1226 *Biochemistry* **2010,** *49,* 1226–1235
DOI: 10.1021/bi901812p

BIOCHEMISTRY
including biophysical chemistry & molecular biology
*Article*

# Structure Determination and Characterization of the Vitamin $B_6$ Degradative Enzyme (*E*)-2-(Acetamidomethylene)succinate Hydrolase[†,‡]

Kathryn M. McCulloch,[§] Tathagata Mukherjee,[§] Tadhg P. Begley,*[,∥] and Steven E. Ealick*[,§]

[§]*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853 and*
[∥]*Department of Chemistry, Texas A&M University, College Station, Texas 77842*

*Received October 22, 2009; Revised Manuscript Received December 23, 2009*

ABSTRACT: The gene identification and kinetic characterization of (*E*)-2-(acetamidomethylene)succinate (*E*-2AMS) hydrolase has recently been described. This enzyme catalyzes the final reaction in the degradation of vitamin $B_6$ and produces succinic semialdehyde, acetate, ammonia, and carbon dioxide from *E*-2AMS. The structure of *E*-2AMS hydrolase was determined to 2.3 Å using SAD phasing. *E*-2AMS hydrolase is a member of the $\alpha/\beta$ hydrolase superfamily and utilizes a serine/histidine/aspartic acid catalytic triad. Mutation of either the nucleophilic serine or the aspartate resulted in inactive enzyme. Mutation of an additional serine residue in the active site causes the enzyme to be unstable and is likely structurally important. The structure also provides insight into the mechanism of hydrolysis of *E*-2AMS and identifies several potential catalytically important residues.

Vitamin $B_6$ is an essential vitamin whose active cofactor, pyridoxal 5′-phosphate (PLP),[1] plays an important role within the cell to stabilize carbanion intermediates adjacent to an amino group (*1*). Vitamin $B_6$ degradation was first shown to occur in select microorganisms in the late 1950s by culturing bacteria on minimal media where the only carbon and nitrogen sources were different $B_6$ vitamers (*2*). Two pathways for the degradation of vitamin $B_6$ have been described; the first has been characterized in *Pseudomonas* sp. MA-1 and the second in *Pseudomonas* IA and *Arthrobacter* Cr-7 (*2*). None of the PLP catabolic genes in these microorganisms were identified. Gene identification was greatly facilitated when the genome of *Mesorhizobium loti* MAFF303099 was sequenced in 2000 (*3*). Shortly thereafter, the pyridoxine 4-oxidase gene was identified, suggesting that this bacterium was capable of degrading vitamin $B_6$ (*4*). Six other catabolic genes were rapidly identified, completing the vitamin $B_6$ degradative pathway in *M. loti*, shown in Figure 1 (*5*−*11*). Crystal structures have been determined for three of these enzymes (*12*−*14*). The *M. loti* degradative pathway is slightly different from that observed in *Pseudomonas* sp. MA-1.

(*E*)-2-(Acetamidomethylene)succinate (*E*-2AMS) hydrolase is the last enzyme to be biochemically characterized on the vitamin $B_6$ catabolic pathway. This enzyme catalyzes the hydrolysis of *E*-2AMS to succinic semialdehyde, ammonia, acetate, and carbon dioxide. The products of this reaction have been conclusively identified, and the steady-state kinetics of the enzyme have been characterized (*7*, *10*). A similar reaction in nicotinate catabolism, catalyzing the amide hydrolysis of 2-(enamine)glutarate, is performed by an enamidase, a member of the amidohydrolase superfamily (*15*). The primary sequence of *E*-2AMS hydrolase, however, indicated that this enzyme was most similar to members of the $\alpha/\beta$ hydrolase superfamily, although the sequence homology was low. This superfamily often utilizes a catalytic triad consisting of an active site nucleophile, an absolutely conserved histidine residue, and an acidic residue. Sequence alignments failed to identify this acidic residue, in *E*-2AMS and it was unclear how this enzyme catalyzed its metal-independent hydrolysis reaction (*10*). The structure reported here identifies this acidic residue and also suggests several other residues within the active site that could be involved in substrate binding and catalysis.

## MATERIALS AND METHODS

*Cloning, Overexpression, and Purification of E-2AMS Hydrolase*. The molecular cloning of the mlr6787 gene from *M. loti* MAFF303099 has been previously described (*7*). The resulting pMl5331.XF1 plasmid contains the mlr6787 gene with an N-terminal polyhistidine tag and was transformed into *Escherichia coli* B834(DE3), a cell line auxotrophic for methionine. Overnight cultures were grown by transferring a single colony to 10 mL of LB medium with 150 $\mu$g/mL ampicillin at 37 °C with shaking. Selenomethionyl-incorporated (SeMet) protein was grown using a minimal medium supplemented with 0.4% glucose, 20 mg/L all amino acids except methionine, 50 mg/L L-selenomethionine, 1× MEM vitamin mix (Invitrogen), 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 25 mg/L $FeSO_4$, and 150 $\mu$g/mL ampicillin. The overnight culture, grown in LB, was harvested at 4 °C at 2000*g* for 15 min and then resuspended in the minimal medium for inoculation of 1 L cultures. The large culture was allowed to grow at 37 °C with shaking to an $OD_{595}$ of 0.6. The temperature was then lowered to 15 °C, and overexpression of protein was induced overnight with 1 mM

[†]This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by Award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract W-31-109-ENG-38. This research was supported by a grant from the National Institutes of Health (GM069618) and a NIH Chemistry and Biology Interface Training Grant (GM008500).

[‡]The Brookhaven Protein Data Bank code for *E*-2AMS hydrolase is 3KXP.

*To whom correspondence should be addressed. Telephone: (607) 255-7961. Fax: (607) 255-1227. E-mail: see3@cornell.edu; begley@chem.tamu.edu.

[1]Abbreviations: PLP, pyridoxal 5′-phosphate; *M. loti*, *Mesorhizobium loti*; *E*-2AMS, (*E*)-2-(acetamidomethylene)succinate; *E. coli*, *Escherichia coli*; SeMet, selenomethionyl; IPTG, isopropyl $\beta$-D-thiogalactopyranoside; DTT, dithiothreitol; NCS, noncrystallographic symmetry.

 Published on Web 01/25/2010 

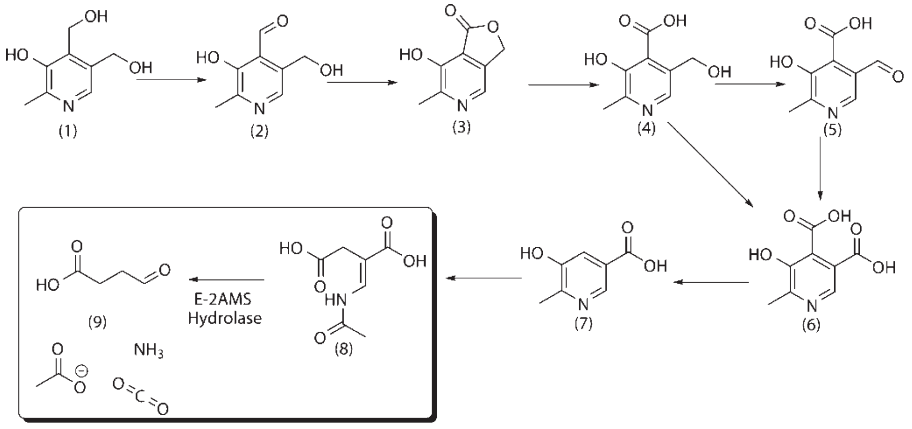


FIGURE 1: Vitamin $B_6$ degradative pathway found in *M. loti*: **1**, pyridoxine; **2**, pyridoxal; **3**, 4-pyridoxolactone; **4**, 4-pyridoxic acid; **5**, 5-formyl-2-methyl-3-hydroxypyridine-4-carboxylic acid; **6**, 2-methyl-3-hydroxypyridine-4,5-dicarboxylic acid; **7**, 2-methyl-3-hydroxypyridine-5-carboxylic acid; **8**, 2-(acetamidomethylene)succinic acid; **9**, succinic semialdehyde. Other final products are ammonia, acetate, and carbon dioxide.

isopropyl $\beta$-D-thiogalactopyranoside (IPTG). After 16 h, cells were pelleted by centrifugation at 6000*g* and stored frozen at −20 °C. For native protein overexpression, 1 L of LB medium with 150 $\mu$g/mL ampicillin was inoculated with 5 mL of overnight culture, then grown at 37 °C with shaking to an $OD_{595}$ of 0.6, and induced with 1 mM IPTG. The cells were harvested as described for the SeMet protein preparation.

Purification of both native and SeMet protein followed the same protocol. Frozen cell pellet was resuspended in 30 mL of purification buffer (50 mM $NaH_2PO_4$ at pH 8.0, 300 mM NaCl, 3 mM $\beta$-mercaptoethanol) with 10 mM imidazole and lysed by sonication. After lysis, the cell extract was clarified by centrifugation at 40000*g* for 1 h at 4 °C. The supernatant was then twice passed over a 2 mL Ni-NTA column (Qiagen) preequilibrated with purification buffer. The column was then washed with 50 mL of purification buffer, and nonspecifically binding contaminants were removed by washing with 25 mL of purification buffer containing 20 mM imidazole. The protein was eluted from the column using purification buffer containing 250 mM imidazole. The resulting sample was further purified using size exclusion chromatography (HiLoad 26/60 Superdex 75 pg;, GE Healthcare) to greater than 95% homogeneity as judged by SDS−PAGE analysis (results not shown). The protein samples were concentrated to ~8 mg/mL as measured by Bradford assay and stored at −80 °C in storage buffer (20 mM Tris, pH 8.0, 50 mM NaCl) with 1 mM dithiothreitol (DTT) added to the SeMet sample (*16*).

*Activity Assay for E-2AMS Hydrolase*. The assay used for determining the activity of *E*-2AMS hydrolase has previously been described (*7*). Briefly, freshly purified enzyme was added to 50 mM phosphate buffer containing a quantified amount of *E*-2AMS, and the decrease in absorbance at 261 nm is monitored. This assay was used to determine the steady-state kinetic parameters previously reported (*7*, *10*).

*Crystallization of E-2AMS Hydrolase*. The hanging-drop vapor diffusion method was used to crystallize *E*-2AMS hydrolase by mixing equal volumes of protein sample with reservoir solution and allowing the drop to equilibrate against 500 $\mu$L of reservoir solution at 22 °C. Initial hits were determined for the native protein using commercially available sparse matrix screens (Hampton Research, Emerald BioSystems). Optimized reservoir conditions consisted of 8−12% PEG 8000, 200 mM $MgCl_2$, and 100 mM Tris buffer, pH 7.4−8.0, and 1.5 $\mu$L of reservoir solution was mixed with an equal volume of protein solution. Tetragonal bipyramidal crystals grew to 200 $\mu$m × 150 $\mu$m within 2 days. SeMet *E*-2AMS hydrolase crystals grew under the same conditions but required an additional microseeding step. Briefly, native crystals were transferred to a seed stabilization solution of 9% PEG 8000, 200 mM $MgCl_2$, and 100 mM Tris at pH 7.8 and then crushed using Seed Bead (Hampton Research). The freshly prepared seeds were then serially diluted, and 0.5 $\mu$L of seed solution was added to SeMet hanging drops preequilibrated overnight. Crystals began growing within 20 min of seeding and reached full size overnight. Native and SeMet *E*-2AMS hydrolase crystals both belong to the space group $P2_12_12_1$ with average unit cell parameters of $a$ = 115 Å, $b$ = 179 Å, and $c$ = 189 Å. The asymmetric unit consists of 12 chains, corresponding to a Matthew's coefficient of 2.40 $Å^3$/Da and a solvent content of 49% (*17*).

*X-ray Data Collection and Processing*. Protein crystals were cryoprotected in the crystallization solution supplemented with 17% glycerol and then flash frozen by plunging into liquid nitrogen. A single wavelength anomalous data set was collected on a single SeMet *E*-2AMS hydrolase crystal on the NE-CAT 24-ID-C beamline at the Advanced Photon Source. A fluorescence scan was performed on a SeMet crystal to determine the maximum $f''$ for selenium, and the data were collected at this wavelength to a maximum resolution of 2.7 Å over a total rotation range of 180°. A native crystal diffracted to 2.3 Å resolution, and data were collected over 110° on the NE-CAT 24-ID-E beamline. Both data sets were collected using Quantum315 detectors (Area Detector Systems Corp.) and used a 0.5° oscillation range to minimize reflection overlaps. The data were then indexed, integrated, and scaled using the HKL2000 suite of programs (*18*). Data collection statistics are given in Table 1.

*Structure Determination, Model Building, and Refinement*. Positions for 36 of a possible 60 selenium atoms were determined using the SAD data set to 2.7 Å and the program HKL2MAP (*19*−*21*). These heavy atom coordinates were then used for SAD phasing, followed by density modification and calculation of initial electron density maps in CNS (*22*). RESOLVE was then used to perform phase extension to 2.3 Å with the native data and automated model building (*23*, *24*). The initial model, consisting of 1511 out of 3336 residues in the asymmetric unit, required extensive manual model building in COOT (*25*). Once 12 chains had been positioned in the asymmetric unit relying solely on the experimental phase information, iterative rounds of refinement were performed in CNS followed by additional manual model building. Strict noncrystallographic

Table 1: Summary of Data Collection Statistics

| | SeMet | native |
|---|---|---|
| source | APS 24-ID-C | APS 24-ID-E |
| resolution (Å) | 2.70 | 2.26 |
| wavelength (eV) | 12662 | 12667 |
| space group | $P2_12_12_1$ | $P2_12_12_1$ |
| *a* (Å) | 115.2 | 115.2 |
| *b* (Å) | 179.0 | 178.5 |
| *c* (Å) | 189.2 | 189.3 |
| Matthew's coeff | 2.4 | 2.4 |
| % solvent | 49 | 49 |
| mol/asu | 12 | 12 |
| measured reflections | 531900 | 727191 |
| unique reflections[a] | 207285 (20629)[b] | 168768 (12441) |
| average $I/\sigma$ | 11.7 (2.4) | 13.1 (1.7) |
| redundancy | 2.6 (2.5) | 4.3 (2.1) |
| completeness (%) | 99.9 (99.4) | 91.6 (68.2) |
| $R_{sym}$[c] (%) | 7.9 (35.7) | 10.3 (45.8) |

[a]Unique reflections include Bijvoet pairs for anomalous data. [b]Values in parentheses are for the highest resolution shell. [c]$R_{sym} = \sum_i\sum|I_i - \langle I\rangle|/\sum\langle I\rangle$, where $\langle I\rangle$ is the mean intensity of the *N* reflections with intensities $I_i$ and common indices *h*, *k*, *l*.

Table 2: Summary of Data Refinement Statistics

| | *E*-2AMS hydrolase |
|---|---|
| resolution (Å) | 50.00−2.30 |
| no. of protein atoms | 24444 |
| no. of ligand atoms | 12 |
| no. of water atoms | 1387 |
| reflections in working set | 279026 |
| reflections in test set | 18821 |
| *R* factor[a] (%) | 20.4 |
| $R_{free}$[b] (%) | 24.5 |
| rmsd from ideal | |
| bonds (Å) | 0.0058 |
| angles (deg) | 1.2 |
| average *B*-factor (Å$^3$) | 28.4 |
| Ramachandran plot | |
| most favored (%) | 87.6 |
| additionally allowed (%) | 11.9 |
| generously allowed (%) | 0.1 |
| disallowed (%) | 0.4 |

[a]$R$ factor $= \sum_{hkl}||F_o| - k|F_c||/\sum_{hkl}|F_o|$, where $F_o$ and $F_c$ are observed and calculated structure factors, respectively. [b]For $R_{free}$, the sum is extended over a subset of reflections (5%) excluded from all stages of refinement.

symmetry (NCS) restraints were utilized and gradually lessened throughout refinement (*26*). Water molecules were added during later rounds of refinement using CNS. The final model, consisting of residues 11−278 for 12 chains, was analyzed using PROCHECK; Ser106 and Asp257 were each identified as having geometry that falls within the disallowed region, although each residue has clear electron density (*27*). Final refinement statistics are summarized in Table 2. Figures were generated using PYMOL and ChemDraw (*28*).

*Mutagenesis of E-2AMS Hydrolase*. Mutants of *E*-2AMS hydrolase were prepared using standard procedures for site-directed mutagenesis (*29*). The plasmid p5331.XFI in the *E. coli* Mach1 cell line was used as template DNA. Primers used for generating S106A are as follows: 5′-(for)CCA TCC TCG TCG GAC ACG CGC TTG GTG CTC GAA ATT CGG-3′; 5′-(rev)CCG AAT TTC GAG CAC CAA GCG CGT GTC CGA CGA GGA TGG-3′. The D130N mutant was prepared

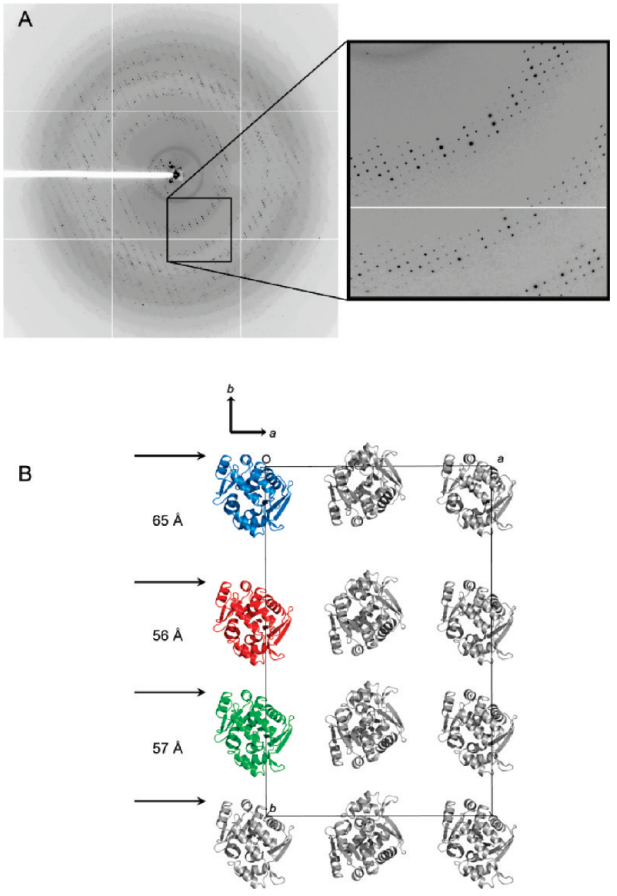


FIGURE 2: Pseudotranslational symmetry in *E*-2AMS hydrolase crystals. (A) A typical diffraction pattern for *E*-2AMS hydrolase. At low resolution, every third row of reflections is much more intense than other rows. (B) Crystal packing of *E*-2AMS hydrolase. For clarity only one-third of the unit cell is shown, with the three unique chains colored blue, red, and green. The crystallographically related protomers are shown in gray.

using the following primers: 5′-(for)GGT GCG GTC GGT CGT CGC GAT TAA CTT TAC GCC GTA CAT CGA G-3′; 5′-(rev)C TCG ATG TAC GGC GTA AAG TTA ATC GCG ACG ACC GAC CGC ACC-3′. To generate the S230A mutant, the following primers were used: 5′-(for)CGT TCG GGG CGA GTC CGC CAA GTT GGT TTC TGC G-3′; 5′-(rev)C GCA GAA ACC AAC TTG GCG GAC TCG CCC CGA ACG-3′. Additionally, the S230C and S230N mutants were generated by replacing the underlined bases in the forward primer with CTG and TTG and CAG and CAA, respectively. These mutants were each overexpressed and purified as described for the wild-type *E*-2AMS hydrolase. The S106A mutant was crystallized using the seeding procedure required for the SeMet sample.

## RESULTS

*Packing of E-2AMS Hydrolase*. A typical diffraction pattern from an *E*-2AMS hydrolase data set is shown in Figure 2A. At low resolution, it is clear that every third row of diffraction maxima exhibited a much greater intensity than the intervening two rows of maxima. This pattern indicated the possibility of pseudotranslational symmetry (*30*), which was confirmed by the structure. Twelve chains were built into the asymmetric unit and resulted in 49% solvent content. Figure 2B shows one layer of monomers in the unit cell, which consists of three unique protomers from the asymmetric unit and their

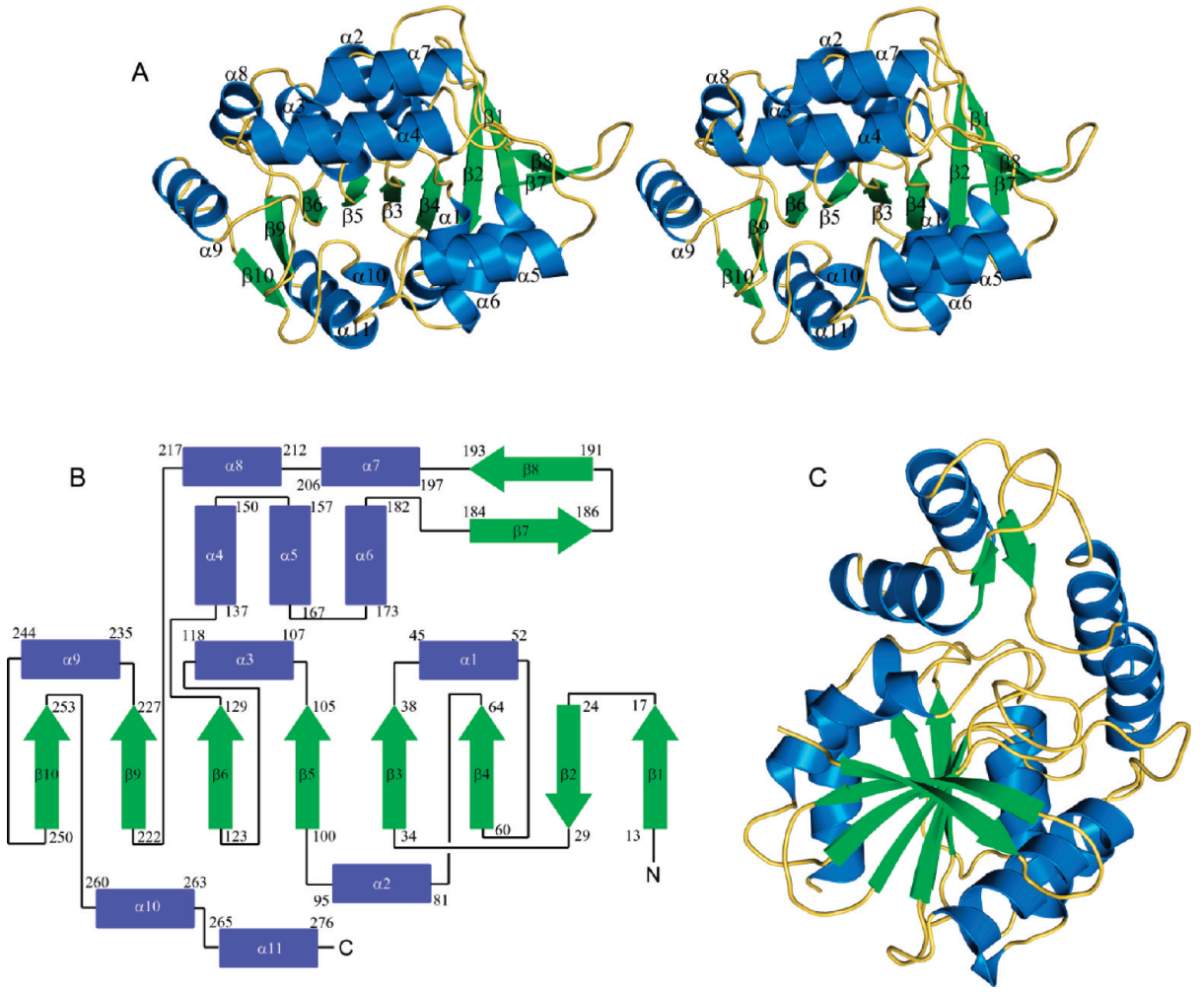


FIGURE 3: Monomeric structure of *E*-2AMS hydrolase. (A) Stereoview ribbon diagram of *E*-2AMS hydrolase with the secondary structures labeled. $\beta$-Strands are shown in green, and α-helices are shown in blue. (B) Topology diagram of *E*-2AMS hydrolase using the same color scheme as (A). (C) Ribbon diagram of *E*-2AMS hydrolase illustrating the sharp twist in the $\beta$-sheet running through the core of the enzyme and the cap domain.

crystallographic equivalents. The pseudotranslational symmetry occurs along the *B*-axis, where the spacing between protomers is nearly equal for two of the three pairs and nearly 10 Å further for the final spacing.

*Overall Structure of E-2AMS Hydrolase*. The structure of *E*-2AMS hydrolase was solved using SeMet SAD phasing at 2.7 Å resolution followed by phase extension with native data to 2.3 Å. Each of the 12 protomers within the asymmetric unit consists of residues 11−278; the first ten residues and the poly-histidine tag are not visible in the electron density. Figure 3 shows the final model of *E*-2AMS hydrolase. The overall fold of this enzyme indicates that *E*-2AMS hydrolase belongs to the α/$\beta$ hydrolase superfamily. The core of the enzyme is an eight-stranded mostly parallel $\beta$-sheet where $\beta$2 is the only strand running antiparallel ($\beta$1↑$\beta$2↓$\beta$4↑$\beta$3↑$\beta$5↑$\beta$6↑$\beta$9↑$\beta$10↑). The $\beta$-sheet has a very tight twist so that $\beta$10 is oriented nearly perpendicular to $\beta$1 when looking down the $\beta$-sheet. One side of the $\beta$-sheet is flanked by three α-helices, α2, α10, and α11. A bundle of α-helices, consisting of α1, α3, α8, and α9, packs on the other side of the $\beta$-sheet. A cap domain is inserted between $\beta$6 and $\beta$9 and sits at the N-terminal side of the $\beta$-sheet. This lid is made up of four α-helices (α4, α5, α6, and α7) and a small two-stranded antiparallel $\beta$-sheet ($\beta$7↑$\beta$8↓).

The 12 protomers of *E*-2AMS hydrolase in the asymmetric unit form six pairs of dimers with identical interfaces. The two protomers are related to each other by local 2-fold symmetry. The $\beta$-sheets of each chain face in the same direction but twist in opposite directions, as seen in Figure 4. The dimer interface involves the loop region between α9 and $\beta$10, residues from $\beta$10, and the C-terminal end of α10 from both protomers. Each dimer is about 95 Å long with a total surface area of nearly 20000 Å$^2$. Roughly 2000 Å$^2$ is buried at the dimer interface (*31*). Two hydrogen bonds are formed between the carbonyl oxygen and amide nitrogen atoms of the protein backbone at Val250 and Val250* (where * indicates the second protomer of the dimer). A salt bridge is formed between Arg243 and Glu228*, a hydrogen bond is found between Asn274 and Aps247*, and Arg124 forms interactions with the oxygen atoms of the carboxy terminus at Ala278*. Additionally, Phe275, Pro221, Leu223, Pro 249, Val250, Val251, and Ala271 from each protomer form an extended hydrophobic patch at the dimer interface.

*Active Site of E-2AMS Hydrolase*. The putative active site was identified through sequence and structural homology with other α/$\beta$ hydrolases and is located at the C-terminal end of the $\beta$-sheet and is shielded from solvent by the cap domain. Three positively charged residues, Arg146, Arg167, and Lys231, are present at the entrance to the active site. The active site pocket has two distinct environments. The narrow cleft adjacent to Ser106 is lined with primarily hydrophobic residues, including Ile41,

Leu207, Leu143, Leu232, and Phe131. The second half of the active site, however, is more hydrophilic. This pocket, roughly 10 Å deep and 6 Å wide, is filled with water molecules coordinated to hydrophilic residues. Arg179 forms hydrogen bonds to two water molecules and is oriented in the active site by Glu262. Three tyrosine residues, Tyr184, Tyr168, and Tyr259, are positioned so the hydroxyl group faces into the active site and forms hydrogen bonds with water molecules. Other hydrophilic residues in the active site include Thr42, Ser43, and His105.

The putative nucleophilic residue, Ser106, adopts a strained conformation at the C-terminal end of $\beta$5 and is hydrogen bonded to a structurally conserved water molecule, which is also coordinated to the backbone amide nitrogen atoms of Leu107 and Leu41. A chloride ion is bound in a pocket next to Ser106. The chloride ion is bound in an oxyanion hole, coordinated to Ser106 2.9 Å away, and is also forming interactions with two backbone amide nitrogen atoms from Ile41 and Leu107 (3.2 and 3.3 Å, respectively). His258, located after $\beta$10, is 3.0 Å from Ser106 and is itself positioned by Asp130, 2.5 Å away. Asp130 is found at the end of $\beta$6, and Ser230 occurs after $\beta$9. Ser230 forms a strong hydrogen bond with Asp130 (2.7 Å) and a second hydrogen bond with the amide nitrogen atom from the protein backbone of Leu232. These interactions are shown in Figure 5.

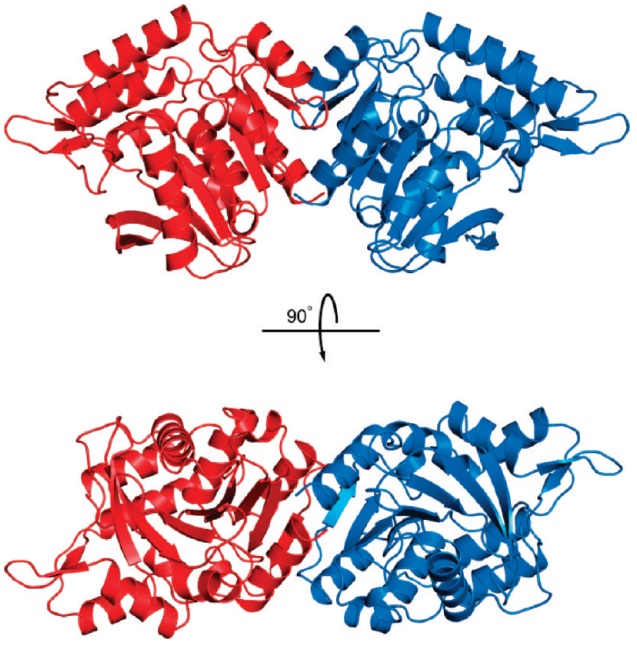


FIGURE 4: Dimeric structure of *E*-2AMS hydrolase. Each chain is colored differently, one red and the second blue. The second orientation was obtained by rotating the first orientation by 90° along the horizontal axis.

*Activity of E-2AMS Hydrolase Mutants*. The extinction coefficient at 261 nm for *E*-2AMS is 18600 $M^{-1}cm^{-1}$ as previously determined, and the $K_M$ and $k_{cat}$ values for *E*-2AMS hydrolase have also been reported (*7*, *10*, *32*). The activity of the active site mutants S106A, D130N, and S230C was assayed by measuring the decrease in the absorbance at 261 nm, the $\lambda_{max}$ for *E*-2AMS as previously described (*7*, *10*). Each of these mutants was found to be completely inactive relative to the native enzyme, which retained an activity comparable to published results. The active site mutants S230A and S230N could not be assayed because these proteins precipitated in the assay buffer. The inactive S106A mutant was then used for cocrystallization and soaking experiments with *E*-2AMS, which were unsuccessful in producing an enzyme−substrate complex.

## DISCUSSION

*Comparison of E-2AMS Hydrolase to Other Members of the α/β Hydrolase Superfamily*. The α/β hydrolase superfamily comprises a very large class of enzymes, catalyzing myriad reactions and related through similar tertiary structures. This family has been structurally well characterized and includes esterases, lipases, epoxidases, alkane dehalogenases, carbon−carbon bond hydrolases, amidohydrolases, and thioesterases (*33*, *34*). Previous analysis of the primary sequence for *E*-2AMS hydrolase using BLAST suggested that this enzyme was a member of the α/β hydrolase superfamily, despite a low sequence identity to this family of enzymes (*10*). A DALI search was performed using the structure of *E*-2AMS hydrolase, and these results confirmed that *E*-2AMS hydrolase is a member of the α/β hydrolase family. Representative members of this superfamily are shown in Table 3, and their similarity to *E*-2AMS hydrolase is given by their *Z* score, rmsd, sequence identity, and the number of residues aligned (*34*−*40*). The core of *E*-2AMS hydrolase is formed by an eight-stranded β-sheet flanked on both sides by α-helices with a strand order that agrees with the other members of this superfamily (*41*, *42*). The cap domain is

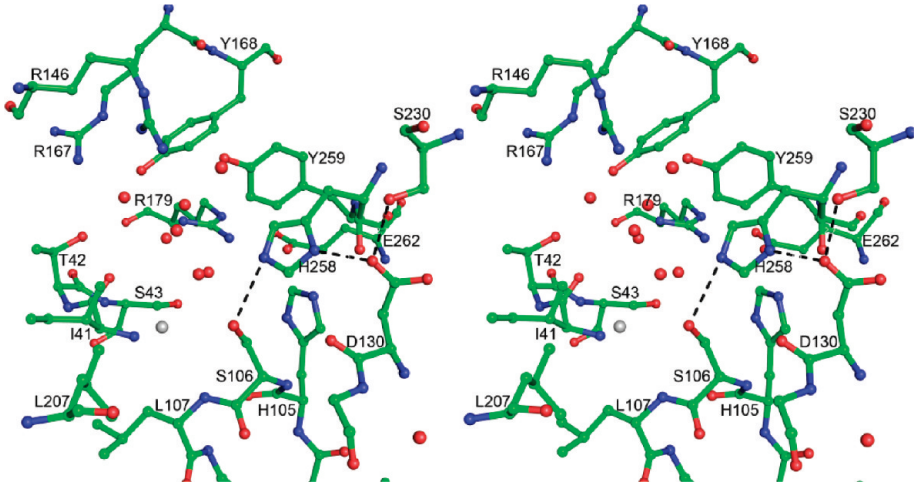


FIGURE 5: Stereoview ball and stick diagram of the active site of *E*-2AMS hydrolase. Water molecules are shown as red nonbonded spheres, and the chloride ion is shown as a nonbonded gray sphere. Key protein interactions are shown as dashed lines.

an insertion between $\beta$6 and $\beta$9, with a short two-stranded antiparallel $\beta$-sheet and four α-helices. The cap domain acts to shield the active site from solvent. While this second domain is not part of the α/$\beta$ hydrolase fold, insertions are common after the fifth $\beta$-strand of the core $\beta$-sheet (*42*). Additionally, the cap domain is not conserved. For example, the cap domains of the serine hydrolase CarC and the carbon−carbon hydrolase MhpC (*34*, *37*) are composed of three long α-helices, and the cap domain from an α/$\beta$ hydrolase of *Novosphingobium aromaticivorans* DSM 12444 has a three-stranded antiparallel $\beta$-sheet flanked by four α-helices (PBD ID 3BWX).

Table 3: Enzymes Identified as Structurally Similar to *E*-2AMS Hydrolase through DALI

| protein | PDB ID | *Z* score | rmsd | % identical | no. of aligned residues |
|---|---|---|---|---|---|
| α/$\beta$ hydrolase, YP_496220.1 | 3BWX | 30.4 | 2.4 | 23 | 263 |
| serine hydrolase (CarC) | 1J1I | 29.1 | 2.2 | 21 | 238 |
| chloroperoxidase L | 1A88 | 28.7 | 2.4 | 21 | 246 |
| arylesterase (PFE) | 1VA4 | 28.7 | 2.5 | 23 | 247 |
| esterase (EST) | 1ZOI | 28.5 | 2.5 | 21 | 247 |
| esterase (ybfF) | 3BF7 | 27.5 | 2.5 | 21 | 242 |
| fluoroacetate dehalogenase | 1Y37 | 27.3 | 2.7 | 24 | 246 |
| C−C bond hydrolase (MhpC) | 1U2E | 26.3 | 2.6 | 18 | 251 |

Figure 6 shows a sequence alignment of similar enzymes identified through DALI (*43*, *44*). The "nucleophilic elbow" begins at the C-terminal end of $\beta$5 in *E*-2AMS hydrolase and is identified using the fingerprint motif Gly-X-Nu-X-Gly, where X is any amino acid and Nu is the nucleophilic amino acid necessary for catalysis. In *E*-2AMS hydrolase, the nucleophile is Ser106. The sequence alignment identified few conserved residues, the majority being glycine residues, presumably important for proper folding. An absolutely conserved histidine residue required for catalysis in the α/$\beta$ hydrolase family is found after the last $\beta$-strand in the core domain. Additionally, the seventh $\beta$-strand of the eight-stranded $\beta$-sheet is mostly conserved, with hydrophobic residues forming interactions with neighboring $\beta$-strands.

*Active Site Comparison.* The α/$\beta$ hydrolase superfamily utilizes a catalytic triad consisting of a nucleophile, a strictly conserved histidine residue, and an acidic residue, usually an aspartate residue. The nucleophile, identified as Ser106, and the catalytic histidine residue, His258, are located in structurally conserved positions at the interface of the two domains of *E*-2AMS hydrolase. However, the third residue of the catalytic triad is not located at the C-terminal end of the seventh $\beta$-strand as expected. Instead, His258 forms a hydrogen bond with Asp130 found after $\beta$6 with Ser230 occupying the turn between $\beta$9 and α9. Figure 7 shows the superposition of the active site of *E*-2AMS hydrolase (green) with the esterase ybfF (white) from *E. coli*

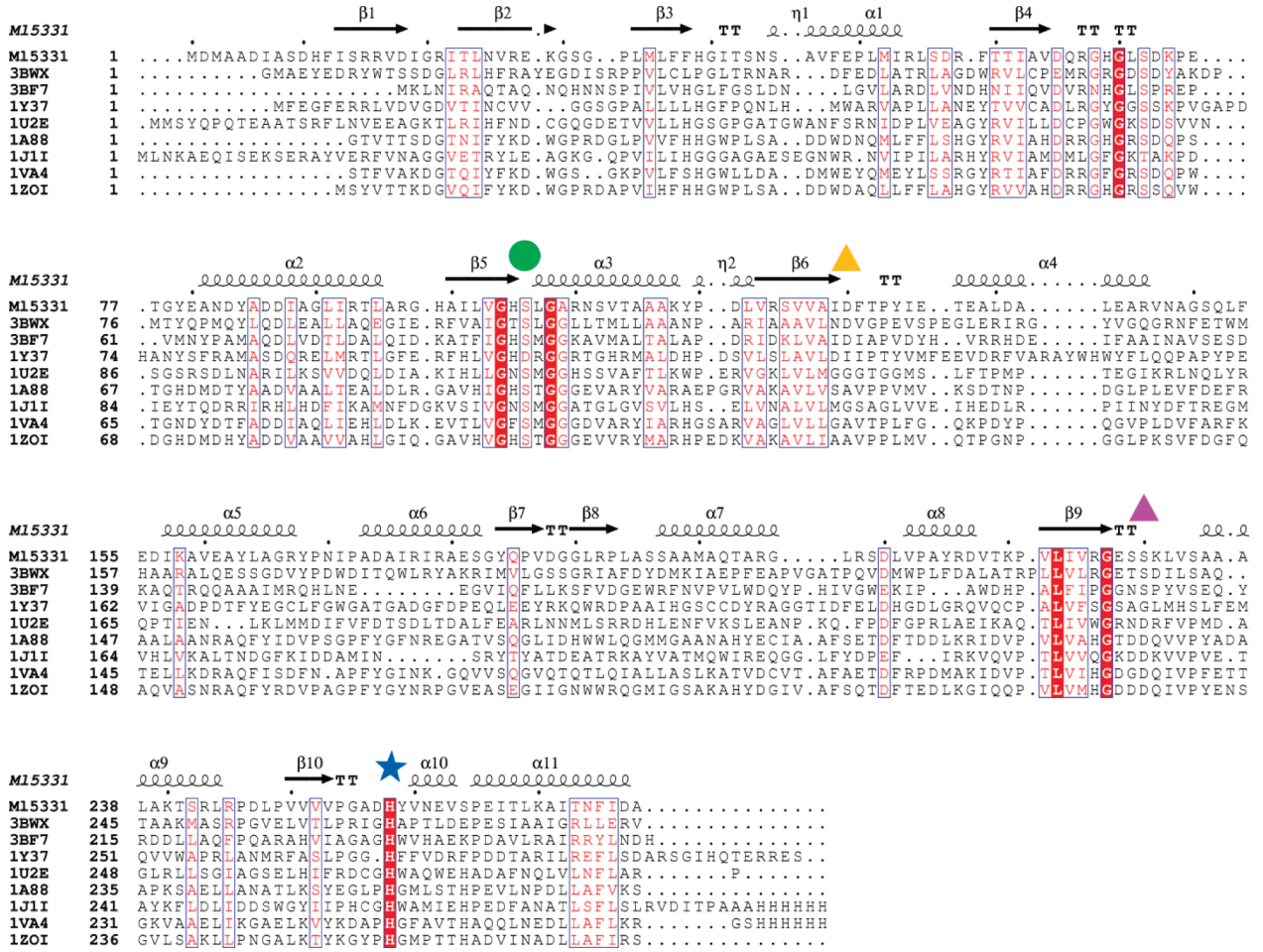


FIGURE 6: Sequence alignment of *E*-2AMS hydrolase with other enzymes identified as structurally homologous by DALI. Included enzymes are the α/$\beta$ hydrolase YP_496220.1 (3BWX), CarC (1J1I), chloroperoxidase L (1A88), the PFE arylesterase (1VA4), the EST esterase (1ZOI), ybfF (3BF7), fluoroacetate dehalogenase (1Y37), and MhpC (1U2E). The nucleophile is marked with a green circle, the catalytic histidine is labeled with a blue star, and the different active site acids are labeled using triangles. The expected acidic residues have a purple triangle, while the location of the acidic residue in *E*-2AMS hydrolase has a yellow triangle.

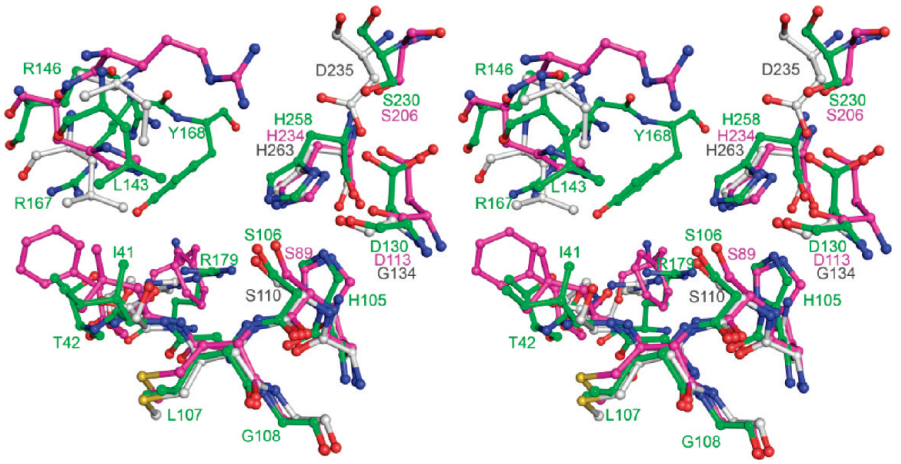


FIGURE 7: Stereoview ball and stick diagram comparing the active site of *E*-2AMS hydrolase to other members of the α/β hydrolase superfamily. The *E*-2AMS hydrolase structure is shown with green carbon atoms, ybfF is shown using magenta carbon atoms, and the carbon−carbon bond hydrolase MhpC has white carbon atoms. All residues of *E*-2AMS hydrolase are labeled, and the catalytic residues of ybfF and MhpC are labeled.

(PDB ID 3BF7) and the carbon−carbon bond hydrolase MhpC (purple) from *E. coli* (PDB ID 1U2E) (*34*, *40*). All three structures have an oxyanion hole formed by the amide nitrogen atoms of the protein backbone using the residue adjacent to the nucleophile (Leu107 in *E*-2AMS hydrolase) and a residue at the C-terminal end of the fourth β-strand (Ile41 at the end of β3 in *E*-2AMS hydrolase). The catalytic triads observed in *E*-2AMS hydrolase and ybfF are similar. The histidine responsible for activation of the nucleophile is 3.0 Å from Ser106 and 2.9 Å in the ybfF structure. The histidine residues in these two structures are each then coordinated to one oxygen atom of the carboxylate group of the acidic residue at the end of β6. The catalytic triad seen in MhpC, shown in purple in Figure 7, adopts the more common arrangement. The nucleophile, Ser110, is positioned 3.3 Å from the Nε2 atom of the side chain of His263. The aspartate residue responsible for activation of His263, Asp235, is coordinated to the Nδ1 atom of the histidine side chain using both oxygen atoms of the carboxylate side chain. Asp235 is located at the end of the seventh β-strand of the β-sheet. The remaining residues lining the active site are not conserved and show no sequence similarity. However, careful sequence analysis has identified several residues apart from the catalytic triad that can help to identify the reaction catalyzed and suggest *E*-2AMS hydrolase is quite similar to the esterase subclass of the α/β hydrolase superfamily (*45*, *46*). The residue preceding the active site nucleophile is a histidine, and two residues before the catalytic histidine is a hydrophobic residue in esterases (His105 and Ala261, respectively, in *E*-2AMS hydrolase). While the reaction catalyzed is a hydrolysis of an amide bond instead of an ester linkage, the requirements and chemistry are very similar. Additionally, the arrangement of active site residues in which the acidic residue is found at the end of β6 and a fourth catalytic residue, a serine at the C-terminal end of the seventh β-strand, is not uncommon in esterases (*40*).

*Mechanistic Implications for E-2AMS Hydrolase*. Previous characterization of *E*-2AMS hydrolase has suggested a possible mechanism for the hydrolysis of *E*-2AMS to form succinic semialdehyde, acetate, ammonia, and carbon dioxide, shown in Figure 8 (*7*). This mechanism utilizes an active site base to activate a water molecule for nucleophilic attack on the carbonyl carbon atom of the amide group. Collapse of the tetrahedral intermediate then leads to the loss of acetate. Tautomerization of intermediate **11** is followed by loss of carbon dioxide and release of ammonia through the addition of water. The last two steps could also occur in the reverse order. Direct addition of an active site nucleophile to the amide carbonyl carbon atom is also a possibility (Figure 8, mechanism 2). Direct attack by the active site nucleophile is most commonly observed in serine proteases, which are not members of the α/β superfamily but have a similar active site triad, composed of serine, histidine, and aspartate (*47*). However, an aryl esterase from *Pseudomonas fluorescens*, which has the α/β hydrolase fold, is believed to function through direct attack of the nucleophilic serine residue (*35*). A third possibility for the hydrolysis of *E*-2AMS to form intermediate **10** utilizes His258 to deprotonate water for attack on the carbonyl carbon, leading to loss of acetate and intermediate **11**. In this mechanism, Ser106 might stabilize the negatively charged tetrahedral intermediate found in the oxyanion hole. This general base mechanism has precedence; the carbon−carbon bond hydrolase MhpC from *E. coli* is believed to use His263 for activation of water and Ser110 for stabilization of a tetrahedral intermediate (*34*, *48*, *49*).

Although all attempts at cocrystallization or soaking *E*-2AMS into crystals of *E*-2AMS hydrolase or into crystals of the *E*-2AMS hydrolase S106A mutant were unsuccessful, examination of the active site suggests several residues that could play important roles in both the binding of *E*-2AMS and the catalysis of *E*-2AMS hydrolase. The catalytic triad, shown in Figure 5, consists of Ser106, His258, and Asp130 and most likely functions as seen in other α/β hydrolase enzymes (*42*). Briefly, Asp130 activates His258 for deprotonation of Ser106, which could act as a nucleophile for direct attack on the *E*-2AMS or activate a water molecule for attack on the substrate. Mutation of Ser106 to alanine abolished all catalytic activity, confirming this residue's critical importance for activity. The D130N mutant also showed no activity, indicating that the activation and stabilization of the catalytic histidine residue are necessary for catalysis. Ser230, found after β9 in the position most commonly occupied by the acidic residue of the catalytic triad, is more than 4 Å from His258 but forms a strong hydrogen bond (2.7 Å) with Asp130. Ser230 may play an important role in the structural integrity of the enzyme because the S230A, S230C, and S230N mutants were all unstable. It is possible that the slightly larger sulfur atom of cysteine in the S230C mutant could disrupt this sharp turn between β9 and α9. Other interactions involving Ser230 include two hydrogen bonds to amide nitrogen atoms of the protein backbone of Leu232 and Val233 in a β-turn.

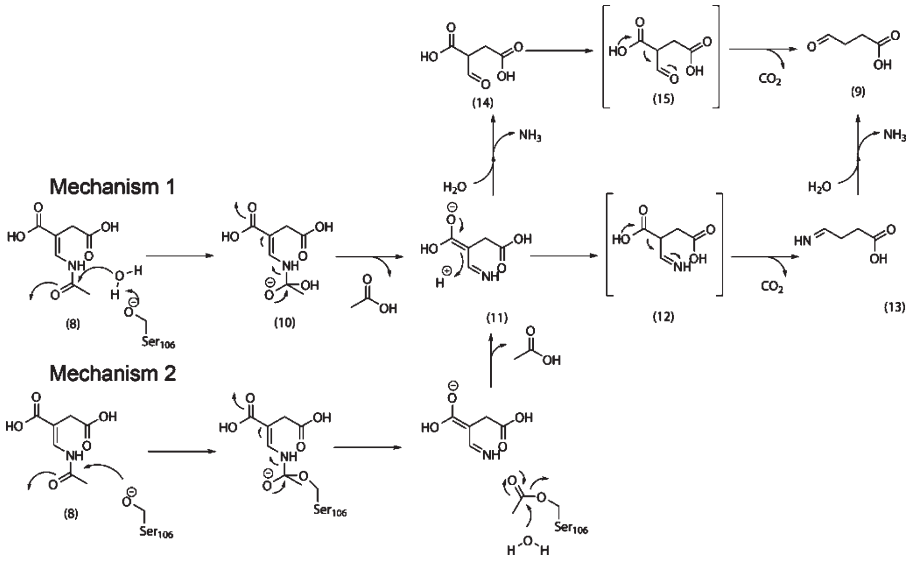


FIGURE 8: Proposed mechanisms for the hydrolysis of *E*-2AMS to produce succinic semialdehyde, acetate, ammonia, and carbon dioxide. Mechanism 1 utilizes Ser106 to activate a water molecule for attack at the carbonyl carbon of the amide bond, while mechanism 2 utilizes a direct attack on *E*-2AMS by Ser106.

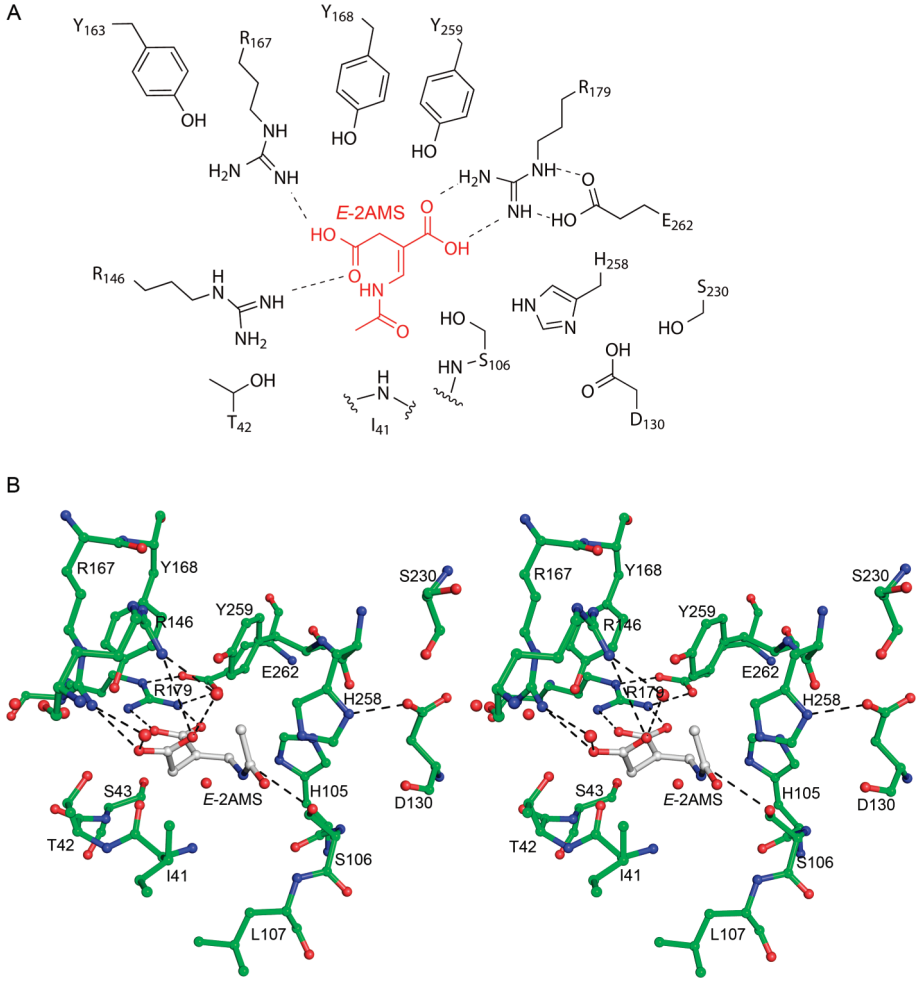


FIGURE 9: Possible coordination of *E*-2AMS within the active site. (A) Schematic diagram where *E*-2AMS (shown in red) was manually positioned such that either Ser106 or an activated water molecule could attack the carbonyl carbon of the amide group. Potential hydrogen bonds are shown as dashed lines. (B) Stereoview ball and stick diagram of the *E*-2AMS hydrolase active site with *E*-2AMS manually positioned in the active site. Carbon atoms of *E*-2AMS hydrolase are colored green, and carbon atoms of *E*-2AMS are colored gray. Possible interactions and hydrogen bonds important for active site orientation are shown in dashed lines.

The active site has several positively charged residues which could be involved in coordination to the carboxylate groups of *E*-2AMS (Figure 5). Three arginine residues, Arg146, Arg167, and Arg179, face into the active site. Additionally, Arg179 forms

two salt bridges with Glu262 to help orient Arg179 within the active site. Manual placement of the substrate, such that the carbonyl carbon of the amide group is located near Ser106 in the active site and the carbonyl oxygen points toward the oxyanion hole, suggests Arg146 and Arg179 are both likely candidates for coordination to the carboxylate groups and could play a role in the proper orientation of *E*-2AMS within the active site (Figure 9). Either arginine residue could also stabilize the enol intermediate and serve as a proton donor in the first mechanism. Protonation of the enol intermediate **11** could utilize Tyr166, Tyr259, or His105, either directly or using an activated water molecule. These same residues could also be involved in catalyzing the decarboxylation reaction leading to succinic semialdehyde.

## ACKNOWLEDGMENT

We are grateful to Dr. Yasunobu Ohkawa of the National Institute of Agrobiological Sciences, Ibaraki, Japan, for sharing the *M. loti* MAFF303099 strain with us. We thank Dr. Cynthia Kinsland and the Cornell Protein Facility for providing the clones and performing the site-directed mutagenesis. We also thank Leslie Kinsland for assistance in preparing the manuscript. Lastly, the assistance provided by the beamline staff scientists at NE-CAT during data collection is appreciated.

## REFERENCES

1. Eliot, A. C., and Kirsch, J. F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem. 73*, 383–415.
2. Snell, E. E., and Haskell, B. E. (1971) The Metabolism of Vitamin $B_6$, in Comprehensive Biochemistry, Vol. *21*, Elsevier/North Holland, New York.
3. Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S. (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res. 7*, 331–338.
4. Yuan, B., Yoshikane, Y., Yokochi, N., Ohnishi, K., and Yagi, T. (2004) The nitrogen-fixing symbiotic bacterium *Mesorhizobium loti* has and expresses the gene encoding pyridoxine 4-oxidase involved in the degradation of vitamin $B_6$. *FEMS Microbiol. Lett. 234*, 225–230.
5. Funami, J., Yoshikane, Y., Kobayashi, H., Yokochi, N., Yuan, B., Iwasaki, K., Ohnishi, K., and Yagi, T. (2005) 4-Pyridoxolactonase from a symbiotic nitrogen-fixing bacterium *Mesorhizobium loti:* cloning, expression, and characterization. *Biochim. Biophys. Acta 1753*, 234–239.
6. Ge, F., Yokochi, N., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2008) Gene identification and characterization of the pyridoxine degradative enzyme 4-pyridoxic acid dehydrogenase from the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti* MAFF303099. *J. Biochem. 143*, 603–609.
7. Mukherjee, T., Hilmey, D. G., and Begley, T. P. (2008) PLP catabolism: identification of the 2-(acetamidomethylene)succinate hydrolase gene in *Mesorhizobium loti* MAFF303099. *Biochemistry 47*, 6233–6241.
8. Mukherjee, T., Kinsland, C., and Begley, T. P. (2007) PLP catabolism: identification of the 4-pyridoxic acid dehydrogenase gene in *Mesorhizobium loti* MAFF303099. *Bioorg. Chem. 35*, 458–464.
9. Yokochi, N., Nishimura, S., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Identification of a new tetrameric pyridoxal 4-dehydrogenase as the second enzyme in the degradation pathway for pyridoxine in a nitrogen-fixing symbiotic bacterium, *Mesorhizobium loti*. *Arch. Biochem. Biophys. 452*, 1–8.
10. Yuan, B., Yokochi, N., Yoshikane, Y., Ohnishi, K., Ge, F., and Yagi, T. (2008) Gene identification and characterization of the pyridoxine degradative enzyme alpha-(N-acetylaminomethylene)succinic acid amidohydrolase from *Mesorhizobium loti* MAFF303099. *J. Nutr. Sci. Vitaminol.* (*Tokyo*) *54*, 185–190.
11. Yuan, B., Yokochi, N., Yoshikane, Y., Ohnishi, K., and Yagi, T. (2006) Molecular cloning, identification and characterization of 2-methyl-3-hydroxypyridine-5-carboxylic-acid-dioxygenase-coding gene from the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *J. Biosci. Bioeng. 102*, 504–510.
12. Mukherjee, T., McCulloch, K. M., Ealick, S. E., and Begley, T. P. (2007) Gene identification and structural characterization of the pyridoxal 5′-phosphate degradative protein 3-hydroxy-2-methylpyridine-4,5-dicarboxylate decarboxylase from *Mesorhizobium loti* MAFF303099. *Biochemistry 46*, 13606–13615.
13. Yoshikane, Y., Yokochi, N., Yamasaki, M., Mizutani, K., Ohnishi, K., Mikami, B., Hayashi, H., and Yagi, T. (2008) Crystal structure of pyridoxamine-pyruvate aminotransferase from *Mesorhizobium loti* MAFF303099. *J. Biol. Chem. 283*, 1120–1127.
14. McCulloch, K. M., Mukherjee, T., Begley, T. P., and Ealick, S. E. (2009) Structure of the PLP degradative enzyme 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase from *Mesorhizobium loti* MAFF303099 and its mechanistic implications. *Biochemistry 48*, 4139–4149.
15. Kress, D., Alhapel, A., Pierik, A. J., and Essen, L. O. (2008) The crystal structure of enamidase: a bifunctional enzyme of the nicotinate catabolism. *J. Mol. Biol. 384*, 837–847.
16. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248–254.
17. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol. 33*, 491–497.
18. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
19. Pape, T., and Schneider, T. R. (2004) HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Crystallogr. 37*, 843–844.
20. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D58*, 1772–1779.
21. Sheldrick, G. M. (2002) Macromolecular phasing with SHELXE. *Z. Kristallogr. 217*, 644–650.
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
23. Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D56*, 965–972.
24. Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D59*, 38–44.
25. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
26. Vellieux, F. M., and Read, R. J. (1997) Noncrystallographic symmetry averaging in phase refinement and extension. *Methods Enzymol. 277*, 18–53.
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.
28. DeLano, W. L. (2002), The PyMOL Molecular Graphics Systems, DeLano Scientific, San Carlos, CA.
29. Sambrook, J., Fritsch, G. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Guide, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Chook, Y. M., Lipscomb, W. N., and Ke, H. (1998) Detection and use of pseudo-translation in determination of protein structures. *Acta Crystallogr. D54*, 822–827.
31. Hasel, W., Hendrickson, T. F., and Still, W. C. (1988) A rapid approximation to the solvent-accessible surface areas of atoms. *Tetrahedron Comput. Methodol. 1*, 103–116.
32. Huynh, M. S., and Snell, E. E. (1985) Enzymes of vitamin $B_6$ degradation. Purification and properties of two N-acetylamidohydrolases. *J. Biol. Chem. 260*, 2379–2383.
33. Du, X., Wang, W., Kim, R., Yakota, H., Nguyen, H., and Kim, S. H. (2001) Crystal structure and mechanism of catalysis of a pyrazinamidase from *Pyrococcus horikoshii*. *Biochemistry 40*, 14166–14172.
34. Dunn, G., Montgomery, M. G., Mohammed, F., Coker, A., Cooper, J. B., Robertson, T., Garcia, J. L., Bugg, T. D., and Wood, S. P. (2005) The structure of the C-C bond hydrolase MhpC provides insights into its catalytic mechanism. *J. Mol. Biol. 346*, 253–265.
35. Cheeseman, J. D., Tocilj, A., Park, S., Schrag, J. D., and Kazlauskas, R. J. (2004) Structure of an aryl esterase from *Pseudomonas fluorescens*. *Acta Crystallogr. D60*, 1237–1243.
36. Elmi, F., Lee, H. T., Huang, J. Y., Hsieh, Y. C., Wang, Y. L., Chen, Y. J., Shaw, S. Y., and Chen, C. J. (2005) Stereoselective esterase from

*Pseudomonas putida* IFO12996 reveals α/β hydrolase folds for D-β-acetylthioisobutyric acid synthesis. *J. Bacteriol. 187*, 8470–8476.
37. Habe, H., Morii, K., Fushinobu, S., Nam, J. W., Ayabe, Y., Yoshida, T., Wakagi, T., Yamane, H., Nojiri, H., and Omori, T. (2003) Crystal structure of a histidine-tagged serine hydrolase involved in the carbazole degravdation (CarC enzyme). *Biochem. Biophys. Res. Commun. 303*, 631–639.
38. Hofmann, B., Tolzer, S., Pelletier, I., Altenbuchner, J., van Pee, K. H., and Hecht, H. J. (1998) Structural investigation of the cofactor-free chloroperoxidases. *J. Mol. Biol. 279*, 889–900.
39. Holm, L., and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res. 26*, 316–319.
40. Park, S. Y., Lee, S. H., Lee, J., Nishi, K., Kim, Y. S., Jung, C. H., and Kim, J. S. (2008) High-resolution structure of ybfF from *Escherichia coli* K12: a unique substrate-binding crevice generated by domain arrangement. *J. Mol. Biol. 376*, 1426–1437.
41. Heikinheimo, P., Goldman, A., Jeffries, C., and Ollis, D. L. (1999) Of barn owls and bankers: a lush variety of alpha/beta hydrolases. *Structure 7*, R141–R146.
42. Nardini, M., and Dijkstra, B. W. (1999) α/β hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol. 9*, 732–737.
43. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics 15*, 305–308.
44. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673–4680.
45. Li, C., Hassler, M., and Bugg, T. D. (2008) Catalytic promiscuity in the alpha/beta-hydrolase superfamily: hydroxamic acid formation, C−C bond formation, ester and thioester hydrolysis in the C−C hydrolase family. *ChemBioChem 9*, 71–76.
46. Li, C., Li, J. J., Montgomery, M. G., Wood, S. P., and Bugg, T. D. (2006) Catalytic role for arginine 188 in the C-C hydrolase catalytic mechanism for *Escherichia coli* MhpC and *Burkholderia xenovorans* LB400 BphD. *Biochemistry 45*, 12470–12479.
47. Hedstrom, L. (2002) Serine protease mechanism and specificity. *Chem. Rev. 102*, 4501–4524.
48. Li, C., Montgomery, M. G., Mohammed, F., Li, J. J., Wood, S. P., and Bugg, T. D. (2005) Catalytic mechanism of C-C hydrolase MhpC from *Escherichia coli*: kinetic analysis of His263 and Ser110 site-directed mutants. *J. Mol. Biol. 346*, 241–251.
49. Li, J. J., and Bugg, T. D. (2007) Investigation of a general base mechanism for ester hydrolysis in C-C hydrolase enzymes of the α/β-hydrolase superfamily: a novel mechanism for the serine catalytic triad. *Org. Biomol. Chem. 5*, 507–513.