [As(III)] by ArsM [As(III) *S*-adenosylmethionine methyltransferease] to the highly toxic trivalent organoarsenicals MAs(III) and dimethylarsenite [DMAs(III)][4,5]. To detoxify MAs(III), microbes demethylate it to inorganic As(III), which increases arsenic mobility, polluting water supplies. Demethylating microbes have been isolated from soil and water [6–9].

However, until recently, no enzyme able to cleave the carbon–arsenic bond had been identified. We identified a pathway for MSMA degradation isolated from Florida golf course soil that has two steps catalyzed by microbial communities; the first step is the reduction of MAs(V) to MAs(III), and the second step is demethylation of MAs(III) to As(III) [9]. Recently, we identified the gene responsible for MAs(III) demethylation step in a *Bacillus* isolated from soil [10]. ArsI is a non-heme, ferrous-dependent extradiol dioxygenase with C–As lyase activity. In extradiol Fe(II)-dependent dioxygenases, the role of Fe(II) is to activate O_2, which can then form a C–O bond, cleaving C–C bonds [11], or, in the case of ArsI, the C–As bond [10]. ArsI degrades both MAs(III) and aromatic organoarsenicals such as Rox(III). There is no precedence for cleavage of the carbon–arsenic bond, so the enzymatic pathway of this novel enzyme is of considerable interest. Knowledge of the three-dimensional structure of this novel enzyme is necessary to understand its catalytic mechanism. Here, we describe the structure of TcArsI (Gene bank accession no. ACY99683.1) from the thermophilic gram-positive bacterium *Thermomonospora curvata* DSM 43183. The conserved triad of Gln8, His65, and Glu117 form the metal binding site. TcArsI is proposed to have a novel binding site for MAs(III) formed by the vicinal Cys98–Cys99 pair, which is in a flexible loop. The loop could be resolved by formation of a mixed disulfide between Cys99 in the loop and β-mercaptoethanol (βME). A Y100H/V102F double mutant was constructed to restrict the mobility of the loop, and its structure was

[†]http://www.atsdr.cdc.gov/SPL/index.html

solved in two crystal forms, each with a different orientation of the loop. Thus, the MAs(III) binding loop is observed in three orientations in three different structures. The results suggest that the movement of this distinctive substrate binding loop controls access of MAs(III) to dioxygen bound in the Fe(II) site.

Results

Overall structure of the ArsI C–As lyase

TcArsI from the moderate thermophile *T. curvata* DSM 43183 was purified and crystallized, as described previously [12]. The purified enzyme exhibited robust activity from 37° to 55 °C (Fig. 1S). The structure of TcArsI co-crystallized with Ni(II) was solved by Ni-single-wavelength anomalous diffraction (Ni-SAD) (PDB 5C4P) phasing. The C–As lyase has an α/β fold composed of two motifs, each of which has an α-helix and four antiparallel β-strands arranged in a βαβββ cupin fold (Fig. 1 and 2S) [13]. The N-terminal segment of the enzyme starts from the β1 strand (Arg6 to Val13), and residues from Leu16 to Phe27 form an α1 helix. Strand β2 follows the α1 helix and forms a βαβ motif. Strand β3 is antiparallel to β2 and is connected by a β hairpin loop that includes Pro36 and Gly37. Strand β4 is located between strands β1 and β3 and is parallel to β1 and antiparallel to β3. Similarly, the second motif starts from strand β5 and follows helix α2 and strand β6. Strand β7 is antiparallel to β6, which is connected by a long hairpin loop containing the substrate binding site (Glu93 to Asp104). Motifs 1 and 2 are connected by a psi loop (Gly55-Thr61). The resulting eight β strands create a β sheet, and two helices are placed at two edges of the sheet. The C-terminal region ends with strand β8, which is parallel to β5 and antiparallel to β7. N-terminal strand β1 and C-terminal β8 are adjacent to each other and are located at the middle of the enzyme core. Motifs 1 and 2 are structurally similar except for the length of the β hairpin loops. The β sheet forms a convexo-concave core similar to other cupin fold structures (Fig. 2). The convex side is mostly occupied by hydrophobic side chains, whereas the concave side includes hydrophilic side chains. As described in more detail below, metal binding residues Gln8, His65, and Glu117 are in strands β1, β5, and β8, respectively, and MAs(III) binding residues Cys98 and Cys99 are in the β hairpin loop that connects β6 and β7. The crystallographic asymmetric unit consists of one molecule. The twofold symmetry generates a crystallographic dimer (Fig. 3S). The hydrophobic residues on the convex side of the β sheet interact with those on the symmetry-related molecule. Each monomer has a single active site, with no contribution of residues from the other monomer, consistent with the observation that TcArsI is a monomer in solution [12]. A search for similar structures using the Dali server [14] identified the closest structural homolog as the non-heme, Fe(II)-dependent catechol 2,3-dioxygenase LapB from *Pseudomonas* sp. KL28 [15]. LapB has two subunits with four βαβββ motifs, but TcArsI has only a single subunit. The crystal packing of TcArsI is similar to LapB, which does not share the active site residues with other molecules in the asymmetric unit.

The metal binding site

ArsI has been shown to be a Fe(II)-dependent dioxygenase [10]. To obtain the TcArsI–Ni(SAD) structure, the protein was co-crystallized with Ni(II). The TcArsI–Ni(II) structure was solved to 1.97 Å resolution with the metal binding site filled by Ni(II) with full

occupancy. TcArsI was also co-crystallized with Co(II) (PDB ID 5C6X) or Fe(III) (PDB ID 5D4F) in the conserved metal binding site. The three metal-bound structures were superimposable, and only the Ni(II) structure is described in detail.

A high-resolution structure with partial Ni(II) occupancy (TcArsI–HR; PDB ID 5C68) was obtained with crystals to which no Ni(II) had been added. That structure was solved to 1.46 Å resolution with 0.37 occupancy with Ni(II). The TcArsI–HR structure is superimposable on the TcArsI–Ni(II) structure. We assume that the Ni(II) came from the Ni-NTA column that was used for purification of the enzyme. Ni(II) is ligated with Gln8, His67, and Glu117, which were predicted to comprise inner coordination sphere of the metal binding site [10] (Fig. 3a). Ni(II) is bound with a six-coordinate octahedral geometry, with three ligands coming from the triad of the three conserved residues and three from water molecules. This geometry, with an average distance between Ni(II) and the six ligands of 2.2 Å, is typical for members of the dioxygenase superfamily [13]. Lys105 in strand β7 interacts directly with the inner coordination sphere. The metal coordination geometry is like other dioxygenase enzymes except that most have two histidine residues in the triad. Some ArsI orthologs have a histidine residue in place of Gln8 (Fig. 4S). It is not clear whether this difference in the composition of the triad has functional consequences.

The substrate binding site

MAs(III) and trivalent roxarsone [Rox (III)] are ArsI substrates that are predicted to bind to a conserved vicinal cysteine pair (Cys98 and Cys99 in TcArsI) (Fig. 4S) [10]. To demonstrate that TcArsI binds those substrates, we employed a fluorescent assay for substrate binding. We have employed intrinsic tryptophan fluorescence to examine substrate binding in arsenic detoxification proteins [16–20]. TcArsI has two tryptophan residues, Trp107 and Trp116 (Fig. 4S), which produce intrinsic fluorescence that is quenched by addition of MAs(III) or Rox(III), but not As(III), DMAs(V) or Rox(V) (Fig. 5S). These results indicate that TcArsI binds MAs(III) and Rox(III). However, neither co-crystallization nor soaking crystals with either substrate produced crystals with bound substrate visible in the electron density, nor was the substrate binding loop visible.

As an alternative, TcArsI was reacted with βME to form a mixed disulfide adduct with the active site cysteines. Even though no Ni(II) was added, the TcArsI–βME structure (PDB ID 5CB9) had Ni(II) in the metal site with 0.26 occupancy. Again, the Ni(II) likely came from the Ni-NTA column that was used for purification of the enzyme. Like the TcArsI–HR structure, in the TcArsI–βME structure, Lys105 in strand β7 also interacts directly with the inner coordination sphere (Fig. 3b). Unlike the TcArsI–HR structure, in the TcArsI–βME structure, the substrate binding loop with bound βME is visible in electron density, with a mixed disulfide formed between the thiols of βME and Cys99 (Fig. 4). Gln103 and βME both interact directly with the inner coordination sphere. In both the TcArsI–HR and TcArsI–βME structures, Tyr38, located in strand β3, interacts with the metal center through the outer coordination sphere. Lys105 and Tyr38 are conserved in other ArsI orthologs, while Gln103 is conserved in some but is a serine in others. The hydroxyl oxygen atom of βME interacts with Ni(II) in the metal binding site via a water molecule. The two carbon atoms of βME face hydrophobic residues Val51, Ile53, and the aromatic ring of Tyr38 in

motif1. The side chain of Tyr100 in the substrate binding loop is facing out from the core of the enzyme. The distance between the hydroxyl group of Tyr38 and the sulfur atom Cys99 is 4.8 Å. The sulfur–sulfur distance between Cys98 and Cys99 is 3.5 Å, consistent with a two-coordinate binding site for trivalent organoarsenicals [21]. While we presume that the substrate binding loop is too flexible to be visible in the TcArsI–HR structure, interaction of βME with the metal binding site and other residues reduces its mobility.

Ligand-free structure

Apparently, TcArsI is able to abstract Ni(II) from the Ni-NTA resin, so there was always a partial occupancy of the metal binding site with Ni(II). To obtain a metal-free structure, we treated the enzyme with EDTA to remove bound Ni(II). The metal-free enzyme was crystallized, and its structure was resolved to 2.27 Å (PDB ID 5DRH). Residues from Asn94 to Asp104 are not visible in the electron density map, consistent with substantial flexibility of the substrate binding loop. The ligand-free structure was superimposed on the TcArsI–HR structure. Their overall conformations are similar, but there are some differences. The orientations of metal-binding residues Gln8 and His65 do not change, but the side chain of Glu117 is flipped out (approximately 4.1 Å) from the metal center (Fig. 5a). Tyr38 has two conformations in the apo structure, in which the orientation of the hydroxyl group of Tyr38 is reoriented nearly 120° different. In one conformation, the hydroxyl group faces the metal binding residues as in the Ni(II)-bound structure, and, in the other conformation, the hydroxyl group faces away from the metal center. The positions of strands β6 and β7 in motif 2 differ between the two structures (Fig. 5b). Strand β6 in the apo structure is shifted approximately 2 Å behind that in the TcArsI–HR structure, which reduces the concaveness of the cupin fold in the apo structure (Fig. 6S). The side chain of Lys105 is displaced away from the metal center. The side chain of Trp107 in strand β7 is reoriented nearly 180° relative to its position in the Ni(II)-bound structure. In TcArsI–Ni structure, the aromatic side chain of Trp107 faces toward the metal, and the ring nitrogen makes a hydrogen bond with Glu93. In the apo structure, reorientation of Trp107 prevents formation of this bond with Glu93. This movement of Trp107 may be responsible for the changes in intrinsic protein fluorescence induced by substrate binding. Residues Tyr38, Lys105, and Trp107 are conserved in ArsI orthologs (Fig. 4S). Their conformational variations suggest that these three residues play a role in catalysis. The results demonstrate that cupin fold is more compact when metal is bound, which brings residues involved in interaction of the inner and outer coordination spheres closer to metal center.

Structure of the Y100H/V102F double mutant

The substrate binding loop is not visible in the electron density in the apo or Ni(II)-bound structures. In the βME-bound structure, the loop has a higher temperature factor than the rest of the protein, indicating greater mobility. In an effort to restrict the mobility of the loop in the crystal, a Y100H/V102F double mutant was constructed. In some ArsI orthologs, a hydrophilic histidine residue is in the position corresponding to the aromatic Tyr100 residue (Fig. 4S). We predicted that changing Tyr100 to histidine would increase interactions with bound metal or water molecules in the core region. We also predicted that replacing the side chain of Val102 with an aromatic ring might stabilize the substrate binding loop. The

TcArsI–Y100H/V102F enzyme was treated with EDTA to remove Ni(II) derived from the Ni-NTA resin and crystallized in two morphologies, diamond-shaped and plate-like crystals.

The diamond-shaped crystals diffracted to 1.95 Å in a tetragonal space group with cell dimensions similar to the crystals of wild-type enzyme. The structure (PDB ID 5DFG) is most similar to the apo–TcArsI structure, with the metal binding site unoccupied. The substrate binding loop is visible in the electron density map, except for residues Ala97, Cys98, and Cys99. However, relative to strand β7, the orientation of the loop is moved nearly 180° away from the position of the loop in the TcArsI–βME structure (Fig. 7S). The loop is nearly entirely solvent exposed and has no strong interactions with other residues or with residues in symmetry-related molecules. Again, the temperature factor of the loop residues is higher than the rest of the protein, consistent with a highly mobile loop.

The plate-like crystals diffracted poorly (2.78 Å). The crystals belong to the monoclinic $P2_1$ space group, with cell dimensions a = 42.9, b = 79.7, c = 70.9 Å, and β = 101.1°. The Mathews coefficient of 2.43 $Å^3D^{-1}$ indicates that there are four molecules in the crystallographic asymmetric unit [22]. The crystal packing also differs from the tetragonal form. In this structure (PDB ID 5HCW), there is similarly no metal in metal binding site. In molecule A, two metal-binding residues (His65 and Glu117) and Lys105 interact with water molecules (Fig. 6a). The side chain of Glu117 has a conformation intermediate between that of the apo and TcArsI–HR structures (Fig. 6b). In molecule B, the entire substrate binding loop is visible. In molecule D, only a portion of the loop is visible, and the loop is not visible in either molecule A or C. In molecule B, the loop intrudes into the core region of the symmetry-related molecule (Fig. 8S), and there are multiple interactions between the loop residues and residues in the symmetry-related molecule (Fig. 6c and d). The side chains of Glu93 and Asp64 interact with backbone atoms of Thr96 and Phe102, respectively (Fig. 6c). The imidazole ring of His100 stacks with the phenyl ring of Tyr38 (Fig. 6d). Thr96 interacts with residues in the metal-binding site of the symmetry-related molecule via a water molecule (Fig. 6a). Cys98 and Cys99, which comprise the MAs(III) binding site, are free from interactions. The electron density is not sufficiently resolved to determine if there is a disulfide bond between the two cysteine residues. Even though there is no bound metal, the orientation of the side chain of Trp105 is similar to that in the TcArsI–HR structure. In the TcArsI–βME structure, strand β6 continues to Glu93, and strand β7 begins with Lys105, while in this structure, strand β6 continues to Asp95, and strand β7 begins with Gln103. The orientation of the loop is midway between that observed in TcArsI–βME and tetragonal TcArsI–Y100H/V102F structures. Thus, the substrate binding loop is observed in three different orientations in three different structures, one near the metal-binding site, one distant from the metal binding site, and one intermediate between those two. This mobility suggests that the movement of the loop may be involved in presentation of the organoarsenical substrate to metal-bound dioxygen.

Discussion

ArsI is a novel C–As lyase, the only known enzyme that can cleave a carbon–arsenic bond. It is a member of the vicinal oxygen chelate superfamily, and, like other vicinal oxygen chelate family enzymes, it has a cupin fold [13]. However, it has a unique substrate binding

site that is quite different from those in other dioxygenases. The site is a vicinal cysteine pair, Cys98–Cys99, which binds the substrate MAs(III) and releases the product As(III). Even though vicinal cysteine pairs can bind As(III), the affinity for As(III) can be several orders of magnitude less than for MAs(III) [23]. The MAs(III) binding site is within a flexible loop consisting of residues from 93 to 104. Attempts to obtain structures with bound MAs(III) or other trivalent organoarsenical substrates were unsuccessful. The loop could be visualized by formation of a mixed disulfide with βME. In that structure, the loop was stabilized by interaction of the βME hydroxyl with a water of the metal binding site. This suggests that the loop is highly mobile but is able to bring its substrate in the vicinity of the metal site to allow for addition of dioxygen to the C–As bond. To restrict mobility of the loop, the double Y100H/V102F was constructed. The mutant enzyme crystallized in two forms, with the loop visible in two different positions. In one structure, the loop is flipped out into solvent. In the other structure, it is located in an intermediate position. Thus, the loop can move from solvent, where substrate could be bound, to the active site, where catalysis occurs and back to the medium for product release (Fig. 7).

The nature of the substrate binding site differentiates TcArsI from the closest structural homolog, the LapB 3-methylcatechol dioxygenase [15]. In LapB, the substrate binds directly with the metal through its hydroxyl group by displacing water molecules of the inner sphere. The aromatic ring and methyl group of 3-methylcatechol is surrounded by mostly aromatic and hydrophobic side chains, which facilitate substrate binding. These residues are not conserved in TcArsI, where the substrate is directly bound to the vicinal cysteine pair. When MAs(III) is bound, its hydroxyl groups are displaced by the sulfurs, and the methyl group decreases its polarity so that the substrate is unable to interact directly with the metal. TcArsI has two hydrophobic residues, Val51 and Ile53, and one aromatic residue Tyr38, which might interact with the methyl group, facilitating its movement toward the inner coordination sphere. Conserved residue Tyr38 has two conformations that may play a role in catalysis. Hypothetically, the side chain of Asn40 could orient Tyr38 toward the catalytic center, where Tyr38 deprotonation of the methyl group of MAs(III) would facilitate attack by dioxygen. A similar mechanism of deprotonation of the methyl group of 3-methylcatechol by a tyrosine has been proposed for LapB [15]. The deprotonated methyl group would be attacked by O_2, resulting in the formation of an alkylperoxo intermediate. Thus, the mechanism of bond cleavage of TcArsI and LapB may be similar, but it is the unique nature of the MAs(III) binding site that makes TcArsI a C–As lyase. We propose that this loop movement provides a gating mechanism for the enzyme (Fig. 8). In this model, βME is replaced with MAs(III), and the loop is moved toward the metal center, where it is available for reaction with dioxygen. Our model is similar to the loop gating that regulates substrate entry into the active site of dimethylsulfoniopropionate lyase [24].

Materials and Methods

Protein purification and crystallization

The TcArsI derivative used in this study lacks 25 nonessential residues from the C terminus [12]. TcArsI and the Y100H/V102F double mutant were purified, as described previously [12]. The enzymes were crystallized by a hanging drop vapor diffusion method at 293 K

temperature [25]. The hanging drop contained of 2 μl of protein at approximately 15 mg ml^{-1} mixed with 2 μl of a reservoir solution consisting of 0.2 M sodium acetate, 0.1 M Tris–HCl (pH 8.5), and 30% polyethylene glycol 4000 [12]. To obtain crystals with bound Ni(II), 1 μl of 10 mM $NiCl_2$ was mixed with protein directly in the hanging drop. The diamond-shaped, chunky crystals were harvested for data collection. Co-crystallation with Co(II) or Fe(III) was performed similarly. For TcArsI–βME, crystals were grown by the addition of 10 mM βME to the hanging drop. TcArsI-Apo was obtained by exchanging with a buffer consisting of 50 mM Mops, 0.5 M NaCl, and 3 mM tris(2-carboxyethyl)phosphine) containing 5 mM EDTA to remove Ni(II). Buffer exchange was performed by 10-fold dilution and concentration with a 10 kDa cutoff Amicon Ultra Centrifuge Filter (Billerica, Massachusetts). The apoenzyme was crystallized as described above.

The double Y100H/V102F mutant was generated by site-directed mutagenesis using a Quick Change mutagenesis kit (Stratagene, La Jolla, CA). The forward and reverse oligonucleotide primers used for mutagenesis were: forward 5′-GCCTGCTGCCACGCCTTCCAGGACAAGGTG-3′ and reverse 5′-CACCTTGTCCTGGAAGGCGTGGCAGCAGGC-3′ (the mutagenic changes are underlined). Purified Y100H/V102F was crystallized by the hanging drop vapor diffusion method. The crystals were obtained with 0.1 M sodium acetate, 0.2 M ammonium sulfate, and 30% polyethylene glycol 4000. Crystals with two different morphologies were obtained in a single drop, chunky diamond-shaped crystals and thin rectangular plate-like crystals. Both crystals were used for data collection.

Assay of TcArsI activity

C–As lyase activity with purified TcArsI was assayed in buffer consisting of 0.1 M Mops and 0.15 mM KCl (pH 7.0), as described [10]. TcArsI (1 μM) was incubated with 3 mM tris(2-carboxyethyl)phosphine, 1 mM cysteine, and 0.1 mM $FeSO_4$, and reaction was initiated by addition of 5 μM MAs(III). After incubation with shaking at 200 rpm for 2 h at the indicated temperatures, the reactions were terminated by addition of EDTA. Arsenic speciation was determined by high pressure liquid chromatography (PerkinElmer Series 2000) with a C18 300 A reverse-phase column (Chromservis s.r.o., Brno, Czech Republic), with arsenic measured by inductively coupled plasma mass spectrometry using an ELAN 9000 (Perkin Elmer, Waltham, MA).

Data collection and structure determination

The TcArsI–Ni(II) structure was solved by the single anomalous dispersion method [26] using Ni(II) as an anomalous scatterer. Single anomalous dispersion data were collected at a wavelength of 1.479 Å [near the Ni(II) X-ray absorption peak] on beamline 22-ID at the Advanced Photon Source (APS), Argonne Laboratory. The crystal belonged to space group $P4_32_12$ with cell dimensions a = b = 42.8 and c = 118.2 Å and diffracted to 1.97 Å. The data were indexed and scaled using HKL-2000 [27]. The data indexing details are shown in Table 1. Heavy atom data preparation, location, and phasing were performed using SHELXC, D, and E, respectively [28]. The phase was further improved by density modification using the DM program [29] in the CCP4 suite [30]. The model was built using automated model building program ARP/wARP program [31] in the CCP4 suit with an R-factor of 27.4%.

The model and electron density map were visualized using COOT software [32], and the structure was refined using REFMAC5 [33]. The Ni-SAD structure was used as a template for molecular replacement with the other data sets. Molecular replacement was performed using the PHASER program [34] in the CCP4 suite. Crystals of the TcArsI–HR diffracted to 1.46 Å, and complete data sets were collected at a wavelength of 1.0088 Å on beamline 22ID at APS. In the structure, a single Ni atom was present in the metal binding site with partial occupancy (0.37).

The X-ray data of the apoenzyme and TcArsi–βME crystals were collected at a wavelength 0.9774 Å on beamline 5.0.3 of the Advanced Light Source at the Lawrence Berkeley National Laboratory. In TcArsi–βME structure, the metal binding site was occupied by a Ni atom with an occupancy of 0.26. The presence of Ni atom and its occupancy in TcArsI–Ni, TcArsI–HR, and TcArsI–βME structures were confirmed by calculating the omit (Fo-Fc) and anomalous maps of those structures (Fig. 9S). These crystals also belong to the $P4_32_12$ space group and diffracted with resolutions of 2.27 and 1.95 Å, respectively. The diamond-shaped and plate-like crystals of the TcArsI–Y100H/V102F double mutant diffracted in two different space groups. The diamond-shape crystals diffracted to 1.97 Å and belong to the $P4_32_12$ space group. The plate-like crystals diffracted to 2.78 Å and belong to the $P2_1$ space group. With all structures, the simulated annealing refinements were performed by the phenix.refine program in PHENIX [35]. All data sets and coordination are deposited to the protein data bank [36]. Molecular models were drawn with PYMOL [37].

Assay of substrate-dependent quenching of intrinsic protein fluorescence

Protein fluorescence quenching experiments were performed using a Quanta Master UV-Vis QM-4 spectrofluorometer (Photon Technology International, Birmingham, NJ) at 23 °C. Samples were excited at 295 nm, with emission at 340 nm in a buffer consisting of 50 mM Mops, 0.5 M NaCl (pH 7.5) and 3 mM tris(2-carboxyethyl)phosphine), and 5 μM TcArsI in a final volume of 1.8 ml. Arsenicals were added with stirring at the indicated concentrations. Fluorescence intensities were corrected for volume changes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grant R37 GM55425 to B.P.R. and by a pilot project award to M.Y. from Florida International University Herbert Wertheim College of Medicine. M.Y. was supported by a Grant-in-Aid for Research Activity Start-up 15H06274 from the Japan Society for the Promotion of Science (JSPS). This project utilized the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline of the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, 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 Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences of the US Department of Energy under Contract No. DE-AC02-05CH11231.

Abbreviations used

MSMA	monosodium methylarsenate
MAs(V)	methylarsenate
Rox	roxarsone (4-hydroxy-3-nitrophenylarsonic acid)
As(III)	inorganic arsenite
As(V)	inorganic arsenate
MAs(III)	methylarsenite
βME	β-mercaptoethanol
APS	advanced photon source

References

[1]. Chen Z, Cai Y, Liu G, Solo-Gabriele H, Snyder GH, Cisar JL. Role of soil-derived dissolved substances in arsenic transport and transformation in laboratory experiments. Sci. Total Environ. 2008; 406:180–189. [PubMed: 18760447]

[2]. Bednar AJ, Garbarino JR, Ranville JF, Wildeman TR. Preserving the distribution of inorganic arsenic species in groundwater and acid mine drainage samples. Environ. Sci. Technol. 2002; 36:2213–2218. [PubMed: 12038832]

[3]. Garbarino JR, Bednar AJ, Rutherford DW, Beyer RS, Wershaw RL. Environmental fate of roxarsone in poultry litter. I. Degradation of roxarsone during composting. Environ. Sci. Technol. 2003; 37:1509–1514. [PubMed: 12731831]

[4]. Thomas, DJ.; Rosen, BP. Arsenic methyltransferases. In: Kretsinger, RH.; Uversky, VN.; Permyakov, EA., editors. Encyclopedia of Metalloproteins. Springer New York; New York: 2013. p. 138-43.

[5]. Zhu YG, Yoshinaga M, Zhao FJ, Rosen BP. Earth abides arsenic biotransformations. Annu. Rev. Earth Planet. Sci. 2014; 42:443–467. [PubMed: 26778863]

[6]. Sanders JG. Microbial role in the demethylation and oxidation of methylated arsenicals in seawater. Chemosphere. 1979:135–7.

[7]. Lehr C, Polishchuk E, Radoja U, Cullen WR. Demethylation of methylarsenic species by Mycobacterium neoaurum. Appl. Organomet. Chem. 2003; 17:831–834.

[8]. Huang JH, Scherr F, Matzner E. Demethylation of dimethylarsinic acid and arsenobetaine in different organic soils. Water Air Soil Pollut. 2007; 182:31–41.

[9]. Yoshinaga M, Cai Y, Rosen BP. Demethylation of methylarsonic acid by a microbial community. Environ. Microbiol. 2011; 13:1205–1215. [PubMed: 21272184]

[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. [PubMed: 24821808]

[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. [PubMed: 18249197]

[12]. 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 Crystallogr. F Struct. Biol. Commun. 2014; 70:761–764. [PubMed: 24915088]

[13]. Fetzner S. Ring-cleaving dioxygenases with a cupin fold. Appl. Environ. Microbiol. 2012; 78:2505–2514. [PubMed: 22287012]

[14]. Holm L, Kaariainen S, Rosenstrom P, Schenkel A. Searching protein structure databases with DaliLite v.3. Bioinformatics. 2008; 24:2780–2781. [PubMed: 18818215]

[15]. Cho JH, Jung DK, Lee K, Rhee S. Crystal structure and functional analysis of the extradiol dioxygenase LapB from a long-chain alkylphenol degradation pathway in pseudomonas. J. Biol. Chem. 2009; 284:34,321–34,330.

[16]. Zhou T, Liu S, Rosen BP. Interaction of substrate and effector binding sites in the ArsA ATPase. Biochemistry. 1995; 34:13,622–13,626. [PubMed: 7819188]

[17]. Chen Y, Rosen BP. Metalloregulatory properties of the ArsD repressor. J. Biol. Chem. 1997; 272:14,257–14,262.

[18]. Liu J, Rosen BP. Ligand interactions of the ArsC arsenate reductase. J. Biol. Chem. 1997; 272:21,084–21,089.

[19]. Marapakala K, Qin J, Rosen BP. Identification of catalytic residues in the as(III) S-adenosylmethionine methyltransferase. Biochemistry. 2012; 51:944–951. [PubMed: 22257120]

[20]. Dheeman DS, Packianathan C, Pillai JK, Rosen BP. Pathway of human AS3MT arsenic methylation. Chem. Res. Toxicol. 2014; 27:1979–1989. [PubMed: 25325836]

[21]. Shi W, Wu J, Rosen BP. Identification of a putative metal binding site in a new family of metalloregulatory proteins. J. Biol. Chem. 1994; 269:19,826–19,829.

[22]. Matthews BW. Solvent content of protein crystals. J. Mol. Biol. 1968; 33:491–497. [PubMed: 5700707]

[23]. Chen J, Zhu YG, Rosen BP. A novel biosensor selective for organoarsenicals. Appl. Environ. Microbiol. 2012; 78:7145–7147. [PubMed: 22843528]

[24]. Li CY, Wei TD, Zhang SH, Chen XL, Gao X, Wang P, et al. Molecular insight into bacterial cleavage of oceanic dimethylsulfoniopropionate into dimethyl sulfide. Proc. Natl. Acad. Sci. U. S. A. 2014; 111:1026–1031. [PubMed: 24395783]

[25]. McPherson A. Current approaches to macromolecular crystallization. Eur. J. Biochem. 1990; 189:1–23. [PubMed: 2185018]

[26]. Wang BC. Resolution of phase ambiguity in macromolecular crystallography. Methods Enzymol. 1985; 115:90–112. [PubMed: 4079800]

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

[28]. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. 2002; 58:1772–1779. [PubMed: 12351820]

[29]. Cowtan K. Error estimation and bias correction in phase-improvement calculations. Acta Crystallogr. D Biol. Crystallogr. 1999; 55:1555–1567. [PubMed: 10489450]

[30]. 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. [PubMed: 21460441]

[31]. Evrard GX, Langer GG, Perrakis A, Lamzin VS. Assessment of automatic ligand building in ARP/wARP. Acta Crystallogr. D Biol. Crystallogr. 2007; 63:108–117. [PubMed: 17164533]

[32]. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]

[33]. 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. [PubMed: 15572771]

[34]. McCoy AJ. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. 2007; 63:32–41. [PubMed: 17164524]

[35]. 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. [PubMed: 20124702]

[36]. Deshpande N, Addess KJ, Bluhm WF, Merino-Ott JC, Townsend-Merino W, Zhang Q, et al. The RCSB protein data Bank: a redesigned query system and relational database based on the mmCIF schema. Nucleic Acids Res. 2005; 33:D233–D237. [PubMed: 15608185]

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

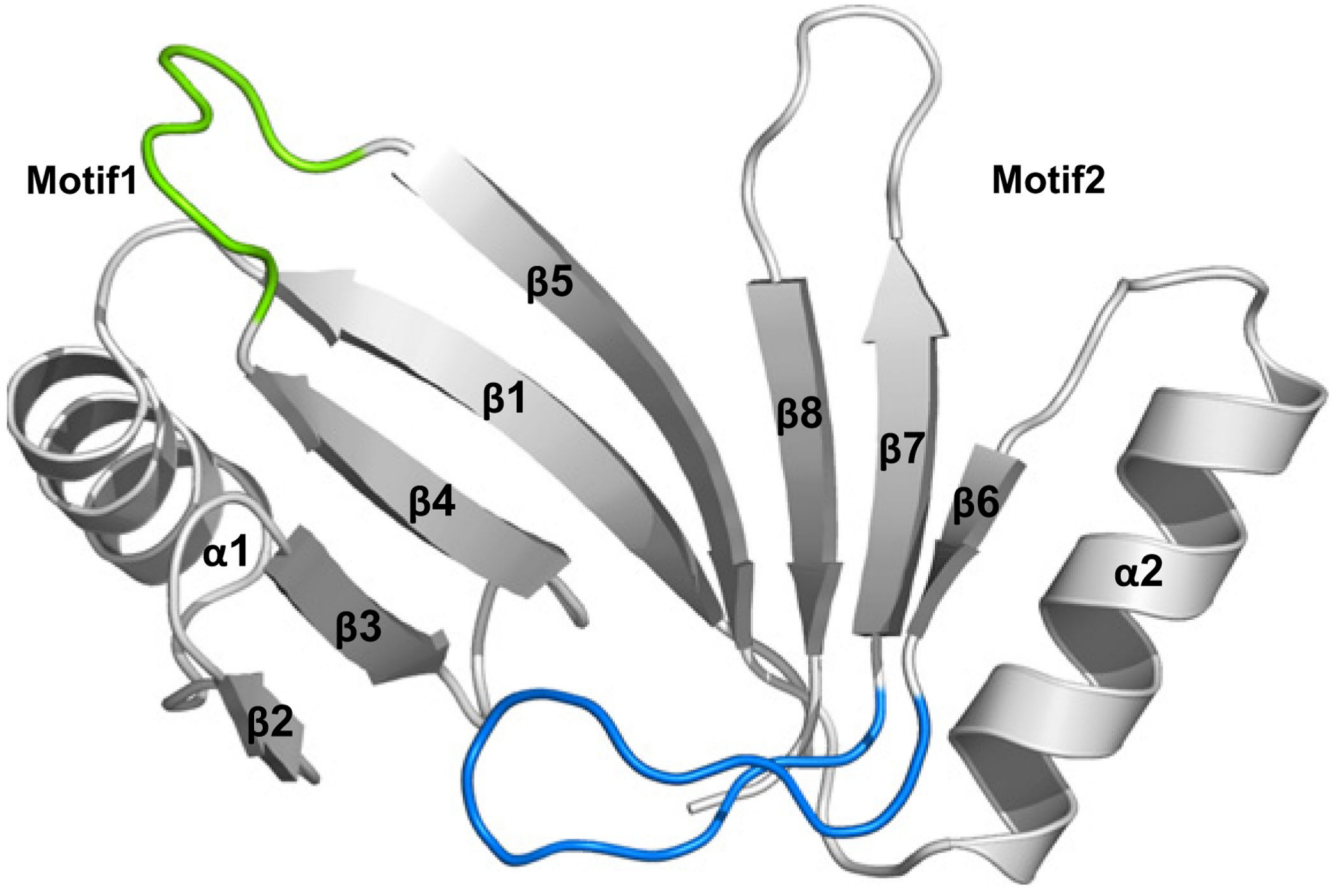

Fig. 1.

Overall structure of TcArsI. The structure has an α/β fold with two domains. Each domain having a βαβββ fold is connected by a loop (residues from 53 to 65) (green). The resulting eight β-strands form the core of the enzyme, curving to create a cleft that orients the active site. The loop that includes the substrate binding site is shown in blue.

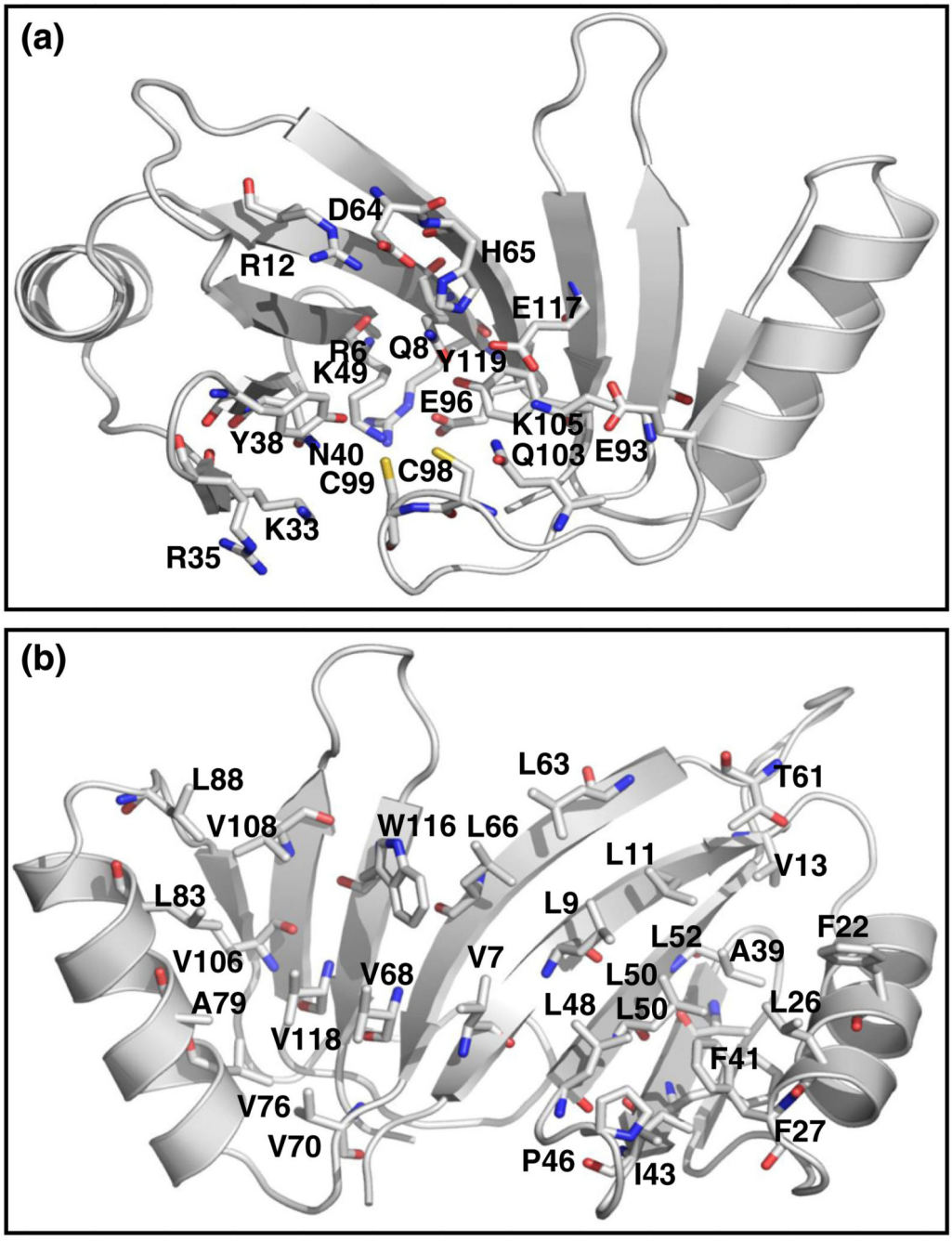

Fig. 2.

The β sheet of TcArsI forms a convexo-concave core. (a) The concave side includes residues with hydrophilic side chains. (b) The residues in the convex side have primarily hydrophobic side chains.

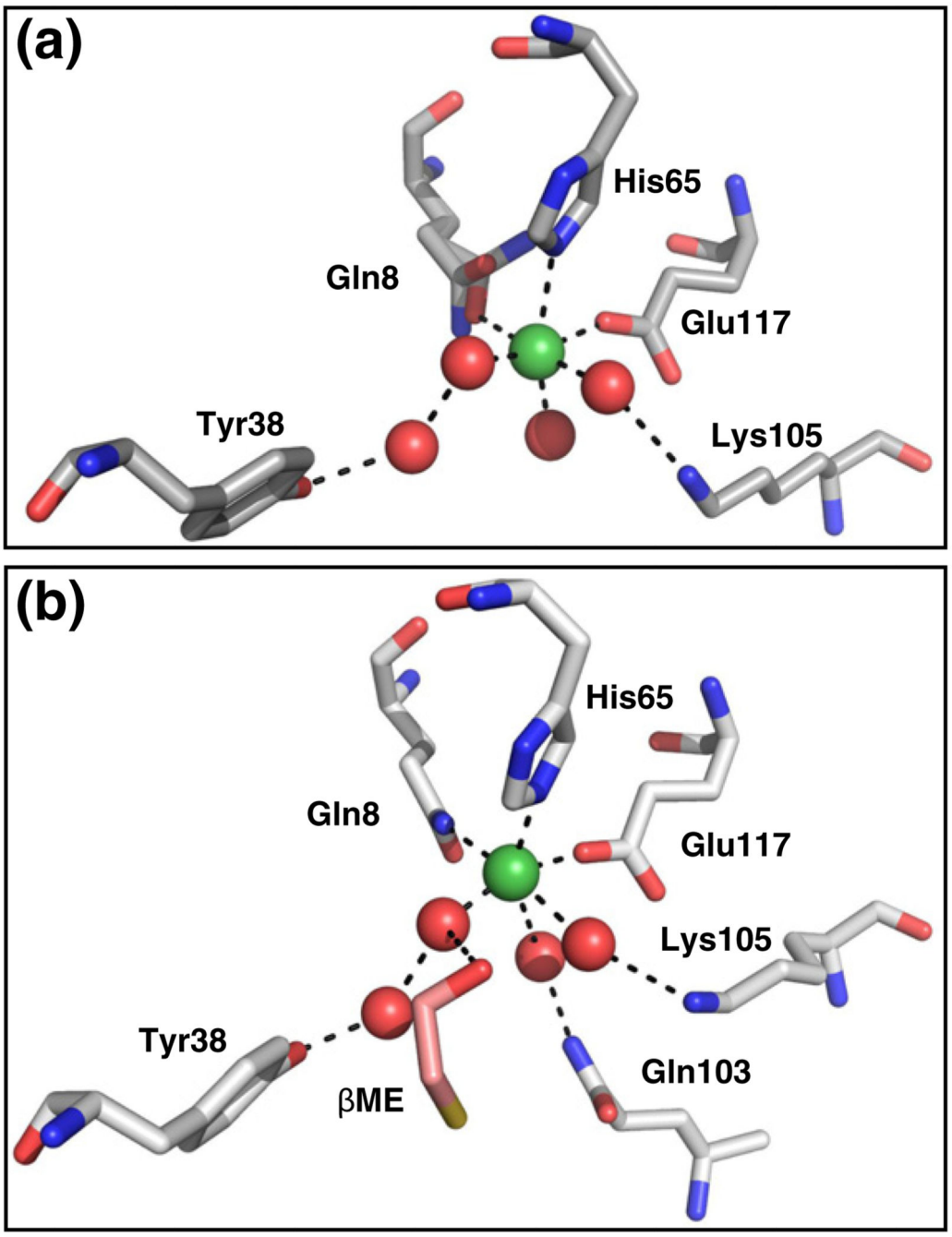

Fig. 3.

The metal binding site. The binding site is filled with Ni(II) (green sphere) in (a) TcArsI–HR and (b) TcArsI–βME. Water molecules are represented by red spheres. Lys105, Gln103, and βME interact with Ni(II) via water molecules of the inner coordination sphere. Tyr38 interacts with the inner coordination sphere via a water molecule in the outer coordination sphere.

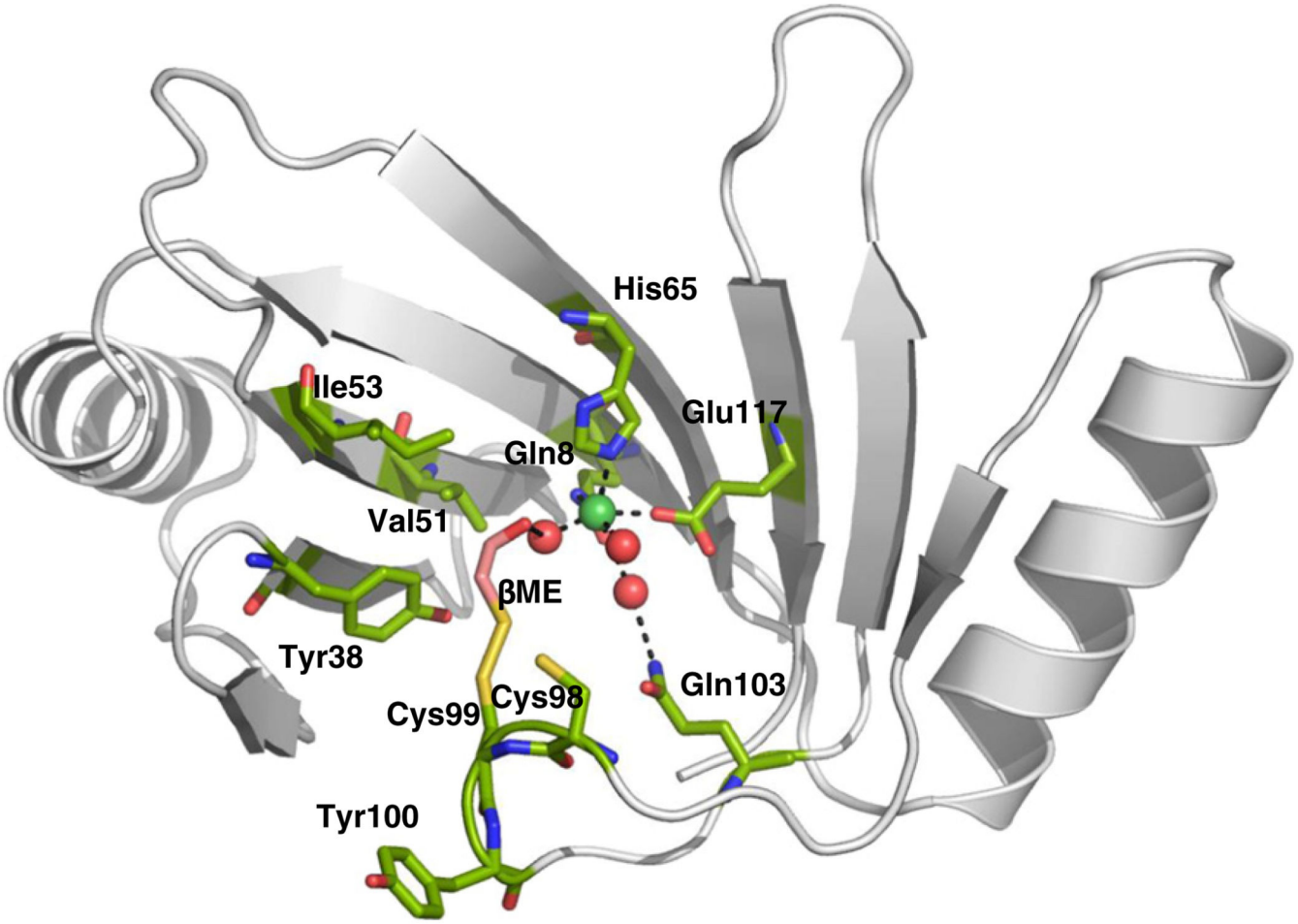

Fig. 4.

The substrate binding site. The substrate binding includes the vicinal pair Cys98–Cys99 [10]. In the TcArsI–βME structure, βME makes a mixed disulfide with the sulfur atom of Cys99. The oxygen atom of βME interacts with Ni(II) via a water molecule in the inner coordination sphere. The carbon atoms of βME face hydrophobic residues Val51, Ile53, and the aromatic ring of Tyr38.

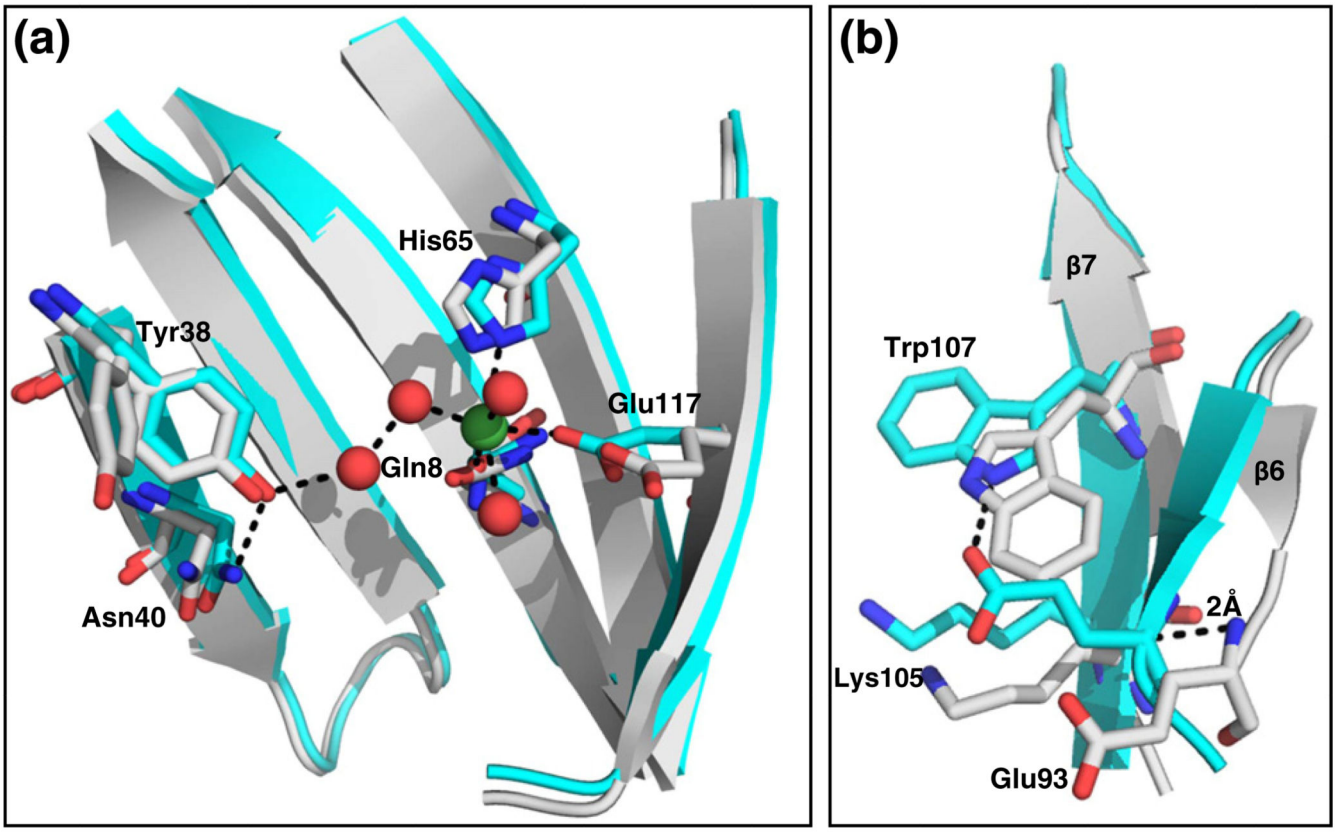

Fig. 5.

Superimposition of the apo TcArsI structure (gray) with the TcArsI–HR structure (cyan). (a) In motif 1, Tyr38 has two conformations in the apo structure. The dotted line shows a hydrogen bond between the hydroxyl and a water of the outer coordination sphere, connecting through a hydrogen bond network to Ni(II) (green sphere). Water molecules are shown as in red spheres. (b) In motif 2, Trp107 is oriented nearly 180° differently in the two structures. The orientation in the TcArsI–HR structure allows for a hydrogen bond between Trp107 and Glu93. In the apo structure, strand β6 is displaced by 2 Å, preventing formation of the hydrogen bond.

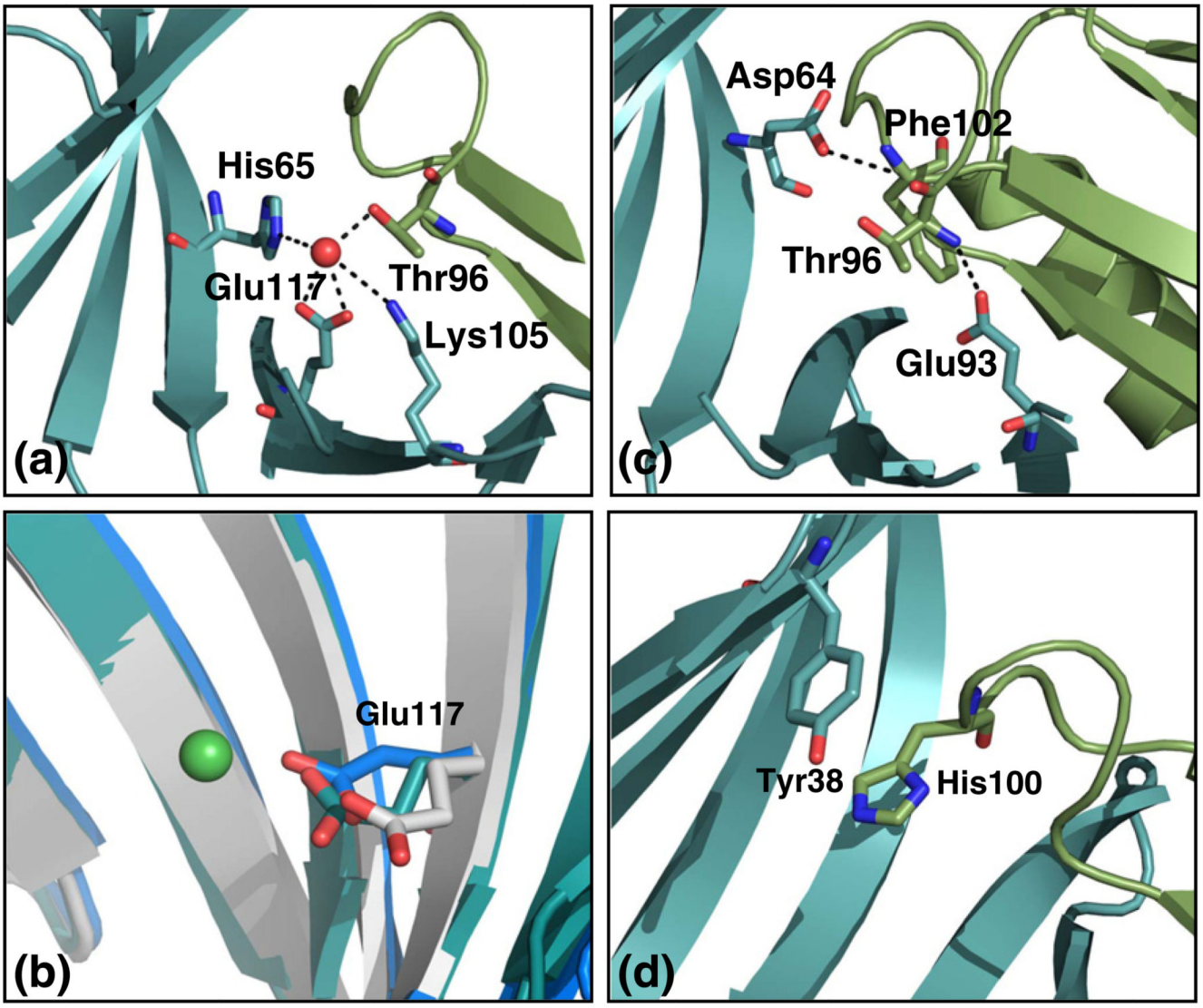

Fig. 6.

Symmetry-related molecules in the Y100H/V102F structures. (a) In the monoclinic TcArsI–Y100F/V102F, a water molecule occupies the metal binding site. In addition to the His65 and Glu117 of the metal binding site, Lys105 from the same molecule and Thr96 from a symmetry-related molecule (green) interact with the water molecule. (b) Superimposition of Glu117 of the monoclinic TcArsI–Y100H/V102F (cyan) with TcArsI–apo (gray) and TcArsI–HR (blue) shows the variations in the orientation of this Ni (green sphere) ligand. (c and d) The substrate binding loop of monoclinic TcArsI–Y100F/V102F interacts with a symmetry-related molecule. The entire substrate binding loop is resolved in molecule B by stabilization with the symmetry-related molecule and by stacking of the phenyl ring of Tyr38 in one molecule with the imidozole ring of His100 in the symmetry-related molecule.

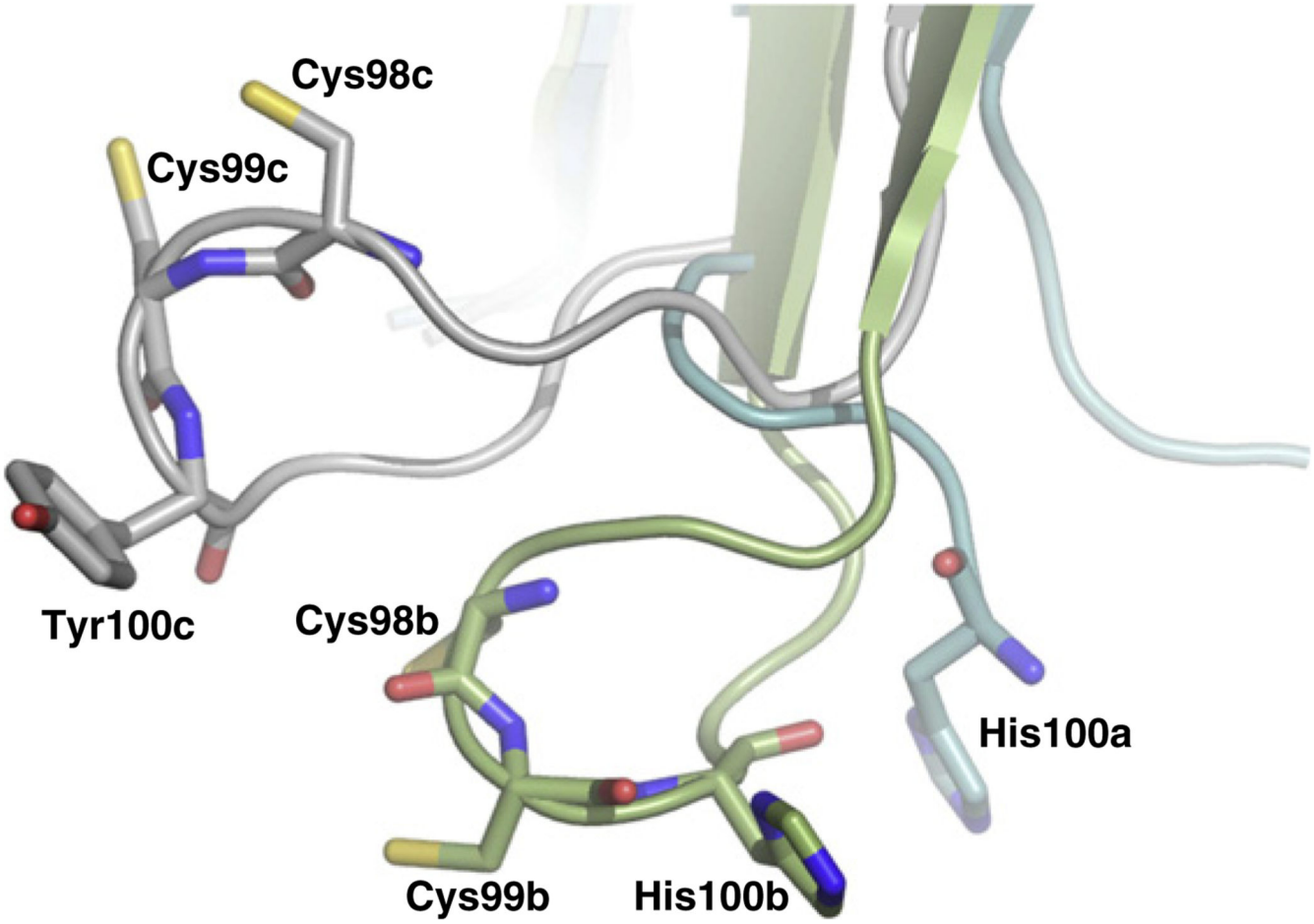

Fig. 7.

Three orientations of the substrate binding loop. The substrate binding loops of the TcArsI–βME (gray), monoclinic TcArsI–Y100H/V102F (green), and tetragonal TcArsI–Y100H/V102F (cyan) are superimposed.

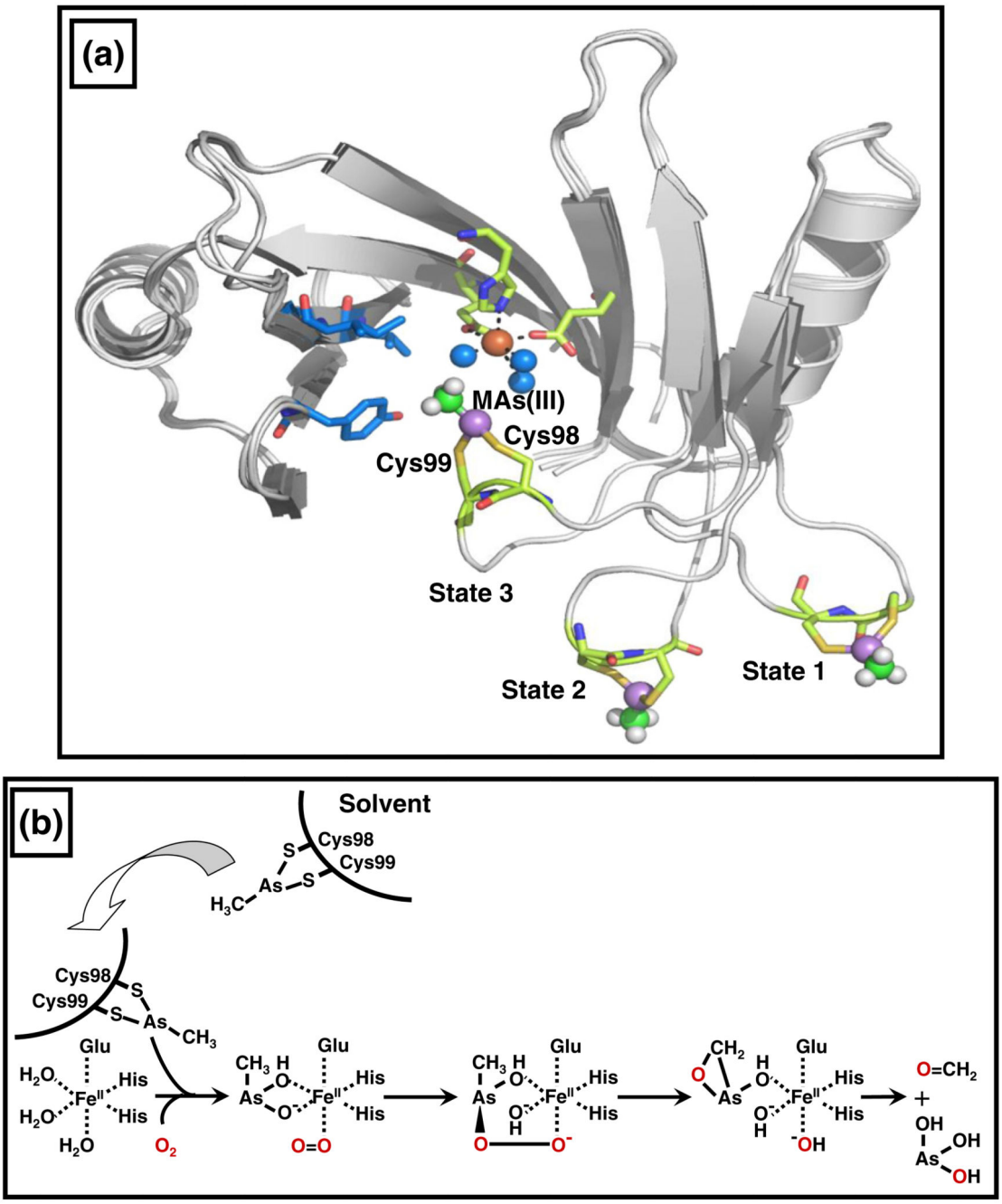

Fig. 8.

Proposed loop-gating mechanism of C–As bond cleavage by ArsI. (a) Loop gating by movement of the loop with bound MAs(III) in three different states. States 1 and 2 are modeled from the tetragonal and monoclinic forms of TcArsI–Y100H/V102F structure, respectively. State 3 was constructed by replacement of the βME molecule with MAs(III) bound to the sulfurs of Cys98 and Cys99. The models of the three states were superimposed. In State 1, the methyl group of MAs(III) has no interaction with residue in ArsI, and the core region is in a solvent accessible open conformation. In State 3, the methyl group of MAs(III)

is oriented toward a hydrophobic face of enzyme (blue sticks), and C–As bond is oriented toward the catalytic center for addition of dioxygen. In this state, the active site is in a solvent inaccessible closed conformation. State 2 is intermediate between the open and close conformations. MAs(III) is in ball and stick, Fe(II) is shown as a brown sphere, and water molecules are shown as blue spheres. (b) Mechanism of C–As bond cleavage. MAs(III) in solvent binds to Cys98 and Cys98 in a loop that repositions to orient the As atom into the active site. Fe(II) is bound in a octahedral geometry with six coordinations, three to the Gln8–His65–Glu117 triad and three with solvent. MAs(III) transfers from the sulfurs of the loop to the active site by replacing the two solvent molecules. This activates the metal center for binding of dioxygen. Binding of O_2 (red) to Fe(II) allows electron transfer, converting dioxygen into a superoxide anion, which attacks MAs(III) to form an alkylperoxo-Fe(II) intermediate, cleaving the C–As bond. The products have one atom of oxygen incorporated into formaldehyde and the other in $As(OH)_3$. Adapted from Ref. [13].

J Mol Biol. Author manuscript; available in PMC 2017 June 05.

Table 1

Data collection, indexing, and refinement statistics. Values in parenthesis are highest resolution shell

	TcArsI–Ni(SAD)	TcArsI–HR	TcArsI–βME	TcArsI–Apo	TcArsI–Y100H/V102F (Tetragonal)	TcArsI–Y100H/V102F (monoclinic)
Source	APS 22-ID	APS 22-ID	ALS Beamline 5.0.3	ALS Beamline 5.0.3	APS 22-ID	APS 22-ID
Wavelength (Å)	1.4789	1.0088	0.9774	0.9774	1.0000	1.0000
Resolution (Å)	50.00–1.97 (2.04–1.97)	50.00–1.46 (1.51–1.46)	42.66–1.95 (2.06–1.95)	44.12–2.27 (2.39–2.27)	50.00–1.97 (2.04–1.97)	50.00–2.78 (2.88–2.78)
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P2_1$
Unit cell	a = b = 42.8, c = 118.2 Å	a = b = 42.2, c = 118.5 Å	a = b = 42.7, c = 117.0 Å	a = b = 44.1, c = 117.6 Å	a = b = 44.2, c = 118.6 Å	a = 42.9, b = 79.7, c = 70.9 Å, β = 101.1°
Unique reflections	8156	19,428	8500	5898	8832	11,380
Multiplicity	20.6 (12.1)	25.3 (12.0)	15.2 (16.1)	12.8 (13.6)	25.1 (27.0)	3.3 (2.4)
Completeness (%)	97.3 (78.9)	99.6 (96.1)	100.0 (100.)	100.0 (100.0)	98.7 (100.0)	96.0 (77.7)
Mean I/sigma(I)	27.5 (5.2)	54.0 (5.3)	27.0 (6.8)	24.1 (8.9)	40.3 (13.6)	10.9 (2.0)
R-merge (%)	7.6 (29.7)	5.0 (44.0)	6.1 (40.0)	7.9 (41.3)	6.7 (26.9)	9.8 (41.2)
Wilson B-factor ($Å^2$)	35.6	16.8	31.7	35.3	23.6	46.1
R-work (%)	20.5 (31.8)	18.7 (19.0)	20.7 (22.0)	22.1 (28.1)	20.2 (24.2)	19.4 (24.5)
R-free (%)	26.4 (35.4)	21.3 (21.5)	26.8 (30.0)	26.7 (31.3)	24.4 (33.0)	28.2 (37.5)
No. of atoms	887	959	1005	850	965	3318
Macromolecules	837	882	901	812	904	3267
Ligands	1	2	6	1	0	0
Water	49	75	98	37	61	51
RMS (bonds) (Å)	0.019	0.005	0.008	0.003	0.005	0.010
RMS (angles) (°)	2.07	1.10	1.23	0.83	1.10	1.42
Ramachandran favored (%)	98.0	98.0	93.0	96.0	98.0	88.0
Ramachandran allowed (%)	2.0	2.0	7.0	4.0	2.0	10
Ramachandran outliers (%)	0.0	0.0	0.0	0.0	0.0	2.0
Average B-factor ($Å^2$)	38.6	20.0	38.6	35.4	28.8	54.2
Macromolecules	37.8	18.9	37.4	35.2	28.1	54.4
Ligands	40.9	18.1	58.3	35.6	0.0	0.0
Water	52.6	33.6	49.8	38.4	39.0	46.8
PDB ID	5C4P	5C68	5CB9	5DRH	5DFG	5HCW