Published in final edited form as:
Biochemistry. 2013 November 5; 52(44): . doi:10.1021/bi400841g.

Structure of a *Clostridium botulinum* C143S thiaminase I/thiamin complex reveals active site architecture†,,‡

Megan D. Sikowitz[§], **Brateen Shome**[‖], **Yang Zhang**[§], **Tadhg P. Begley**[‖,*], and **Steven E. Ealick**[§,*]

[§]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853

[‖]Department of Chemistry, Texas A&M University, College Station, Texas 77843

Abstract

Thiaminases are responsible for the degradation of thiamin and its metabolites. Two classes of thiaminases have been identified based on their three-dimensional structures and in their requirements for a nucleophilic second substrate. While the reactions of several thiaminases have been characterized, the physiological role of thiamin degradation is not fully understood. We have determined the three-dimensional X-ray structure of an inactive C143S mutant of *Clostridium botulinum* (Cb) thiaminase I with bound thiamin at 2.2 Å resolution. The C143S/thiamin complex provides atomic level details of the orientation of thiamin upon binding to Cb-thiaminase I and the identity of active site residues involved in substrate binding and catalysis. The specific roles of active site residues were probed using site directed mutagenesis and kinetic analyses, leading to a detailed mechanism for Cb-thiaminase I. The structure of Cb-thiaminase I is also compared to the functionally similar but structurally distinct thiaminase II.

Thiamin (vitamin B_1) is an essential vitamin in all living organisms. Thiamin diphosphate (ThDP) is a cofactor in many biological processes including carbohydrate metabolism and amino acid biosynthesis. Most plants, fungi, and bacteria can synthesize thiamin, but animals must obtain it from their diets. Numerous investigations over the past sixty years provide a nearly complete understanding of thiamin biosynthesis,[1] yet the physiological basis for its degradation and the fate of the breakdown products is largely unknown. Thiaminases catalyze the degradation of thiamin into its thiazole and pyrimidine components (Figure 1).[2–10] These enzymes are found in a wide variety of organisms, including plants, fish, and bacteria. Consequently, ingestion of thiaminase containing foods by humans or other animals can cause symptoms of thiamin deficiency.[11, 12] In humans, this deficiency can affect the cardiovascular system (wet beriberi) or nervous system (dry beriberi) with potentially fatal consequences.[13, 14]

†This work was supported by NIH grants DK67081 (to SEE), DK44083 (to TPB) and T32GM008500, and by the Robert A. Welch Foundation (A-0034 to TPB). This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award GM103403 from the NIH. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

‡The coordinates of the C143S/thiamin complex have been deposited in the Protein Data Bank under accession code 4KYS.

*To whom correspondence should be addressed at the Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853. Telephone: (607) 255–7961. Fax: (607) 255-1227.begley@mail.chem.tamu.edu, see3@cornell.edu.

SUPPORTING INFORMATION

The experimental method for determining the pH rate profile of native Cb-thiaminase I is provided in the supplementary material. Supplementary Figure 1 is the pH rate profile for the thiaminase I reaction. Supplementary Figure 2 is a multiple sequence alignment that shows the conserved residues in thiaminase I. This material is available free of charge via the Internet at http://pubs.acs.org.

Biochemical analyses have established two distinct classes of thiaminases. Thiaminase I utilizes a variety of nucleophiles,[15, 16] whereas thiaminase II exclusively utilizes water as the nucleophile.[17] Thiaminase II[17] is distinct from thiaminase I in both structure and sequence.[18] A crystal structure of the thiaminase I from *Bacillus thiaminolyticus* (Bt-thiaminase I) has been previously reported.[18] Bt-thiaminase I is a monomer composed of two domains joined by three crossover segments and is structurally homologous to the group II periplasmic binding proteins (PBPs).[18, 19] PBPs bind small molecules and deliver them to ABC transporters for eventual uptake into the cytoplasm. Structural studies reveal group II PBPs in an open conformation when no ligand is bound and in a closed conformation when a ligand is bound.[19] Bt-thiaminase I is a rare example of an enzyme in the group II PBP superfamily. Bt-thiaminase I is structurally similar to TbpA,[20] which is the PBP in many prokaryotes for thiamin, thiamin phosphate (ThMP), and ThDP. Interestingly, THI5, the thiamin pyrimidine synthase in eukaryotes, is also a member of the group II PBP superfamily. [21, 22] THI5 is structurally homologous to ThiY,[21] which is the PBP in some prokaryotes for the thiamin degradation product *N*-formyl-4-amino-5-(aminomethyl)-2-methylpyrimidine.[23]

Previous biochemical and structural studies established that an active site cysteine residue is involved in thiaminase I catalysis. [15, 16, 18] A crystal structure of Bt-thiaminase I with the mechanism-based, irreversible inhibitor 4-amino-6-chloro-2, 5-dimethylpyrimidine (ACDP) showed that the active site is located in a V-shaped cleft created between the two Bt-thiaminase I α/β domains.[18] The active site location is structurally similar to the ligand binding site in the PBPs. The active site cysteine residue is activated by a nearby glutamate residue for addition to the thiamin pyrimidine at C6′. Loss of thiazole followed by reversal of these steps, using a variety of nucleophiles, completes the thiaminase I-catalyzed reaction. Despite their highly distinct sequences and structures, thiaminase I and most thiaminase IIs utilize an activated cysteine residue to form a covalently bound intermediate. [17, 18]

Thiaminase I from *Clostridium botulinum* (Cb-thaminase I) is a 46.3 kDa protein with 51% sequence identity to Bt-thiaminase I. The catalytic cysteine residue (Cys135) was predicted by sequence alignment with other thiaminase Is and was mutated to generate a catalytically inactive Cb-thiaminase I (C143S). C143S was co-crystallized with thiamin to obtain the C143S/thiamin complex reported here. This is the first structure of a thiaminase I/substrate complex and reveals atomic level details of thiamin binding prior to degradation. This structure was used to identify the key amino acids involved in the Cb-thiaminase I reaction. Site directed mutagenesis and kinetic studies of thiamin degradation were used to assign roles to active site amino acid residues. Together, the C143S/thiamin structure and subsequent kinetic studies led to a detailed mechanistic proposal for thiaminase I. The active sites and catalytic mechanisms of Cb-thiaminase I and thiaminase II are also compared.

MATERIALS AND METHODS

Cloning, Overexpression and Purification of Cb-thiaminase I

The thiaminase I gene, *bcmE*, was cloned from *C. botulinum* A str. ATCC 19397 genomic DNA. Standard DNA manipulation methods were used for all of the cloning procedures. The gene was inserted into pTHT, a modified pET-28 plasmid with an N-terminal His_6 tag, followed by a tobacco etch virus (TEV) protease cleavage site. All active site mutants were obtained by site-directed mutagenesis of the native gene using standard, PCR-based mutagenesis.[24]

After DNA sequencing to verify plasmid accuracy, the plasmids were transformed into *Escherichia coli* BL21(DE3) cells and grown overnight at 37 °C on selective kanamycin (30 μg/mL) containing agar. Starter cultures were then grown from a single colony. A selected

colony was placed in 10 mL sterile Luria-Bertani (LB) media containing 30 μg/mL kanamycin at 37 °C with shaking overnight. The following day, 5 mL of the overnight starter culture were added directly to 1.5 L volumes of sterile LB media. Cells grew with shaking at 37 °C until reaching an OD_{600} of 0.6, at which point the incubator temperature was reduced to 15 °C. Upon reaching an induction temperature of 15 °C and an OD_{600} of 0.8, protein overexpression was initiated by adding isopropyl β- D-1-thiogalactopyranoside to the cultures with a final concentration of 0.5 mM. Cells were harvested 18 h after induction of protein overexpression by centrifugation at 2000*g* for 20 min. The cell pellet was collected and frozen at −20 °C for storage.

The cell pellet was later thawed and resuspended in 45 mL of lysis buffer (50 mM Tris-HCl, 300 mM NaCl, and 20 mM imidazole at pH 8.0) before lysing by sonication. The lysed cell extract was centrifuged at 40,000*g* for 30 min at 4 °C. The supernatant was collected and loaded onto a 2 mL Ni-nitrilo acetic acid (NTA) column (Qiagen), preequilibrated with lysis buffer. The column was then washed with 45 mL of lysis buffer to remove any nonspecifically bound contaminants from the column. The protein was eluted from the column by passing 20 mL of elution buffer (50 mM Tris-HCl, 300 mM NaCl, and 250 mM imidazole at a final pH of 8.0) through the column. The elution volume containing Cb-thiaminase I was collected. This procedure was followed for purification of Cb-thiaminase I and all active site mutant proteins. In addition to Ser143, Tyr46, Tyr80, Asp94, Glu271, and Asp302 were selected for mutation on the basis that the side chain interacted directly with the substrate or is potentially involved in activating residues that interact with the substrate.

To remove the His_6 tag for crystallization, the purified C143S Cb-thiaminase I was incubated with TEV protease during dialysis into 3.5 L of 10 mM Tris HCl at pH 8.0, 150 mM KCl and 1 mM dithiothreitol for 18 h at 4 °C, utilizing a ratio of 0.5 mg of TEV for every 10 mg of C143S. The sample was then passed over a Ni-NTA column preequilibrated with lysis buffer. The elution volume containing C143S with the His_6 tag removed was collected in the flow through. Complete cleavage and purity were assessed by SDS-PAGE analysis. The sample was buffer exchanged by overnight dialysis into 10 mM Tris-HCl at pH 8.0 and 150 mM KCl, then concentrated to 15 mg/mL using a centrifugal filter device with a molecular weight cutoff of 10,000 Da (Vivaspin). The purification process yielded approximately 3.5 mg protein per L of cell culture. The purified protein was aliquoted and flash frozen in liquid nitrogen for storage at −80 °C.

Cocrystallization of C143S and Thiamin

For crystallization using the hanging drop vapor diffusion method, the protein was diluted with a solution of 10 mM Tris pH 8.0, 150 mM KCl and 10 mM thiamin with final concentrations of 7.5 mg/mL protein and 5 mM thiamin. Equal volumes of protein and reservoir solutions were mixed and equilibrated at 18 °C against a total volume of 500 μL well solution. The initial crystallization condition was determined using the commercially available Wizard III sparse matrix screen (Emerald Biosystems). The optimized crystallization condition was 22% (w/v) polyethylene glycol 10000, 100 mM sodium citrate pH 4.4, and 2% (v/v) dioxane. Needle-like C143S/thiamin crystals grew approximately 200 μm long and 10–20 μm thick in three to five days. Crystals were cryoprotected for data collection in a solution composed of the mother liquor supplemented with 20% (v/v) ethylene glycol.

X-ray Data Collection and Processing

X-ray diffraction data were measured at the Northeastern Collaborative Access Team (NE-CAT) beamline 24-ID-C, at the Advanced Photon Source, Argonne National Laboratory, using a Quantum 315 detector (Area Detector Systems Corp.) at a distance of 200 mm. The

oscillation method was used with 1.0° rotation per frame for 180° with cryocooling. Diffraction data were measured at a wavelength of 0.97918 Å. The NE-CAT in-house automated RAPD data collection and processing system, which utilizes XDS,[25, 26] was used for indexing, integration, and scaling of the data. Data collection and processing statistics are listed in Table 1.

Structure Determination, Model Building, and Refinement

The structure of the C143S/thiamin complex was determined by the molecular replacement method using thiaminase I from *B. thiaminolyticus* (PDB ID: 2THI), with a sequence identity of 51%, as the search model. The crystal belongs to space group $P2_1$ and contains two molecules per asymmetric unit corresponding to a Matthews coefficient[27] of 1.96 $Å^3$/Da and estimated solvent content of 37%. MolRep[28] positioned two chains in the asymmetric unit. Iterative rounds of model refinement were conducted using COOT[29] for manual model building followed by refinement with PHENIX.refine[30] using default parameters. The thiamin substrate was added using PHENIX.ligandfit for placement into F_o - F_c electron density. Water molecules were added during later rounds of refinement. The quality of the structure was analyzed using MolProbity[31] and COOT.[29] Structure refinement converged with a final R-factor of 15.9% and R_{free} of 21.1%. Complete refinement statistics are listed in Table 1.

Determination of Kinetic Parameters

Steady state kinetic analysis was performed for six Cb-thiaminase I mutant proteins selected based on the crystal structure of the C143S/thiamin complex. Reactions were carried out in 100 mM potassium chloride and 50 mM phosphate buffer at pH 8.0 using 713 mM β-mercaptoethanol as the exogenous nucleophile. All reactions were incubated at 25 °C except the E271Q Cb-thiaminase I, which was measured at 37 °C due to low activity levels. The reactions were initialized by the addition of enzyme to the reaction volume containing thiamin, with thiamin concentrations ranging from 100 μM to 10 mM. Aliquots were taken from the reaction mixture after defined time intervals, quenched with 1 M HCl, filtered through a 10,000 Da cutoff membrane to remove the enzyme and analyzed by HPLC.

In the HPLC method, the ratio of 100 mM phosphate buffer at pH 6.6 (P) to methanol (M)/water mixture was varied over time in minutes (t). The gradient was carried out following the scheme (t, M%, H_2O%, P%): (0, 0, 0, 100), (5, 0, 10, 90), (9, 15, 25, 60), (14, 65, 20, 15), (19, 0, 0, 100), (25, 0, 0, 100). Peak areas of the product in the enzyme assays were measured and the concentrations calculated based on a calibration curve relating peak area to 4-methyl-5-hydroxyethylthiazole concentration. The initial rates were calculated from plots of concentration versus time. Initial rates were then plotted and fit to the Michaelis–Menten equation for calculation of kinetic parameters.

Figure Preparation

All figures were prepared using PyMOL[32] or ChemDraw (Cambridge Biosoft) and compiled in Photoshop (Adobe).

RESULTS

Overall Structure of Cb-thiaminase I

The structure of the C143S/thiamin complex, with two molecules per asymmetric unit, was determined at 2.2 Å resolution. Of the 404 possible amino acid residues in the Cb-thiaminse I sequence, the structure contains residues 39–404 for molecule A and 40–404 for molecule B. Each molecule contains thiamin, citrate, and two metal ions modeled as Mg. The two

molecules (A and B) make a total of 16 non-bonded interactions along an interface area of approximately 210 Å2, suggesting only crystal packing contacts. Size exclusion chromatography (data not shown) predicted that Cb-thiaminase I is a monomer in solution. This is also consistent with the oligomeric state of Bt-thiaminase I.[18]

Cb-thiaminase I is a member of the group II PBP superfamily and can be divided into two distinct domains (Figure 2). The N-terminal domain is composed of residues 39–120 and 297–384, while residues 145–296 and 385–403 form the C-terminal domain. Both domains have a central β-sheet flanked on either side by α-helices with three crossover segments connecting the N- and C- domains. Residues 144–145, 287–296, and 375–385 comprise the three domain linker regions. The crossover segment formed by residues 287–296 connects β_{10} to β_{11} and the crossover segment formed by residues 375– 384 connects α_{14} and α_{15}. The third crossover connects β_6 to β_7 at residue 144.

Cb-thiaminase I Active Site and Substrate Binding

The structure of the C143S/thiamin complex reveals the substrate binding site and atomic level details of binding interactions (Figure 3). Each monomer contains one thiamin molecule in a V-shaped cleft between the two domains. This cleft is approximately 17 Å deep, 15 Å long, and 12 Å wide. Six tyrosine residues, Tyr46, Tyr48, Tyr80, Tyr252, Tyr269, and Tyr300, form an outer collar of the cleft with most of the active site residues located at the bottom of the cleft. The catalytic cysteine residue Cys143, which is represented by Ser143 in the C143S/thiamin complex, is located on β_6.

Thiamin binds to the active site in the F-conformation with torsion angles of $_{\phi T}$ = −10° (C5′-C7′-N3-C2) and ϕ_P = −93.6° (N3-C7′-C5′-C4′). The N3- *C7′*- C5′angle is 108°. The thiamin pyrimidine and thiazolium moieties are approximately perpendicular with a dihedral angle of 85.9° between their two planes. The pyrimidine portion of the thiamin is positioned towards the bottom of the V-shaped binding pocket (Figure 4). The thiamin N3′ atom is positioned 2.8 Å from the Asp302 carboxylic acid group, within hydrogen bond distance if Asp302 is protonated. The 4′-amino group hydrogen bonds with Asp94 (2.9 Å), which in turn hydrogen bonds to the Tyr46 hydroxyl group (2.5 Å). Tyr80 is 3.5 Å from the positivity charged N3 atom, and hydrogen bonds to the Tyr46 hydroxyl group (2.9 Å). The thiazolium group does not interact with the enzyme except through van der Waals contacts and is extended towards the top of the binding groove with the 5-hydroxyethyl group extended out toward the solvent region.

Steady-state Kinetics of Cb-thiaminase I and Mutant Proteins

Thiamin degradation by Cb-thiaminase I was determined using HPLC to detect the reaction products. Wild type Cb-thiaminase I catalyzes the thiaminase reaction with a k_{cat} of 2.31 × 10^2 s^{-1} and a k_{cat}/K_m value of 5.10 × 10^5 M^{-1} s^{-1}. All activity was abolished by mutation of the catalytic cysteine in C143S. The kinetic parameters for five other Cb-thiaminase I active site mutants are reported in Table 2. A description of the determination of the pH rate profile for the native enzyme is provided in the supplementary material and the profile is shown in Supplementary Figure 2.

DISCUSSION

Structural Comparison of the Cb-thiaminase I with Other Proteins

A DALI[34] search was performed to identify proteins with high structural similarity to Cb-thiaminase I (Table 3). Cb-thiaminase I showed the highest structural homology to Bt-thiaminase I (PDB ID 2THI) [18] with a root mean squared deviation (RMSD) of 0.9 Å. Bt-thiaminase I, which was used as the search model for molecular replacement, has 51%

sequence identity with Cb-thiaminase I. Cb-thiaminase I is a member of the group II PBP superfamily, with three characteristic crossover segments. Most superfamily members bind small molecules including metals, vitamins, sugars, or amino acids and are components of ABC transporters. [35] The small molecule binding site is located in a cleft near the surface that corresponds structurally to the Cb-thiaminase I active site. The DALI search identified several PBPs with RMSD values in 3.0–3.5 Å range. The thiamin binding protein, TbpA, shows an RMSD of 3.9 Å. Thiaminase I is a rare example of a PBP superfamily member with enzymatic activity. The structural similarity between thiaminase I and TbpA suggests an evolutionary relationship between these two PBP group II superfamily members.[20]

PBP proteins undergo a conformational change when the small molecule ligand binds. The two domains are in an open state when no ligand is bound and close like a Venus fly trap after ligand binding.[36] A comparison between the structures of the C143S/thiamin complex and unliganded Bt-thiaminase I (2THI)[18] suggests that thiaminase I does not undergo a conformational change upon binding thiamin (Figure 5A). Both the liganded and unliganded forms appear to most closely resemble the closed conformation of group II PBPs.

Comparison of Cb-thiaminase I and Bt-thiaminase I

The previous structure of Bt-thiaminase I with the mechanism based inhibitor 4-amino-6-chloro-2, 5-dimethylpyridimine (ACDP), which lacks the thiazolium moiety, identified the catalytic cysteine residue and the general location of the active site.[18] Attempts to model the thiamin binding geometry using the inactivated thiaminase I structure were unreliable because the covalent bond between C6 of the inhibitor and Sγ of the cysteine side chain, followed by rearomatization of the pyrimidine ring, distorted the active site interactions. Therefore, the structure of thiaminase I with the suicide inhibitor ACDP does not represent any possible intermediate along the reaction pathway. Consequentially, the structure of the Bt-thiaminase I-ACDP adduct showed few active site interactions and only one hydrogen bond. In contrast the C143S/thiamin complex reveals all interactions within the enzyme/substrate complex. The thiamin pyrimidine ring in the C143S/thiamin complex forms four hydrogen bonds with active site residues (Figure 4). These include interactions between the 4′-amino group of the thiamin pyrimidine and the side chain of Asp94. Additionally, Asp302 hydrogen bonds to both the 4′-amino group and N3′, and Glu271 hydrogen bonds to Glu271. Comparison of the structures of the C143S/thiamin complex and the Bt-thiaminase-ACDP adduct shows that ACDP and the thiamin pyrimidine are approximately perpendicular (Figure 5B).

The C143S/thiamin complex structure also reveals the binding geometry of the thiamin thiazolium moiety, which is pointed towards the solvent at the top of the binding cleft. In the crystal structure, no hydrogen bonds or stacking interactions with any active site residues are observed. Previous studies showed that thiaminase I accepts thiamin analogs with modified thiazolium ring moieties; however, modifications to the pyrimidine are not tolerated.[37] These observations are consistent with the positioning of the thiamin molecule in the active site of Cb-thiaminase I.

Comparison to Thiamin Binding Sites in Other Proteins

In addition to thiaminases, biological macromolecules that bind thiamin, or its mono- or diphosphorylated forms, include ThDP-dependent enzymes, thiamin biosynthetic enzymes, thiamin transport proteins and riboswitches. Thiamin has been observed bound in three different conformations: F,V and S. The F-conformation is the lowest energy form in solution and is defined by $\phi T \approx 0°$ and $\phi_P \approx \pm 90°$, where ϕ_T is the C5′-C7′-N3-C2 torsion angle and ϕ_P is the N3-C7′-C5′-C4′ torsion angle. The V-conformation is defined by $\phi_T \approx \pm 90$ and $\phi_p \approx 90°$, and the rare S-conformation is defined by $\phi_T \approx \pm 100°$, $\phi_P \approx \pm 150°$. A

common thiamin-binding motif is found in enzymes that utilize ThDP as a cofactor, where thiamin binds in the V-conformation.[33, 38] This 30-residue motif begins with –GDG- and concludes with –NN. ThDP binding is usually facilitated by phosphate binding interactions with a divalent metal, aspartate side chains, and asparagine side chains.[39]

The thiamin, ThMP, and ThDP binding sites found in enzymes that are not ThDP-dependent show significant divergence. Thiamin pyrophosphokinase (TPK) and ThMP kinase, like Cb-thiaminase I, bind thiamin in the F-conformation; however their thiamin binding motifs show little similarity. TPK (PDB ID: 1IG3) pyrophosphorylates thiamin to ThDP and binds thiamin ($\phi_T = -6°$, $\phi_P = -78°$) through multiple interactions.[40] In addition to stacking interactions with a nearby tryptophan residue, the thiamin pyrimidine moiety hydrogen bonds through its 4′-amino group to an aspartate side chain and a main chain carbonyl group. An additional hydrogen bond forms between N1′ and a nearby serine residue. ThDP, the TPK product, is further anchored through a phosphate-binding pocket. In ThMP kinase (PDB ID: 3C9T), ThDP ($\phi_T = -20°$, $\phi_P = -64°$) is bound in a cleft with stacking interactions between the thiazolium moiety and a nearby tryptophan residue.[41] In addition to phosphate binding pocket interactions, ThDP hydrogen bonds to a glutamate side chain through its 4′-amino group. While all three enzymes hydrogen bond to the 4′-amino group using either a glutamate or aspartate side chain, they share few other thiamin binding features. TPK and ThMP kinase contain phosphate-binding pockets that facilitate binding of ThMP or ThDP; however, Cb-thiaminase I binds thiamin with its hydroxyethyl substituent at C5 extended outwards from the binding cleft towards the solvent. ThMP in thiamin phosphate synthase (PDB ID: 2TPS) displays a non-standard “cis” V-conformation ($\phi_T = -95.6°$, $\phi_P = -113°$).[42] Like Cb-thiaminase I, its pyrimidine is located at the bottom of its active site. Thiamin binding is aided by hydrogen bonding between its 4′-amino group and N3′ and with a glutamine side chain within an otherwise hydrophobic pyrimidine binding pocket.

Despite sharing only 14% sequence identity, both Cb-thiaminase I and *E. coli* thiamin binding protein have a group II PBP fold and have similar binding site locations for thiamin; however, their thiamin binding motifs differ. In the reported structure of thiamin binding protein (PDB ID: 2QRY), ThMP spans the length of the binding cleft in an F-conformation ($\phi_T = 0°$, $\phi_P = -83°$).[20] Thiamin binding protein binds ThMP by anchoring the phosphate group within a hydrogen bond rich pocket and by sandwiching the thiazolium group between two tyrosine residues. The ThMP pyrimidine interacts mainly through water mediated hydrogen bonding with one hydrogen bond to a serine residue side chain. In contrast, in Cb-thiaminase I the thiamin pyrimidine is buried at the bottom of a cleft, with multiple hydrogen bonds, and with the thiazolium portion extending upward towards the solvent.

Thiamin biosynthesis is often under the control of a ThMP riboswitch in messenger RNA. A structure for the ThMP riboswitch (PDB ID: 2HOM) shows that ThMP binds ThMP in the rare S-conformation ($\phi_T = -85°$, $\phi_P = -173°$).[43] The riboswitch forms hydrogen bonds between N3′ and the 4′-amino group of the ThMP pyrimidine and the 2-amino group and N3, respectively, of an adenine base. An additional hydrogen bond forms between ThMP N1′ and a ribosyl 3-hydroxyl group. A weak hydrogen bond is observed between the ThMP sulfur atom and a ribosyl 2-hydroxyl group. The phosphate group interacts with the O6 atoms of a pair of guanine bases, through a magnesium ion.

Cb-thiaminase I Mechanism

A mechanistic proposal for thaminase I is shown in Figure 6. In this mechanism, Glu271 functions as the base activating Cys143 for addition to the pyrimidine as previously proposed.[15, 18] This reaction is essential for activating the pyrimidine for side chain substitution: the C143S mutant is inactive and k_{cat} for the E271Q mutant is reduced 21,000-

fold (Table 2). Asp302 is hydrogen bonded to N3 of the pyrimidine suggesting that this residue also activates the pyrimidine ring for nucleophilic attack. This is supported by the observation that k_{cat} of the D302N mutant is reduced 28-fold. In the next step, the thiazole departs in a reaction assisted by deprotonation of the C4 amino group by Asp94. In support of this, the D94N mutant shows a 1000-fold reduction in k_{cat}. The thiazole departure is also facilitated by Tyr80 and the Y80F mutant shows a 215-fold reduction in k_{cat}. A possible explanation for this effect is that the tyrosine phenol, anchored by hydrogen bonding to Asp94 and Tyr46, holds the thiamin thiazolium in a conformation that maximizes its reactivity as a leaving group with the breaking C-N bond perpendicular to the plane of the pyrimidine. The Tyr80-Tyr46 hydrogen bond is not critical because the Y46F mutant shows only a 4-fold reduction in k_{cat}. Replacement of the thiazolium by a variety of nucleophiles involves the microscopic reverse of these steps. Alignment of thiaminase I sequences from several organisms shows conservation of Cb-thiaminase I residues Tyr80, Asp94, Cys143, Glu271 and Asp302 (Supplementary Figure 2). Tyr46 is only partially conserved, consistent with the high mutant activity.

Relationship to Thiaminase II

Thiaminase II from *B. subtilis* (also known as TenA) cleaves thiamin into its pyrimidine and thiazole moieties; however, its main physiological function is conversion of 2-methyl-4-amino-5-aminomethylpyrimidine (MeAP) to 2-methyl-4-amino-5-hydroxymethylpyrimidine (HMP). In contrast to thiaminase I, thiaminase II utilizes only water as the nucleophile.[17] In addition, thiaminase II will not cleave ThMP and ThDP, whereas thiaminase I tolerates modification of the thiazolium moiety and degrades many thiamin analogs.[17,37] Interestingly, like thiaminase I, thiaminase II is often clustered with other enzymes in the thiamin biosynthetic pathway. This grouping of the biosynthetic and degradation enzymes was initially thought to be counterproductive. It was later discovered that thiamin does not appear to be the preferred physiological substrate for *B. subtilis* thiaminase II. Instead, thiaminase II salvages base-degraded pyrimidine from the environment for incorporation into thiamin.[23, 44, 45]

Thiaminase II adopts an all α-helical fold that shows no sequence or structural homology to the Cb-thiaminase I structure. The active site of thiaminase II is buried, leading to exclusive use of water as the nucleophile. In contrast, the Cb-thiaminase I has a two-domain PBP fold, and an accessible active site location near the surface of the enzyme, thus allowing a variety of nucleophiles to participate in catalysis. No thiaminase II structure has been reported with bound thiamin; however, the structure of *B. subtilis* thiaminase II (PDB ID: 1YAK) with the reaction product HMP reveals the key active site residues.[17]

Despite their major structural differences, thiaminase II and Cb-thiaminase I share some active site features (Figure 7). Both feature a catalytic cysteine that is activated by glutamate, and both use an aspartic acid residue to anchor the pyrimidine through hydrogen bonds to N3′ and the 4′-amino group.[45] In both thiaminase I and thiaminase II the cysteine side can exist two conformations.[17, 18] In one conformation the cysteine residue is positioned to be deprotonated by a glutamate side chain. In the other conformation the cysteine sulfur atom is about 3 Å from the pyrimidine C6 atom. In thiaminase II the cysteine residue may also require interactions with two intermediate tyrosine residues.[45] While both thiaminase I and II appear to share a similar catalytic mechanism for thiamin degradation, their major structural differences and substrate preferences demonstrate that they evolved independently and therefore may serve different physiological purposes.

We currently differentiate between thiaminase I and thiaminase II by sequence analysis of the entire thiaminase gene. Thiaminase I is rarely found while thiaminase II is abundant. As the active sites of the two enzyme families could in principle catalyze either substitution

reaction, the possibility remains that some genes annotated as thiaminase II may have thiaminase I activity. However, since thiaminase II appears to accept a variety of modifications to the thiamin pyrimidine moiety, we are not yet able to address this interesting question by analyzing the binding sites of the leaving group.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the staff of NE-CAT at the Advanced Photon Source for assistance with data collection, Dr. Cynthia Kinsland for cloning wild type thiaminase I and mutants, and Leslie Kinsland for help in preparing the manuscript.

ABBREVIATIONS

ACDP	4-amino-6-chloro-2, 5-dimethylpyridimine
PBP	periplasmic binding protein
LB	Luria-Bertani
TEV	tobacco etch virus
NTA	nitriloacetic acid
NE-CAT	Northeastern Collaborative Access Team
RMSD	root mean squared deviation
ThMP	thiamin monophosphate
ThDP	thiamin diphosphate
TPK	thiamin pyrophosphokinase
HMP	2-methyl-4-amino-5-hydroxymethylpyrimidine
MeAP	2-methyl–4-amino-5-aminomethylpyrimidine

REFERENCES

1. Jurgenson CT, Begley TP, Ealick SE. The Structural and Biochemical Foundations of Thiamin Biosynthesis. Annu. Rev. Biochem. 2009; 78:569–603. [PubMed: 19348578]
2. Carvalho PSM, Tillitt DE, Zajicek JL, Claunch RA, Honeyfield DC, Fitzsimons JD, Brown SB. Thiamine Deficiency Effects on the Vision and Foraging Ability of Lake Trout Fry. J. Aquat. Anim. Health. 2009; 21:315–325. [PubMed: 20218505]
3. Honeyfield DC, Tillitt DE, Fitzsimons JD, Brown SB. Variation in Lake Michigan Alewife (*Alosa pseudoharengus*) Thiaminase and Fatty Acids Composition. J. Freshwater Ecol. 2010; 25:65–71.
4. Jaroszewska M, Lee BJ, Dabrowski K, Czesny S, Rinchard J, Trzeciak P, Wilczynska B. Effects of vitamin B-1 (thiamine) deficiency in lake trout (*Salvelinus namaycush*) alevins at hatching stage. Comp. Biochem. Phy.s A. 2009; 154:255–262.
5. Ketola HG, Isaacs GR, Robins JS, Lloyd RC. Effectiveness and retention of thiamine and its analogs administered to Steelhead and landlocked Atlantic salmon. J. Aquat. Anim. Health. 2008; 20:29–38. [PubMed: 18536500]
6. Kimura Y, Iwashima A. Occurrence of Thiaminase-II in. Saccharomyces cerevisiae, Experientia. 1987; 43:888–890.
7. Mccleary BV, Chick BF. Purification and Properties of a Thiaminase I Enzyme from Nardoo (Marsilea-Drummondii). Phytochemistry. 1977; 16:207–213.

8. Muller IB, Bergmann B, Groves MR, Couto I, Amaral L, Begley TP, Walter RD, Wrenger C. The Vitamin B1 Metabolism of *Staphylococcus aureus* Is Controlled at Enzymatic and Transcriptional Levels. PLOS ONE. 2009:4.
9. Puzach SS, Gorbach ZV. Characteristics of Products Developed after Degradation of Thiamin by Means of Mollusk Thiaminase-I. Vop. Med. Khim. 1989; 35:82–84.
10. Sato M, Hayashi S, Nishino K. Subcellular-Localization of Thiaminase-I in the Kidney and Spleen of Carp. Cyprinus carpio, Comp. Biochem. Physiol. 1994; A108:31–38.
11. Earl JW, Mccleary BV. Mystery of the Poisoned Expedition. Nature. 1994; 368:683–684. [PubMed: 8152477]
12. Honeyfield DC, Brown SB, Fitzsimons JD, Tillitt DE. Early mortality syndrome in great lakes salmonines. J. Aquat. Anim. Health. 2005; 17:1–3.
13. Hazell AS, Todd KG, Butterworth RF. Mechanisms of neuronal cell death in Wernicke's encephalopathy. Metab. Brain Dis. 1998; 13:97–122. [PubMed: 9699919]
14. Singleton CK, Martin PR. Molecular mechanisms of thiamine utilization. Curr. Mol. Med. 2001; 1:197–207. [PubMed: 11899071]
15. Costello CA, Kelleher NL, Abe M, McLafferty FW, Begley TP. Mechanistic studies on thiaminase I - Overexpression and identification of the active site nucleophile. J. Biol. Chem. 1996; 271:3445–3452. [PubMed: 8631946]
16. Lienhard GE. Kinetic Evidence for a (4-Amino-2-Methyl-5-Pyrimidinyl)Methyl-Enzyme Intermediate in Thiaminase-I Reaction. Biochemistry. 1970; 9:3011. [PubMed: 5474802]
17. Toms AV, Haas AL, Park JH, Begley TP, Ealick SE. Structural characterization of the regulatory proteins TenA and TenI from *Bacillus subtilis* and identification of TenA as a thiaminase II. Biochemistry. 2005; 44:2319–2329. [PubMed: 15709744]
18. Campobasso N, Costello CA, Kinsland C, Begley TP, Ealick SE. Crystal structure of thiaminase-I from Bacillus thiaminolyticus at 2.0 Å resolution. Biochemistry-Us. 1998; 37:15981–15989.
19. Spurlino JC, Lu GY, Quiocho FA. The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. J. Biol. Chem. 1991; 266:5202–5219. [PubMed: 2002054]
20. Soriano EV, Rajashankar KR, Hanes JW, Bale S, Begley TP, Ealick SE. Structural similarities between thiamin-binding protein and thiaminase-I suggest a common ancestor. Biochemistry. 2008; 47:1346–1357. [PubMed: 18177053]
21. Bale S, Rajashankar KR, Perry K, Begley TP, Ealick SE. HMP Binding Protein ThiY and HMP-P Synthase THI5 Are Structural Homologues. Biochemistry. 2010; 49:8929–8936. [PubMed: 20873853]
22. Lai RY, Huang S, Fenwick MK, Hazra A, Zhang Y, Rajashankar K, Philmus B, Kinsland C, Sanders JM, Ealick SE, Begley TP. Thiamin pyrimidine biosynthesis in Candida albicans : a remarkable reaction between histidine and pyridoxal phosphate. J. Am. Chem. Soc. 2012; 134:9157–9159. [PubMed: 22568620]
23. Jenkins AH, Schyns G, Potot S, Sun G, Begley TP. A new thiamin salvage pathway. Nat. Chem. Biol. 2007; 3:492–497. [PubMed: 17618314]
24. Sambrook, J.; Fritsch, EF.; Maniatis, T. Molecular Cloning: A Laboratory Manual. Vol. Vol. 3. Plainview, New York: Cold Spring Harbor Laboratory Press; 1989.
25. Kabsch W. Xds. Acta Crystallogr. D. 2010; 66:125–132. [PubMed: 20124692]
26. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D. 2010; 66:133–144. [PubMed: 20124693]
27. Matthews BW. Solvent content of protein crystals. J. Mol. Biol. 1968; 33:491–497. [PubMed: 5700707]
28. Vagin A, Teplyakov A. An approach to multi-copy search in molecular replacement. Acta Crystallogr. D. 2000; 56:1622–1624. [PubMed: 11092928]
29. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D. 2004; 60:2126–2132. [PubMed: 15572765]
30. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC,

Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta. D. 2010; 66:213–221.

31. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. 2010; 66:12–21. [PubMed: 20057044]
32. DeLano, WL. The PyMOL Molecular Graphics System. San Carlos, CA: DeLano Scientific; 2002.
33. Shin W, Oh DG, Chae CH, Yoon TS. Conformational analyses of thiamin-related compounds. A stereochemical model for thiamin catalysis. J. Am. Chem. Soc. 1993; 115:12238–12250.
34. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010; 38(Suppl):W545–W549. [PubMed: 20457744]
35. Dwyer MA, Hellinga HW. Periplasmic binding proteins: a versatile superfamily for protein engineering. Curr. Opin. Struct. Biol. 2004; 14:495–504. [PubMed: 15313245]
36. Mao B, Pear MR, McCammon JA, Quiocho FA. Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. J. Biol. Chem. 1982; 257:1131–1133. [PubMed: 7035444]
37. Bos M, Kozik A. Some molecular and enzymatic properties of a homogeneous preparation of thiaminase I purified from carp liver. J. Protein Chem.y. 2000; 19:75–84.
38. Andrews FH, Tom AR, Gunderman PR, Novak WR, McLeish MJ. A bulky hydrophobic residue is not required to maintain the v-conformation of enzyme-bound thiamin diphosphate. Biochemistry. 2013; 52:3028–3030. [PubMed: 23607689]
39. Hawkins CF, Borges A, Perham RN. A common structural motif in thiamin pyrophosphate-binding enzymes. Febs Letters. 1989; 255:77–82. [PubMed: 2792374]
40. Timm DE, Liu J, Baker LJ, Harris RA. Crystal structure of thiamin pyrophosphokinase. J. Mol. Biol. 2001; 310:195–204. [PubMed: 11419946]
41. McCulloch KM, Kinsland C, Begley TP, Ealick SE. Structural studies of thiamin monophosphate kinase in complex with substrates and products. Biochemistry. 2008; 47:3810–3821. [PubMed: 18311927]
42. Chiu HJ, Reddick JJ, Begley TP, Ealick SE. Crystal structure of thiamin phosphate synthase from Bacillus subtilis at 1.25 Å resolution. Biochemistry. 1999; 38:6460–6470. [PubMed: 10350464]
43. Edwards TE, Ferre-D'Amare AR. Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. Structure. 2006; 14:1459–1468. [PubMed: 16962976]
44. Barison N, Cendron L, Trento A, Angelini A, Zanotti G. Structural and mutational analysis of TenA protein (HP1287) from the *Helicobacter pylori* thiamin salvage pathway - evidence of a different substrate specificity. FEBS J. 2009; 276:6227–6235. [PubMed: 19780837]
45. Jenkins AL, Zhang Y, Ealick SE, Begley TP. Mutagenesis studies on TenA: A thiamin salvage enzyme from. Bacillus subtilis Bioorg. Chem. 2008; 36:29–32.

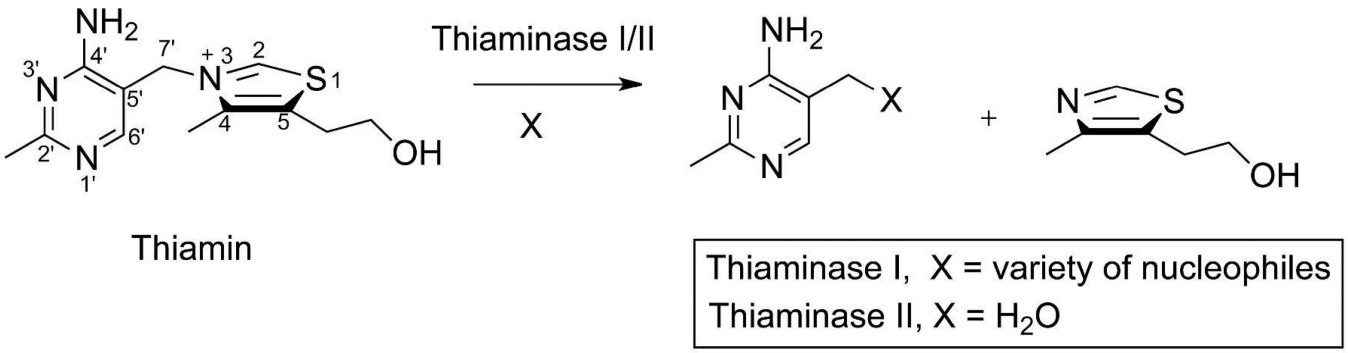

Figure 1.
The thiamin cleavage reaction catalyzed by thiaminases. Thiaminase I utilizes a variety of nucleophiles, whereas thiaminase II exclusively uses water.

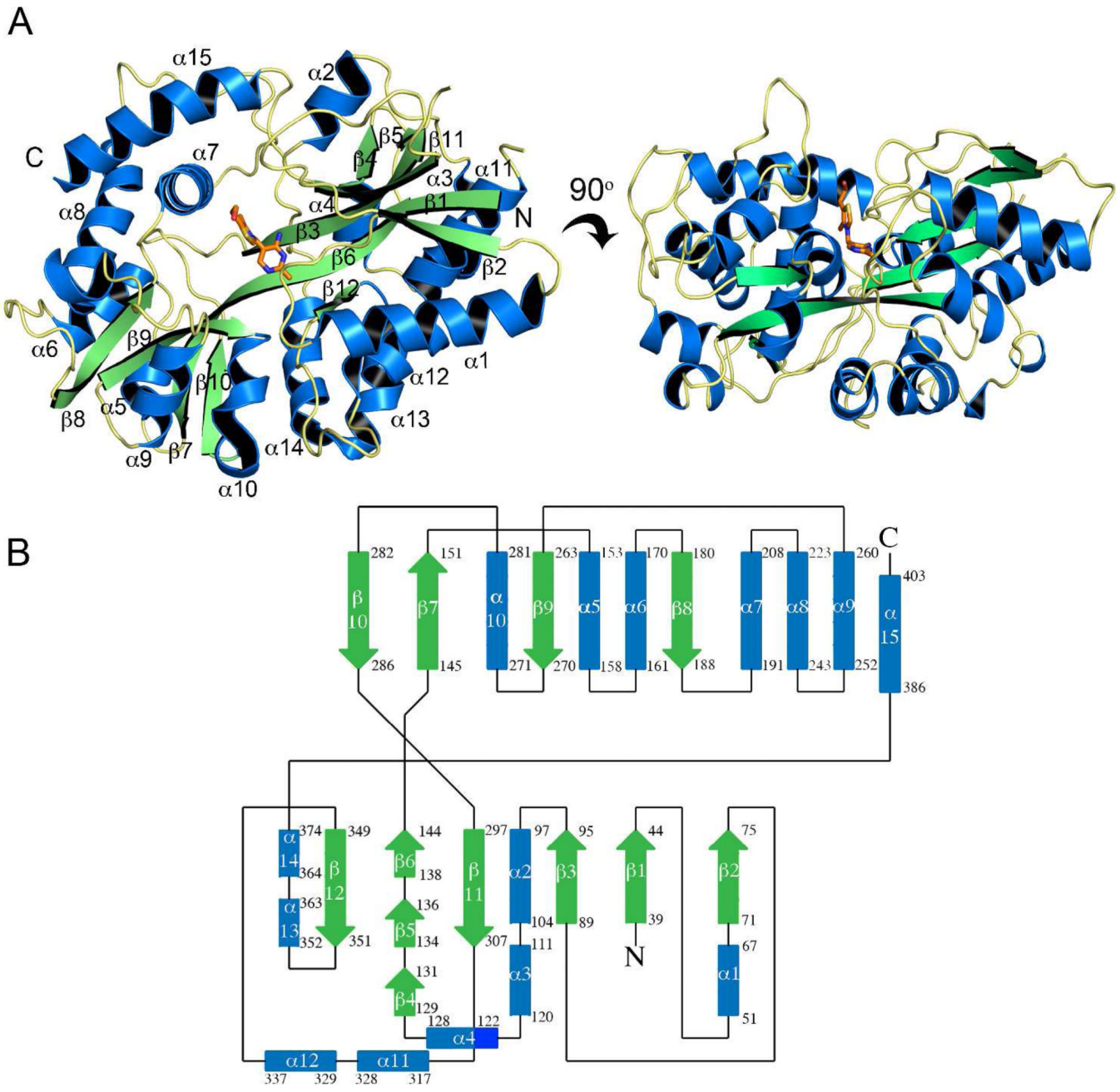

Figure 2.

Overall structure of C143S/thiamin. (A) The ribbon diagram shows the α/β fold of C143S with α-helices depicted in blue and β-strands in green. The two domains are composed of β-sheets flanked by α-helices on both sides, connected by a central β-strand. The thiamin-binding site is located in the central groove between the N- and C-terminal domains. The 90° rotation shows the orientation of the thiamin molecule in the binding cleft. (B) The topology diagram of C143S following the same coloring scheme as in (A).

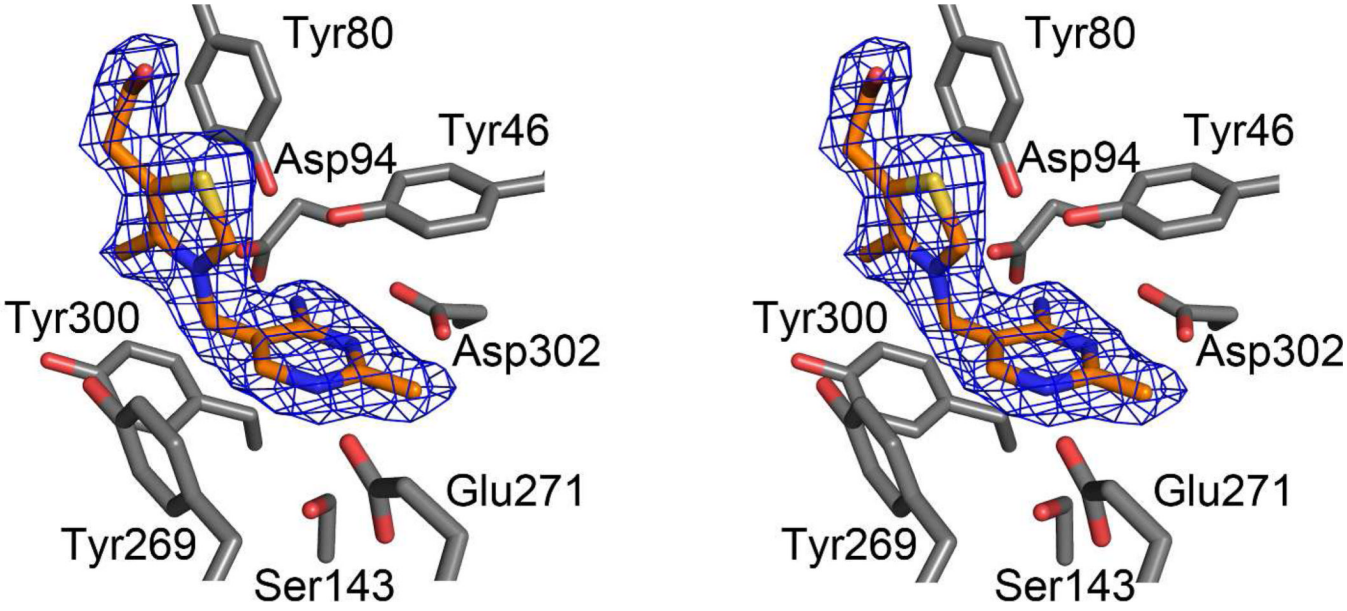

Figure 3.
C143S/thiamin active site interactions. The stereodiagram displays active site residues near the thiamin binding site with thiamin modeled into a F_o-F_c electron density map contoured at 1σ.

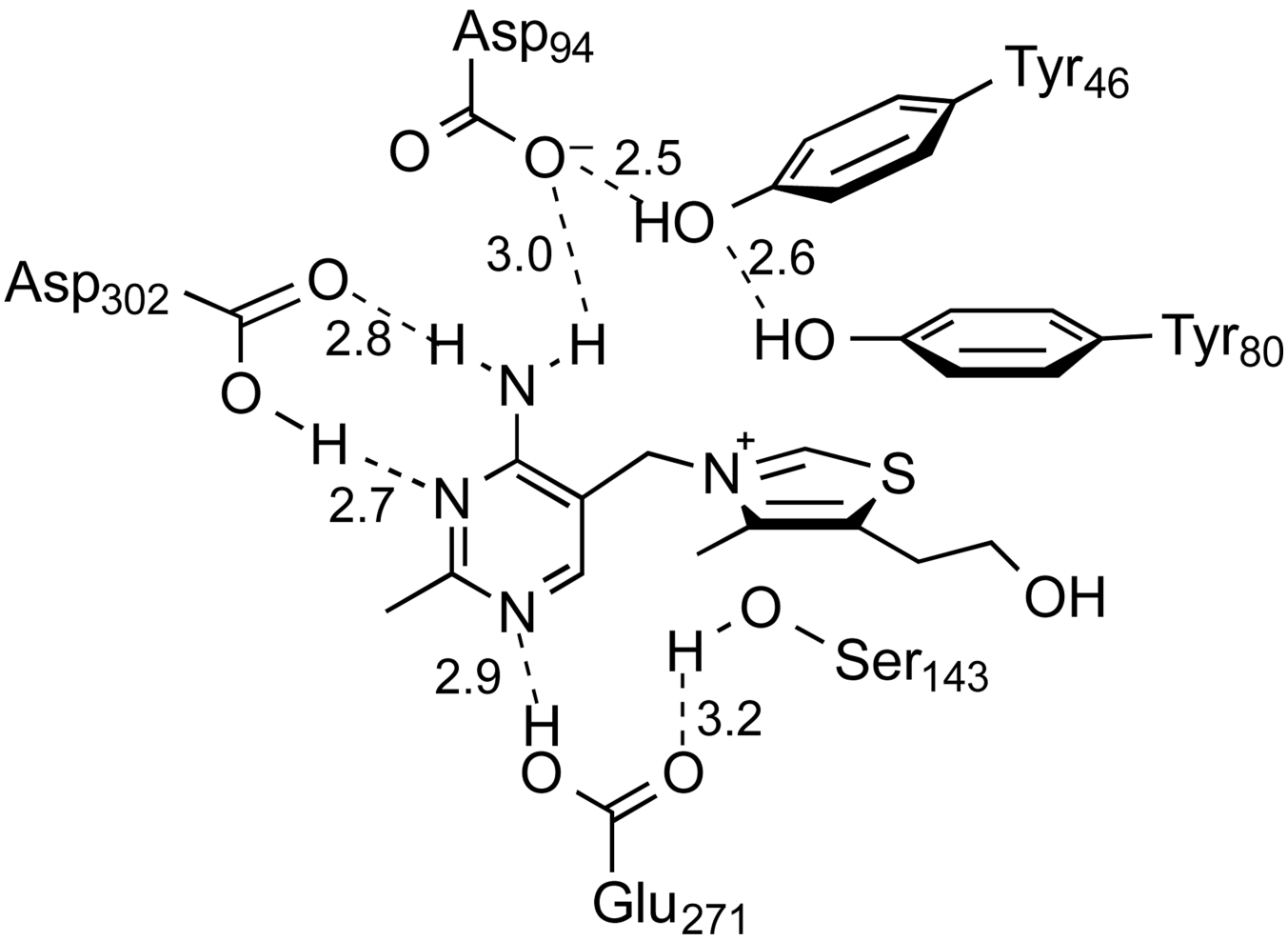

Figure 4.
Thiamin binding interactions in C143S thiaminase I. All distance values given are in Å.

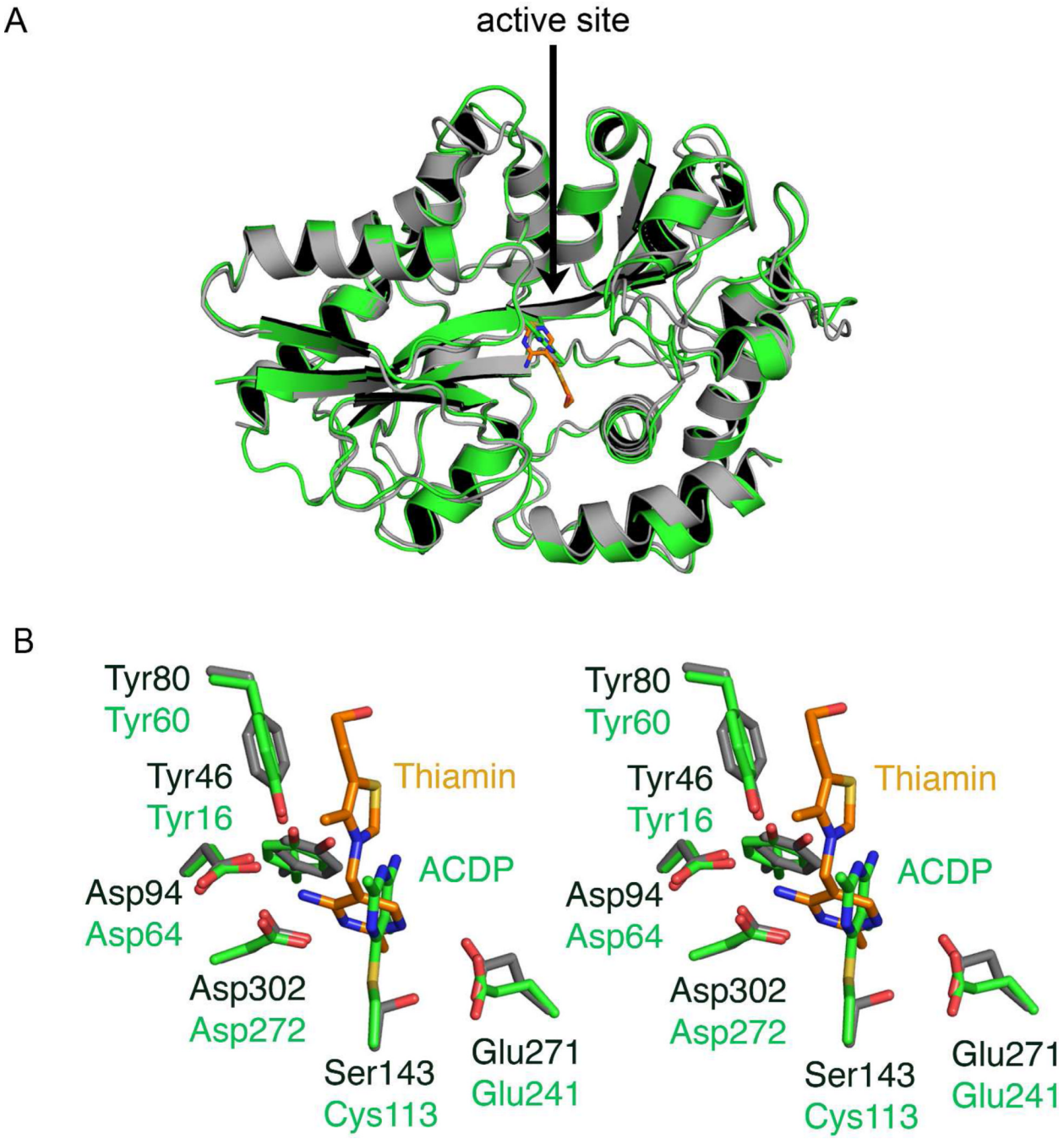

Figure 5.

Comparison with thiaminase I/ACDP structure from *B. thiaminolyticus*. (A) C143S structure (gray) with thiamin (orange) superimposed on Bt-thiaminase I with covalently bound inhibitor (green, PDB ID: 4THI). (B) Active site comparison between C143S and Bt-thiaminase I following the same coloring scheme as in (A). The pyrimidine portion of thiamin from the C143S/thiamin complex is oriented perpendicular to the ACDP inhibitor covalently bound to the active site cysteine in Bt-thiaminase I.

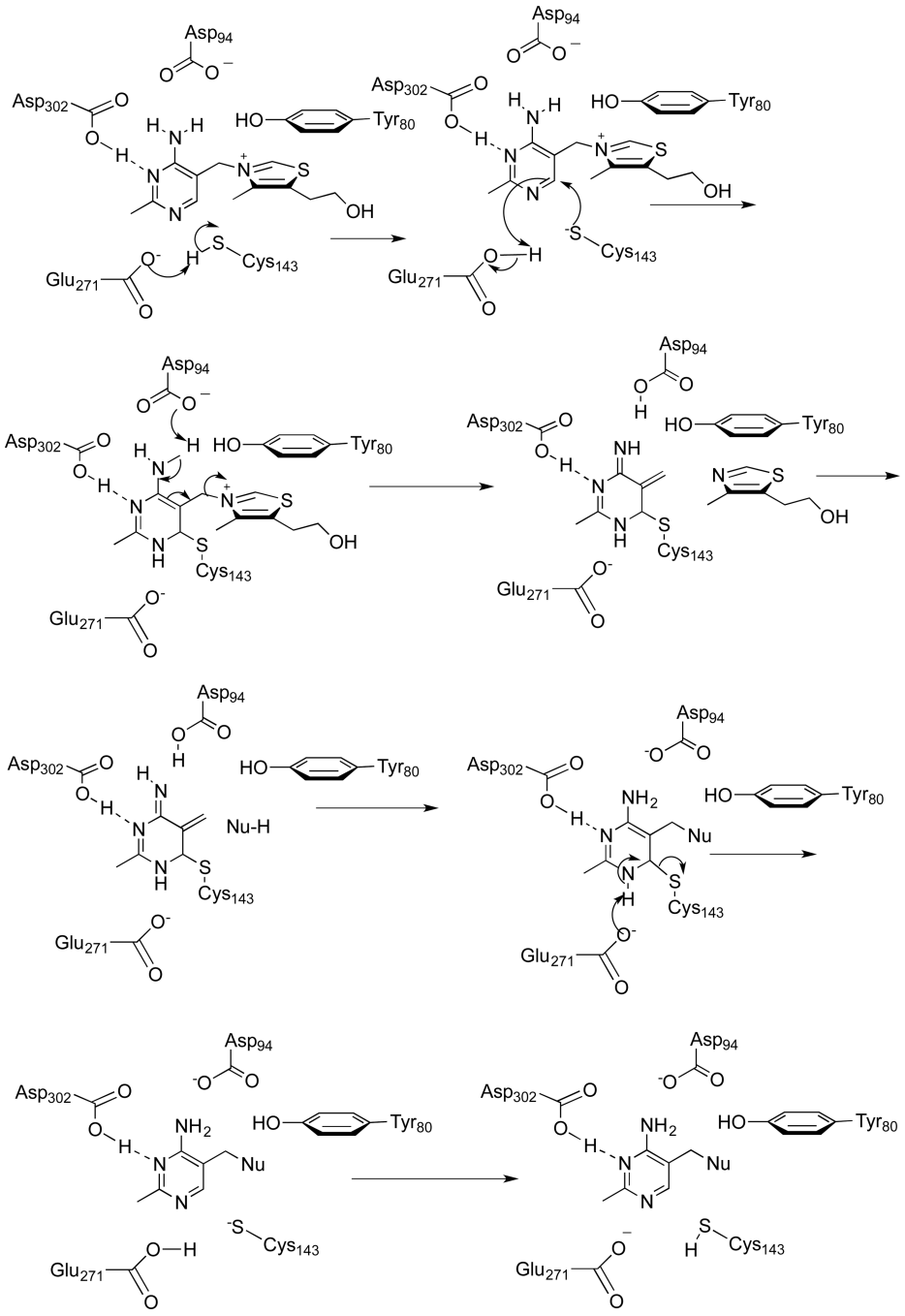

Figure 6.
Mechanism of thiamin degradation by Cb-thiaminase I. This mechanism is based on kinetic studies of Cb-thiaminase I mutants and structural data.

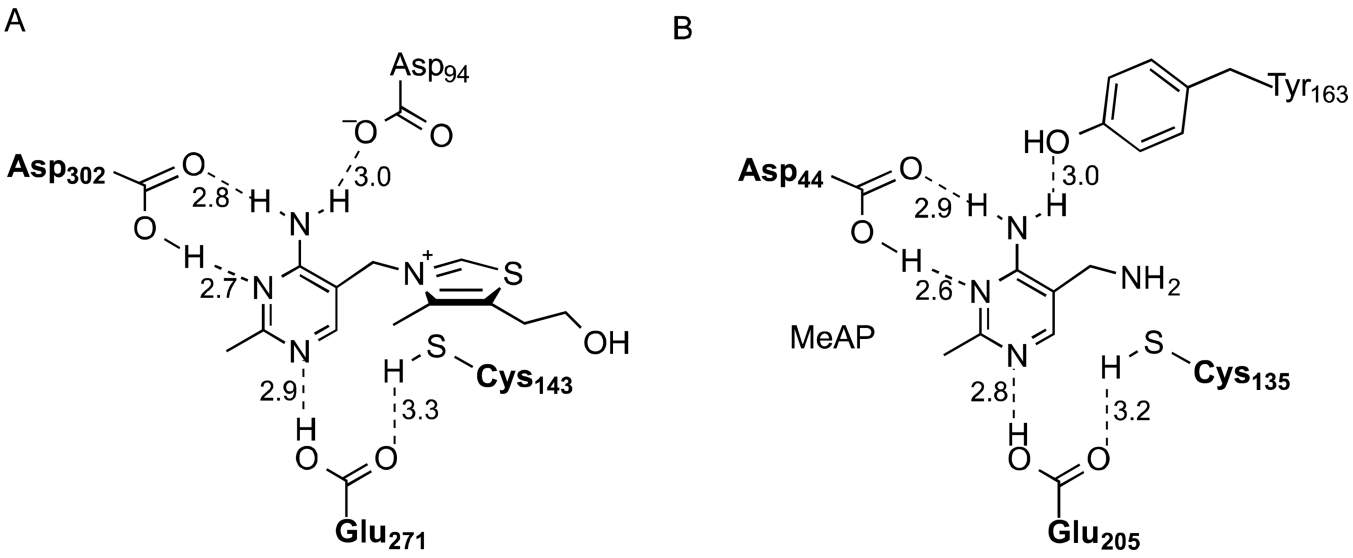

Figure 7.

Comparison of active sites for thiaminase I and II. (A) Thiamin binding interactions in Cb-thiaminase I modeled from the C143S/thiamin structure. Ser143 was replaced by cysteine while the χ_1 was kept constant. (B) Model for MeAP binding in thiaminase II (PDB ID: 1YAK). The substrate was generated by replacing the HMP hydroxyl group with an amino group. Functionally conserved residues are labeled in bold font.

Table 1

Summary of data collection and refinement statistics

	C143S/Thiamin
beamline	APS 24-ID-C
resolution (Å)	2.18
wavelength (Å)	0.97918
space group	$P2_1$
a (Å)	63.5
b (Å)	36.3
c (Å)	156.8
β (°)	91.4
Matthews coefficient	1.96
% solvent	37.4
mol/asu	2
Measured reflections	132,181
Unique reflections	37,678
Average I/σ	16.3 (5.0)[a]
Redundancy	3.4 (3.0)
Completeness (%)	99 (92)
R_{sym} (%)[b]	6.5 (29.4)
No. of protein atoms	5814
No. of ligand atoms	66
No. of water atoms	430
Reflections in working set	37,536
Reflections in test set	1,877
R-factor/R_{free} (%)[c]	15.9/21.1
rms deviation from ideals	
bonds (Å)	0.008
angles (°)	1.045
average *B* factor for protein (Å^2)	21.7
average *B* factor for water (Å^2)	26.3
average *B* factor for ligand (Å^2)	26.3
Ramachandram plot	
most favored (%)	99.0
allowed (%)	1.0
disallowed (%)	0.0

[a] Values in parentheses are for the highest-resolution shell.

[b] $R_{merge} = \Sigma\Sigma_i | I_i - \langle I \rangle | / \Sigma \langle I \rangle$, where $\langle I \rangle$ is the mean intensity of the N reflections with intensities I_i and common indices h,k,l.

[c] R-factor = $\Sigma_{hkl} ||F_{obs}| - k |F_{cal}|| / \Sigma_{hkl} |F_{obs}|$ where F_{obs} and F_{cal} are observed and calculated structure factors, respectively, calculated over all reflections used in the refinement. R_{free}, is similar to R_{work} but calculated over a subset of reflections (5%) excluded from all stages of refinement.

Table 2

Kinetic properties of thiaminase activity for wild type Cb-thiaminase I and active site mutants

Mutant	k_{cat} (s^{-1})	K_M (M)	k_{cat}/K_M ($M^{-1}s^{-1}$)
wild type	$(2.31 \pm 1.10) \times 10^{2}$	$(4.52 \pm 0.51) \times 10^{-4}$	5.10×10^{5}
C143S	nd*	nd*	-
Y80F	1.07 ± 0.12	$(2.11 \pm 0.81) \times 10^{-3}$	4.96×10^{2}
D302N	8.14 ± 0.74	$(3.07 \pm 1.09) \times 10^{-5}$	2.64×10^{4}
E271Q	$(1.1 \pm 0.03) \times 10^{-2}$	$(7.47 \pm 1.25) \times 10^{-5}$	1.50×10^{2}
Y46F	$6.09 \pm 0.38 \times 10^{1}$	$(2.45 \pm 0.51) \times 10^{-4}$	2.52×10^{5}
D94N	$(2.3 \pm 0.2) \times 10^{-1}$	$(3.49 \pm 0.76) \times 10^{-3}$	6.61×10^{1}

*Activity not detected at the level of sensitivity of the assay.

Table 3

Enzymes structurally similar to Cb-thiaminase I

Protein	PDB ID	Z score	rmsd	% identical	no. of aligned residues
Thiaminase I	2THI	57.2	0.9	51	361
Maltose PBP	1EU8	28.6	3.3	17	328
β-D-galctopyranose PBP	3OO6	26.2	3.4	11	390
Cyclodextrin PBP	2DFZ	25.2	3.4	16	323
Iron PBP	1O7T	21.2	3.0	10	287
Thiamin PBP	2QRY	18.3	3.9	14	308