Published in final edited form as:
Biochemistry. 2013 July 16; 52(28): . doi:10.1021/bi400567a.

Kinetic, Mutational, and Structural Analysis of Malonate Semialdehyde Decarboxylase from *Coryneform* bacterium strain FG41: Mechanistic Implications for the Decarboxylase and Hydratase Activities

Youzhong Guo[‖,§], **Hector Serrano**[‖,§], **Gerrit J. Poelarends**[‡], **William H. Johnson Jr.**[‖], **Marvin L. Hackert**[⊥], and **Christian P. Whitman**[‖,*]

[‖]Division of Medicinal Chemistry, College of Pharmacy, The University of Texas, Austin, Texas 78712-1074 [‡]Department of Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands [⊥]Department of Chemistry and Biochemistry, The University of Texas, Austin, Texas 78712-1074

Abstract

Malonate semialdehyde decarboxylase from *Pseudomonas pavonaceae* 170 (designated Pp MSAD) is in a bacterial catabolic pathway for the nematicide 1,3-dichloropropene. MSAD has two known activities: it catalyzes the metal-ion independent decarboxylation of malonate semialdehyde to produce acetaldehyde and carbon dioxide, as well as a low-level hydration of 2-oxo-3-pentynoate to yield acetopyruvate. The latter activity is not known to be biologically relevant. Previous studies identified Pro-1, Asp-37, and a pair of arginines (Arg-73 and Arg-75) as critical residues in these activities. MSAD from *Coryneform* bacterium strain FG41 (designated FG41 MSAD) shares 38% pairwise sequence identity with the *Pseudomonas* enzyme including Pro-1 and Asp-37. However, Gln-73 replaces Arg-73, and the second arginine is shifted to Arg-76 by the insertion of a glycine. In order to determine how these changes relate to the activities of FG41 MSAD, the gene was cloned and the enzyme expressed and characterized. The enzyme has a comparable decarboxylase activity, but a significantly reduced hydratase activity. Mutagenesis along with crystal structures of the native enzyme (2.0 Å resolution) and the enzyme modified by a 3-oxopropanoate moiety (resulting from the incubation of enzyme and 3-bromopropiolate) (2.2 Å resolution) provided a structural basis. The roles of Pro-1 and Asp-37 are likely the same as those proposed for MSAD. However, the side chains of Thr-72, Gln-73, and Tyr-123 replace those of Arg-73 and Arg-75 in the mechanism and play a role in binding and catalysis. The structures also show that Arg-76 is likely too distant to play a direct role in the mechanism. FG41 MSAD is the second functionally annotated homologue in the MSAD family of the tautomerase superfamily and could represent a new subfamily.

[*]Corresponding author (CPW) Tel: 5127471-6198; Fax: 512-232-2606; whitman@austin.utexas.edu..
[§]The authors contributed equally to this work

The atomic coordinates and structure factors have been deposited with the Brookhaven Protein Data Bank (PDB codes 3MJZ and 3MLC).

SUPPORTING INFORMATION

The experimental procedures used for the construction of the FG41 MSAD mutants (P1A, D37N, Q73A, Q73R, R76A, and Y123F) and the R73Q mutant of Pp MSAD are provided in the Supporting Information. In addition, protocols used for the expression and purification of wild-type enzymes and the seven mutant proteins are reported in the Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

The soil bacteria *Pseudomonas pavonaceae* 170 converts the nematicide 1,37 dichloropropene (**1**, Scheme 1) to acetaldehyde in five enzyme-catalyzed steps (Scheme 1).[1-4] In the first three enzyme-catalyzed steps, the isomeric mixture of **1** is converted to the *cis* and *trans* isomer of 3-chloroacrylic acid (**2** and **3**, respectively). Subsequent hydrolytic dehalogenation by *cis*- or *trans*-3-chloroacrylic acid dehalogenase (*cis*-CaaD and CaaD, respectively) produces **4**.[4-6] Decarboxylation of **4** by malonate semialdehyde decarboxylase (designated Pp MSAD) affords acetaldehyde (**5**), which is likely routed to the Krebs cycle.[7] Interestingly, Pp MSAD, *cis*-CaaD, and CaaD represent three of the five known families in the tautomerase superfamily, which is a group of proteins constructed from a β–α–β building block and characterized by a catalytic amino-terminal proline.[8-10] Extensive mechanistic and structural studies have been carried out on these three enzymes in order to determine the evolutionary basis for their presence in the same pathway.

As a result of these studies, the properties of MSAD from *P. pavonaceae* 170 are well characterized. The enzyme catalyzes a metal-ion independent decarboxylation of **4** to afford acetaldehyde and carbon dioxide.[7] In addition to its biological activity, the homotrimeric Pp MSAD has a hydratase activity using three acetylenic compounds, 2-oxo-3-pentynoate (**6**, Scheme 2) and 3-bromo- and 3-chloropropiolate (**7** and **8**, Scheme 3).[11-13] These activities occur at the same active site, but do not have known biological roles. The first compound is converted to acetopyruvate (**9**, Scheme 2), whereas the 3-halopropiolates are transformed into acylating agents (**10** or **11**, Scheme 3) that inactivate the enzyme by the covalent modification of Pro-1 **(**i.e., **12)**.[11,12]

Crystal structures of Pp MSAD and the inactivated enzyme (modified at the prolyl nitrogen by a 3-oxopropanoate adduct, i.e., **12**) coupled with mutagenesis studies are the primary basis for the working hypotheses for both the decarboxylase and hydratase activities.[7,11,13] The key residues for these activities are Pro-1, Asp-37, and a pair of arginines (Arg-73 and Arg-75). Roles for these residues have been assigned based primarily on their interactions with the covalent adduct (i.e., **12**) in the crystal structure of the covalently modified Pp MSAD.[13]

MSAD from *Coryneform* bacterium strain FG41 (designated FG41 MSAD) has not been previously studied. The enzyme shares 38% sequence identity (65% similarity) with the *Pseudomonas* enzyme and includes three of the four key residues, Pro-1, Asp-37, and Arg-76 (shifted to position 76 by the insertion of a glycine). The enzyme is part of a pathway for the degradation of **2**, and the genomic context suggests it is a malonate semialdehyde decarboxylase (unpublished results, G.J. Poelarends, H. Serrano, and C.P. Whitman, 2009). However, a glutamine replaces the “missing” Arg-73. These observations prompted us to clone and express the FG41 MSAD gene and characterize the properties of the gene product. It was found that FG41 MSAD has a comparable decarboxylase activity to that of the *Pseudomonas* enzyme, but the hydratase activity (using **6**) is significantly reduced. Mutagenesis experiments showed that Pro-1, Asp-37, and Arg-76 are essential for decarboxylase activity, whereas Gln-73 is important, but not essential. Crystal structures of the native FG41 MSAD and the 3-oxopropanoate-modified enzyme suggest roles for Pro-1, Asp-37, and Gln-73, and identified Tyr-123 and Thr-72 as additional participants in the mechanism. Subsequent mutagenesis confirmed the importance of Tyr-123 for decarboxylase activity. The hydratase activity requires Pro-1, Asp-37, and Arg-76, but is less dependent on Gln-73 and Tyr-123. The combined results suggest reaction mechanisms for both activities and will assist in the future identification of family members.

EXPERIMENTAL PROCEDURES

Materials

Chemicals, biochemicals, buffers, column resins, and solvents were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO), Fisher Scientific Inc. (Pittsburgh, PA), Fluka Chemical Corp. (Milwaukee, WI), or EM Science (Cincinnati, OH), unless stated otherwise. Literature procedures were used for the synthesis of 2-oxo-3-pentynoate (**6**)[14] and 3-bromopropiolate (**7**).[15,16] Enzymes, reagents, and kits used for molecular biology procedures were obtained from New England Biolabs, Inc. (Ipswich, MA) or F. Hoffmann-La Roche, Ltd. (Basel, Switzerland). The protease inhibitor cocktail tablets were purchased from F. Hoffmann-La Roche. The sources for the components of Luria-Bertani (LB) media are reported elsewhere.[17] The QuikChange Mutagenesis Kit was purchased from Agilent Technologies (Santa Clara, CA). The Amicon concentrator and the YM10 ultrafiltration membranes were obtained from Millipore Corp. (Billerica, MA). The Econo-Column® chromatography columns (1 × 10 cm) were obtained from BioRad (Hercules, CA). Oligonucleotides for DNA amplification and sequencing were synthesized by Genosys (The Woodlands, TX), Integrated DNA Technologies (Coralville, IA), or Invitrogen (Carlsbad, Ca).

Bacterial Strains, Plasmids, and Growth Conditions

Coryneform bacterium strain FG41 was obtained from Professor Dick B. Janssen (Department of Biochemistry, University of Groningen, The Netherlands). *Escherichia coli* strain BL21-Gold(DE3) (Agilent Technologies) was used in combination with the pET3b vector for expression of wild type FG41 MSAD, the FG41 MSAD mutants, and the Pp MSAD mutant. *E. coli* cells were grown at 30 °C (for protein expression) or 37 °C (for plasmid preparation) in LB media supplemented with ampicillin (Ap, 100 μg/mL), as indicated.

General Methods

Protein was analyzed by SDS-PAGE under denaturing conditions on gels containing 15% polyacrylamide.[18] The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the method of Waddell.[19] Kinetic data were obtained on a Hewlett Packard 8452A Diode Array spectrophotometer or an Agilent 8453 UV-Visible spectrophotometer. The kinetic data were fitted by nonlinear regression data analysis using the Grafit program (Erithacus Software Ltd., Horley, U.K.) obtained from Sigma Chemical Co. Techniques for restriction enzyme digestion, ligation, transformation, and other standard molecular biology manipulations were based on methods described elsewhere.[17] The PCR was carried out in a Perkin-Elmer DNA thermocycler Model 480 obtained from PerkinElmer Inc. (Wellesley, MA). DNA sequencing was performed at the DNA core facility in the Institute for Cellular and Molecular Biology (ICMB) at the University of Texas at Austin. Mass spectral data were obtained on an LCQ electrospray ion-trap mass spectrometer (Thermo, San Jose, CA) in the ICMB Protein and Metabolite Analysis Facility. Samples were prepared as described elsewhere.[5] A BLAST search of the National Center for Biotechnology Information (NCBI) databases was performed using the FG41 MSAD and Pp MSAD amino acid sequences as the query sequences.[20] The search identified 500 sequences still at high significance for each query. For each set, three of the most closely related sequences and the query were subjected to multiple sequence alignment using CLUSTAL W.[21] The ^{1}H NMR experiments to verify product formation for the decarboxylation of **4** and the hydration of **6** were carried out on a Varian Unity INOVA-500 spectrometer using the protocols described previously.[5-7]

Construction of the FG41 MSAD Expression Vector

The FG41 MSAD gene was cloned from the genomic DNA of *Coryneform* bacterium strain FG41 using the following primers. The forward primer, designated primer F1, 5′-ATA**CATATG**CCTCTCATCCGCATCGATCTGACCTCGGATCGCTCC-3′, contains an *Nde*I restriction site (in bold) followed by 36 bases corresponding to the coding sequence of the FG41 MSAD gene. The reverse primer, designated primer R1, 5′-CAT**GGATCC**TCAGGCTGCTCCGGTGGCGGGGATCGCGAGTTCCCCGGTGACG TACTGGGCGGACCCGAAGCCGAA-3′, contains a *Bam*HI restriction site (in bold) followed by 66 bases corresponding to the complementary sequence of the gene with the exception of the underlined base. The underlined base is a silent mutation that results in the deletion of an internal *Bam*HI restriction site. The genomic DNA was isolated by a phenol extraction procedure described elsewhere.[22] The amplification mixtures (100 μL) contained the appropriate synthetic primers (100 ng each), the deoxynucleotide triphosphates (dNTPs) (200 μM each), genomic DNA (25-50 ng), *Taq* DNA polymerase (2 units) and the accompanying buffer. The product was purified by electrophoresis on an 0.8% agarose gel. The PCR amplification protocol consisted of an initial 10-min denaturation cycle at 94° C, followed by 35 cycles of 94 °C for 1 min, 58 °C for 1 min, and 72 °C for 3 min, and a 10-min elongation cycle at 72 °C. The reaction mixtures were purified by electrophoresis using 0.8% agarose gels. The resulting PCR product and the pET3b vector were digested with *Nde*I and *Bam*HI restriction enzymes, purified, and ligated using T4 DNA ligase. An aliquot of the ligation mixture was transformed into competent *E. coli* BL21-Gold(DE3) cells. Transformants were selected at 37 °C on LB/Ap plates (100 μg/mL). Plasmid DNA was isolated from several randomly selected colonies and analyzed by restriction analysis for the presence of the insert. The cloned FG41 MSAD gene was sequenced to verify that no other mutations had been introduced during the amplification of the gene. The newly constructed expression vector was named pET(FG41MSAD).

Construction of the FG41 MSAD and Pp MSAD Mutants and Expression and Purification of Wild-type and Mutant Proteins

The experimental procedures used for the construction of the FG41 MSAD mutants (P1A, D37N, Q73A, Q73R, R76A, and Y123F) and the R73Q-Pp MSAD mutant are provided in the Supporting Information. In addition, protocols used for the expression and purification of the wild-type enzymes and the seven mutant proteins are reported in the Supporting Information.

Enzyme Assays

The decarboxylation of malonate semialdehyde (**4**) was monitored by following the conversion of NADH to NAD^+, in a coupled assay, as described elsewhere.[7] The assay mixtures (total volume of 1 mL) were made up in 20 mM K_2HPO_4 buffer (pH 9.0) and contained dithiothreitol (0.1 mM), NADH [5 μL of a 44 mg/mL stock solution in 100 mM Na_2HPO_4 buffer (pH 9.0)], alcohol dehydrogenase [10 YL of a 30 mg/mL stock solution in 100 mM Na_2HPO_4 buffer (pH 9.0)], and *cis*-3-chloroacrylic acid [**2**, 10 μL from a 100 mM stock solution made up in 100 mM Na_2HPO_4 buffer (pH 9.0)]. Malonate semialdehyde was generated by the addition of 10 μL of a 13 mg/mL stock of *cis*-CaaD. When the reaction was complete (5 min), an aliquot of the appropriate enzyme (FG41 MSAD, its mutants, Pp MSAD, or the R73Q mutant of Pp MSAD) was added [~0.1 mg, 10 μL of a 10 mg/mL stock solution in 10 mM Na_2HPO_4 buffer (pH 8.0)]. The rate was monitored over a period of 60 s by following the decrease in absorbance at 340 nm ($\varepsilon = 6220\ M^{-1}\ cm^{-1}$). One unit of enzyme activity is defined as the amount of enzyme required to convert 1 μmol of substrate to product in 1 min.[7]

The hydration of 2-oxo-3-pentynoate (**6**) was monitored by following the formation of acetopyruvate (**9**) at 294 nm (ε = 7000 $M^{-1}cm^{-1}$) in 20 mM Na_2HPO_4 buffer (pH 9.0).[5] An aliquot of Pp MSAD (2 μL of a 10.7 mg/mL solution, 0.06 μM), the R73Q mutant of Pp MSAD (200 μL of a 9.3 mg/mL solution, 6.6 μM), FG41 MSAD (10 μL of a 14.87 mg/mL solution, 10.22 μM) or the Q73R mutant of FG41 MSAD (0.5 μL of a 6.8 mg/mL solution, 0.241 μM) was diluted into buffer (20 mL) and incubated for 1 h. Subsequently, a 1 mL-portion of the diluted enzyme was transferred to a cuvette and assayed by the addition of a small quantity of **6** from either a 10 or 100 mM stock solution. The stock solution (100 mM) was made by dissolving the appropriate amount of **6** in 100 mM Na_3PO_4 buffer. The final pH of the stock solution was adjusted to 9.0. The 10 mM stock solution was made by dilution of an aliquot of the 100 mM stock into 100 mM Na_2HPO_4 buffer. The concentrations of **6** used in the assay ranged from 0.05 to 12 mM.

Inactivation of FG41 MSAD by 7 for Mass Spectral and Crystallographic Analysis

The enzyme (20 μM based on the molecular mass of the native enzyme) was incubated with a large excess of 3-bromopropiolate (**7**) (0.58 mM) in 1 mL of 20 mM NaH_2PO_4 buffer (pH 7.3) for 28 h at 4°C. In a separate control experiment, the same quantity of enzyme was incubated without inhibitor under otherwise identical conditions. The sample treated with **7** had no activity, whereas the control sample retained full decarboxylase activity. Subsequently, the two samples were treated as described elsewhere and analyzed by electrospray ionization mass spectrometry (ESI-MS).[5]

For the crystallographic analysis, stock solutions of FG41 MSAD (25 mg/mL) and **7** (45 mg of inhibitor in 10 mL of 10 mM TRIS-SO_4 buffer, pH 7.0-8.0, 30 mM) were made up. Aliquots of the FG41 MSAD (2 mL) and the inhibitor solutions (1 mL) were combined and incubated at 4 °C. The sample was processed through a PD-10 Sephadex G-25 column to remove unbound inhibitor.

Crystallization and Determination of the Ligand-free FG41 MSAD Structure

The initial crystallization screens were determined in a 96-well plate using sitting drops with Hampton Screens I and II and a home-made pH grid screen at both room temperature and 4 °C. A portion (100 μL) of the screen solution was placed in the large reservoir at the bottom of each well and sitting drops containing a mixture of 3 μL protein sample and 3 μL of the screen solution were equilibrated against the well solution. Final crystals were obtained after five months at 4 °C from 6-μL sitting drops consisting of equal amounts of precipitant solution [0.1 M sodium HEPES buffer, pH 7.5, 2% (v/v) polyethylene glycol 400, 2.0 M $(NH_4)_2SO_4$] and protein solution (20.5 mg/mL in 10 mM NaH_2PO_4 buffer, pH 8). The rod shaped, single crystals grew to about 0.25 × 0.25 × 0.4 mm.

Diffraction data to 2.0 Å resolution were collected in-house using a Rigaku RU200H rotating anode X-ray source with Cu K_α radiation and equipped with a R-AXIS IV++ image plate detector with a crystal-to-detector distance of 150 mm. Diffraction images (720) were collected over a total rotation of 360 degrees. The data were indexed, integrated and scaled using the HKL2000 software package.[23] The crystal belongs to the space group $P2_12_12_1$ with cell parameters a = 88.96, b = 94.69 and c = 190.70 Å. The asymmetric unit of FG41 MSAD contains twelve monomers (4 complete trimers), each composed of 136 amino acid residues, corresponding to a calculated Matthews' coefficient of 2.30 $Å^3$/Da and an estimated solvent content of 47%.[24] The *Pseudomonas* ligand-free MSAD (PDB code 2AAL) was used as the search model for molecular replacement using MOLREP[25] or the on-line server, CaspR.[26] Rotation and translation functions were calculated using data between 15 and 3 Å resolution, yielding the position and orientations of all twelve monomers in the asymmetric unit. Refinement of the top seven solutions by AMORE

resulted in a correlation coefficient of 0.51 and an R-Factor of 0.52.[27] After a round of automatic CNS refinement, a new set of coordinates resulted in an R_{test} of 0.51 and an R_{work} = 0.41. Examination of the molecular packing within the unit cells with the Program O showed that one trimer overlapped with another trimer and a large area of electron density present within the unit cell with no molecules.[28] This may account for the relatively large values of R_{test} and R_{work}. After manual model building of the initial model, twelve monomers were placed that occupy most of the electron density in the asymmetric unit. Iterative cycles of refinement with CNS and REFMAC5,[29,30] manual model building with the Program O and the addition of water molecules, followed by final rounds of refinement using Phenix,[31] resulted in a final structural model. A summary of the refinement statistics and geometric quality of the model is given in Table 1.

Crystallization and Structure Determination of Inactivated FG41 MSAD

The initial crystallization screens were performed as described above. The best crystals for inactivated FG41 MSAD were obtained at 4 °C from 6-μL hanging drops consisting of equal amounts of precipitant solution [0.1 M TRIS buffer, pH 8.5, 2.0 M $(NH_4)H_2PO_4$] and protein solution (24.5 mg/mL in 10 mM sodium phosphate buffer, pH 8). The diamond-shaped crystals grew to about 0.4 × 0.4 × 0.4 mm within one week. Diffraction data to 2.2 Å resolution were collected in-house using a Rigaku MicroMaX™-007HF rotating anode X-ray source with Cu K_α radiation and equipped with a Mar-345 image plate detector using a crystal-to-detector distance of 200 mm. Diffraction images (720) were collected over a total rotation of 360 degrees. The crystal belongs to the space group $P2_13$ with cell parameters a = 144.23 Å. The asymmetric unit contains five monomers (one complete trimer and two other subunits that form complete trimers using the crystallographic 3-fold symmetry, each subunit with 136 amino acid residues) corresponding to a calculated Matthews' coefficient of 3.44 $Å^3$/Dalton and an estimated solvent content of 64%.[24] The ligand-free FG41 MSAD was used as the search model for molecular replacement with the on-line server, CaspR.[26] Rotation and translation functions were calculated using data between 15 and 3 Å resolution, yielding the position and orientations of the five monomers in the asymmetric unit. Refinement of the top solutions by AMORE[27] resulted in a correlation coefficient of 0.62. After a round of automatic CNS refinement,[29] a new set of coordinates resulted in R_{test} = 0.39 and R_{work} = 0.32. A final structural model was obtained using CNS, REFMAC5, and the Program O, followed by final rounds of refinement using Phenix, as described above.[28-31] A summary of the refinement statistics and geometric quality of the model is given in Table 1.

RESULTS

Sequence Analysis of FG41 MSAD and Pp MSAD Homologues

A sequence similarity search in the NCBI database was performed with the BLAST program using the FG41 MSAD and the *Pseudomonas* MSAD amino acid sequences as the query sequences.[20,21] The search resulted in several related sequences with the eight most closely related ones shown in Figure 1. The sequences are divided into those related to FG41 MSAD (PDB Code 3MJZ) and those related to *P. pavonaceae* 170 MSAD (PDB Code 2AAJ). The first set (FG41 MSAD and homologues from *M. abscessus* ATCC 19977, *M. massiliense* CCUG 48898, and *Saccharopolyspora erythraea* NRRL 2338) shows Pro-1, Asp-37, and Arg-76, and the replacement of Arg-73 with Gln-73. The insertion of Gly-75 shifts arginine to position 76. The second set (the *P. pavonaceae* 170 MSAD, and homologues from *Pseudomonas* sp. GM67, *Mycobacterium colombiense* CECT 3035, and *Lactobacillus casei* strain BL23) shows the conservation of Pro-1, Asp-37, Arg-73, and Arg-75. Five other amino acids are intriguing because they are conserved only in the Pp MSAD or FG41 MSAD homologues. Tyr-39, Thr-55, Ser-72, Tyr-84, and Phe-123 in the

Pseudomonas MSAD (and related homologues) are replaced respectively with Phe-39, Ala-55, Thr-72, Phe-85, and Tyr-123 in the FG41 MSAD (and related homologues). (There are other amino acids conserved in Pp MSAD or FG41 MSAD, but the differences are not significant, involving the substitution of valine for leucine, as one example, or they are outside the active site.) Hence, kinetic, mutagenesis, and crystallography studies were carried out to examine the consequences of substitutions at these sites.

Expression, Purification, and Characterization of the FG41 MSAD and Mutants

FG41 MSAD and the mutants were purified in a three-step protocol (anion exchange, hydrophobic, and gel-filtration chromatography). Typically, this procedure yielded ~50 mg of homogeneous protein (as assessed by SDS-PAGE) per liter of cell culture. The purified proteins were analyzed gel filtration and electrospray ionization mass spectrometry (ESI-MS). In all cases, the enzymes eluted as trimers. ESI-MS analysis showed that the initiating N-formylmethionine has been removed so that the amino-terminal group is Pro-1. This observation is consistent with all previous results on tautomerase superfamily members.[5-7] The initiating methionine is removed by a methionyl aminopeptidase, and removal is correlated with the amino acid in the second position.[32]

Product Confirmation for the FG41 MSAD-catalyzed Reactions

It has previously been shown that Pp MSAD catalyzes the decarboxylation of **4** to afford **5** and the hydration of **6** to afford acetopyruvate (**9**).[7] ^{1}H NMR analysis of an incubation mixture containing **3**, CaaD, and FG41 MSAD showed signals corresponding to acetaldehyde (**5**) and its hydrate (data not shown).[5-7] (The rate of decarboxylation is minimal in the absence of enzyme.[7]) Likewise, ^{1}H NMR analysis of an incubation mixture containing **6** and FG41 MSAD showed signals corresponding to **9**, its hydrate, and enol (data not shown).[5-7] This analysis shows that the products of the reactions using FG41 MSAD are the same as those using Pp MSAD.

Characterization of the Decarboxylase and Hydratase Activities of FG41 MSAD and Mutants

The specific activity of FG41 MSAD for decarboxylation varied from 25,000-36,000 units, depending on the preparation (Table 2). The P1A, D37N, and R76A mutants had little detectable decarboxylase activity beyond that of non-enzymatic decarboxylation, reflecting the essential nature of these residues. The Q73A mutant retained 8.4% of the wild type activity, implicating Gln-73 in the decarboxylase activity as well. Interestingly, replacing Gln-73 in FG41 MSAD with an arginine results in substantially higher decarboxylase activity compared to the activity observed for the Q73A mutant (~61% of wild type FG41 MSAD). Moreover, replacing Arg-73 in Pp MSAD with a glutamine decreased decarboxylase activity substantially (~2% of wild type). Finally, replacing Tyr-123 with a phenylalanine decreased the decarboxylase activity of FG41 MSAD to 6% of wild type. This analysis implicates Pro-1, Asp-37, and Arg-76 as essential residues for the decarboxylase activity. Gln-73 and Tyr-123 are also critical, but mutagenesis does not lead to a complete loss of activity. The presence of Arg-73 in both FG41 MSAD and MSAD clearly results in an efficient decarboxylase.

As previously reported, the hydratase activity of Pp MSAD (measured by the conversion of 2-oxo-3-pentynoate to acetopyruvate, **6** to **9** in Scheme 2) falls between that of CaaD and *cis*-CaaD (Table 3).[7,11] The k_{cat}/K_m for the FG41 MSAD-catalyzed hydration reaction is reduced substantially from that of Pp MSAD (~56-fold, Table 3), and is comparable to that of *cis*-CaaD. The reduced efficiency is due to a 20-fold reduction in k_{cat} and a 2.4-fold increase in K_m. However, replacing Gln-73 with an arginine (in FG41 MSAD) increases k_{cat} 7-fold and decreases the K_m slightly. The net effect is a 9-fold increase in the catalytic

efficiency. The R73Q mutant of Pp MSAD has no detectable activity (data not shown). The Y123F mutant of FG41 MSAD results in a 4-fold increase in k_{cat} and a 3-fold increase in K_m. The P1A, D37N, and R76A mutants of FG41 MSAD show no detectable activity. The Q73A-FG41 MSAD appears to show hydratase activity comparable to that of wild type, but consistent kinetic data could not be obtained. This analysis indicates that Pro-1, Asp-37, and Arg-76 are critical for activity. Tyr-123 is important, but less so. The presence of Arg-73 results in a more efficient hydratase.

Mass Spectral Analysis FG41 MSAD Inactivated by 7

FG41 MSAD was incubated with **7**, and the inactivated protein was isolated and analyzed by ESI-MS. A control sample containing only FG41 MSAD was processed and analyzed similarly. (The control sample retained full activity.) Mass spectral analysis of the FG41 MSAD control sample showed a major peak corresponding to a mass of 14545 ± 2 Da and a minor peak of 14058.0 ± 2 Da. This latter signal could not be identified and appeared in all preparations (despite various treatments). Mass spectral analysis of FG41 MSAD incubated with **7** showed two major peaks corresponding to masses of 14630 ± 2 and 14586 ± 2 Da. The mass of the first species (i.e., 14630 Da) is in agreement with that expected for the enzyme modified with a 3-oxopropanoate group (+85 Da), the adduct resulting from the enzyme-catalyzed hydration of **7** (i.e., **12** in Scheme 3). This observation is consistent with the mass spectral analysis of CaaD, *cis*-CaaD, and Pp MSAD inactivated by **7**,[5-7] which are also modified by a molecule with a mass of 86 Da. The species with a mass of 14586 Da reflects a covalent adduct with a mass of 42 Da. This adduct is an acetyl group and likely results from the decarboxylation of the 3-oxopropanoate moiety, which is a β-keto acid. Two minor peaks corresponding to masses of 14144 ± 2 and 14101 ± 2 Da were also observed and represent the covalent modification of an unidentified species in the protein preparation by a 3-oxopropanoate and acetyl moiety, respectively.

Crystal Structure of FG41 MSAD

The crystal structure of the native FG41 MSAD was determined to 2.0 Å resolution using molecular replacement with the native *Pseudomonas* MSAD as the search model (PBD code 2AAL), and refined to R and R_{free} values of 19.6% and 24.1%, respectively. The asymmetric unit contains 12 monomers where each monomer shows two repeats of a β–α–β unit, which is the signature tautomerase superfamily structural motif.[9,10] The electron densities for 131 amino acids making up a single monomer are well defined. The crystal structure also confirms that FG41 MSAD is a homotrimer. A three-dimensional structural comparison of the native structure of Pp MSAD with that of FG41 MSAD shows strong similarity with RMSD values of 1.10 Å (Cα), 1.10 Å (backbone), and 1.46 Å (global structure).

The Active Site of FG41 MSAD

The active site of FG41 MSAD is identified by the presence of Pro-1 at the bottom of a cavity. With the exception of Ala-55′ and Phe-103′, the active site consists of residues from a single monomer. Looking down at Pro-1 from the top of the cavity, there are roughly four hydrophobic sides to the cavity (Figure 2).[33] The left side is formed by Trp-114, Phe-116, Leu-128, Phe-103′, Ala-55′, Phe-39 and Tyr-123. The right side is formed by Val-84, Phe-85, and Ile-88. The front side, which is more of a lip, is formed by Ala-26, Leu-27, Val-28, Val-30, Leu-31, Ala-32, Ile-33, and Pro-34. Asp-37 fills a gap between front lip and left wall. The lower backside is formed by Leu-2, Ile-3, Ile-68, Val-70, Phe-71, Ile-104, Ala-105, and Ile-106. The hydroxyl group of Tyr-123 (from the left side) points into the active site cavity. Glu-108 interacts with Lys-80 at right end of cavity and is part of a string of hydrophilic residues that lies over the top of the cavity.

Sequence analysis shows a high degree of identity in the C-terminal part of the 8 sequences beginning with Trp-114 (Figure 1). In a 17-amino acid stretch, 8 residues are identical (including Trp-114, Ser-115, Phe-116, and Gly-117) and 3 residues are highly similar. These residues make up a region of hydrophobicity that includes Trp-114, Phe-116, Tyr-123 (Phe-123 in Pp MSAD) and Leu-128. The hydrophobic "pocket" could destabilize the negative charge of **4** and facilitate decarboxylation.

Examination of the active site shows that the putative phosphate ion sits above Pro-1, and that the phosphate oxygens interact with the prolyl nitrogen of Pro-1 (2.8 Å) and the side chains of Asp-37 (3.2 Å), Tyr-123 (3.2 Å), Gln-73 (2.7 Å), and Thr-72 (2.7 Å) (Figure 3). There is also an interaction with the backbone amide of Gln-73 (3.2 Å). Arg-76 is further away (~7.1 Å between the η_2-nitrogen of Arg-76 and the δ-carbon of Pro-1). The closest interactions are those between the side chain amide of Gln-73 (2.7 Å) and the side chain hydroxyl group of Thr-72 (2.7 Å). Near the active site, there is an interaction between the side chains of Glu-108 and Lys-80 (2.9 Å). These residues are conserved in 7 of the 8 homologues shown in Figure 1.

Crystal Structure of the Inactivated FG41 MSAD

The mass spectral analysis of FG41 MSAD isolated from an incubation mixture containing enzyme and **7** is consistent with the covalent modification of the prolyl nitrogen by the 3-oxopropoanoate moiety (Scheme 3).[12,13] Pro-1 is the sole site of modification on the protein. The crystal structure of the inactivated FG41 MSAD was determined to 2.2 Å resolution using molecular replacement with the native FG41 MSAD as the search model, and refined to R and R_{free} values of 16.4% and 20.2%, respectively. Inspection of the F_o -F_c electron density map showed unambiguous electron density for the covalent attachment of the 37 oxopropoanoate moiety to the prolyl nitrogen.

Examination of the structure shows hydrogen bonds between the adduct and the side chains of Asp-37, Tyr-123, Gln-73, and Thr-72 (Figure 4). Asp-37 appears to be an acid because each of its carboxylate oxygens is within hydrogen bonding distance of the carbonyl oxygen of the adduct (2.8 Å), suggesting that Asp-37 has abstracted a proton from water. The fact that both carboxylate oxygens of Asp-37 are equidistant from the carbonyl group is probably the result of statistical averaging. One carboxylate oxygen of the adduct forms hydrogen bonds with the side chain amide of Gln-73 (3.0 Å) and the side chain hydroxyl group of Tyr-123 (3.1 Å), while the other carboxylate oxygen forms hydrogen bonds with the side chain hydroxyl group of Thr-72 (2.8 Å) and the backbone amide of Gln-73 (2.8 Å). The carboxylate group of the adduct is poised just "in front" and parallel to the side chain indole ring of Trp-114. Arg-76 does not appear to interact with the adduct.

Comparison of the FG41 MSAD and Pp MSAD Active Sites

A comparison of the active sites shows the conservation of Pro-1 and Asp-37. Arg-73 in Pp MSAD is replaced by Gln-73 in FG41 MSAD, where the side chain amide is almost in an identical position as the ε nitrogen of the side chain of Arg-73. Arg-75 in Pp MSAD is replaced with Arg-76 because of the insertion of Gly-75 in FG41 MSAD. Phe-123 in Pp MSAD is replaced by Tyr-123 in FG41 MSAD. Furthermore, Tyr-39 and Thr-55′, which do not have assigned roles in the Pp MSAD mechanism, but are located near the active site, are replaced with non-polar residues, Phe-39 and Ala-55′ in FG41 MSAD. Thr-72 in FG41 MSAD replaces Ser-72 in Pp MSAD. However, the carboxylate group of the adduct is only 2.8 Å from the side chain hydroxyl group of Thr-72 compared to 3.3 Å in the case of Ser-72 in Pp MSAD (Figure 4).

DISCUSSION

The tautomerase superfamily consists of five known families bearing the names of their founding members, 4-oxalocrotonate tautomerase (4-OT), which includes the CaaD subfamily, 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI), macrophage migration inhibitory factor (MIF), *cis*-CaaD, and MSAD. 4-OT and CHMI are tautomerases in bacterial pathways for the degradation of aromatic hydrocarbons (4-OT) or aromatic amino acids (CHMI).[8-10,34] MIF is a mammalian cytokine with phenylpyruvate tautomerase (PPT) activity.

The tautomerase superfamily is not as well characterized as other superfamilies such as the enolase and haloacid dehalogenase (HAD) superfamilies.[35,36] In fact, most tautomerase superfamily members have not been functionally annotated except for a handful in the 4-OT and *cis*-CaaD families, and the founding members of each of the other three families. In the 4-OT family, CaaD and 4-OT represent the diversity of the family: CaaD catalyzes a hydrolytic dehalogenation (**3** to **4** in Scheme 1)[4] and 4-OT converts 2-hydroxymuconate (**13**, Scheme 4) to 2-oxo-3-hexenedioate (**14**).[37] Two other enzymes in the family (a heterohexamer 4-OT designated hh4-OT and TomN) catalyze the same reaction (**13** to **14**) with comparable efficiencies, but have different biological niches. The hh4-OT is found in a putative catabolic pathway for aromatic hydrocarbons in the thermophile *Chloroflexus aurantiacus* J-10-fl[38] and TomN is found in a biosynthetic pathway for the C ring of the anti-tumor antibiotic agent, tomaymycin.[39] TomN likely has a different biological substrate.[39]

The only characterized members of the *cis*-CaaD family are *cis*-CaaD (**2** to **4**, Scheme 1),[6] the homologue designated Cg10062 from *Corynebacterium glutamicum*,[40] and a homologue MsCCH2 from *Mycobacterium smegmatis* MC2.[41] Cg10062 and MsCCH2 do not have known functions or a genomic context that provides clues about their functions. Although Cg10062 has the six residues (Pro-1, His-28, Arg-70, Arg-73, Tyr-103, Glu-114) that are critical for *cis*-CaaD activity as well as a very similar active site, it is not a very efficient *cis*-CaaD.[40] MsCCH2 is a more distant family member of *cis*-CaaD than Cg10062. Only four of the six residues that are critical for activity of *cis*-CaaD are conserved in MsCCH2 (Pro-1, His-28, Arg-70, Glu-114), but the enzyme displays low-level *cis*-CaaD activity. Whereas *cis*-CaaD is highly specific for the *cis*-isomer, Cg10062 and MsCCH2 process both isomers of 3-chloroacrylate (**2** and **3**), with Cg10062 having a preference for the *cis*-isomer and MsCCH2 having a slight preference for the *trans*-isomer.[40,41] These observations reinforce the well-known difficulties in the functional annotation of closely related homologues.[42]

FG41 MSAD attracted our attention because it has decarboxylase activity, but it is missing one critical residue (Arg-73) and a second one is shifted in position (Arg-76). It also has a significantly reduced hydratase activity (using **6**). Hence, it represents an interesting contrast to the *cis*-CaaD/Cg10062 pair in that a different set of residues likely contributes to the decarboxylase and hydratase activities. Characterization can assist in the identification of new homologues in the MSAD family, and potentially, their functional annotation, which is a fundamental problem in biochemistry.[42]

Previous kinetic, mutagenesis, inhibition, and crystallographic studies suggested mechanisms for the decarboxylase and hydratase activities of Pp MSAD (using **6**) as well as its inactivation by the 3-halopropiolates (i.e., **7** and **8**).[7,11-13] All three mechanisms involve Pro-1, Asp-37, Arg-73, and Arg-75. For the decarboxylase activity, the cationic Pro-1 (pK_a ~ 9.2)[11] is proposed to polarize the 3-keto group of **4** (Scheme 5).[13] The arginine pair might bind the carboxylate group and stabilize the developing enolate species. The two arginines could also position the carboxylate group such that the scissile carbon-carbon bond (C1-C2

of **4**) is parallel to the p-orbitals of the carbonyl group.[43] This orientation and the position of the carboxylate group relative to the hydrophobic wall (consisting of Trp-114, Phe-116, Phe-123, and Leu-128) facilitate decarboxylation. It is further proposed that Asp-37 maintains the pK_a of Pro-1 by participating in a hydrogen bond network along with the amino-terminal proline.[11] Changing Asp-37 to an asparagine reduces the decarboxylase activity of Pp MSAD (0.5% of wild-type), which can be attributed (at least in part) to a disruption of this hydrogen bond network and perhaps a decrease in the pK_a of Pro-1. Finally, the position of Pro-1 and pK_a suggest that it adds a proton to the enolate to produce acetaldehyde (**5**). Precise roles for these residues in decarboxylation (and whether they contribute to binding and/or catalysis) cannot be determined because it is not possible to measure K_m or k_{cat} values.[7,13]

The Pp MSAD-catalyzed hydration of the 3-halopropiolates (i.e., **7** and **8**) leads to enzyme inactivation, whereas hydration of **6** produces acetopyruvate (**9**).[11,12] Based on the crystal structure of the inactivated enzyme and positions of the residues interacting with the covalent adduct, it is proposed that Asp-37 activates water for attack at C-3 of **7** or **8** (Scheme 6).[13] The arginine pair serves dual roles: it binds the carboxylate group and polarizes the α,β-unsaturated acid. Polarization creates a partial positive charge at C-3, which facilitates the addition of water at C-3. Collapse and protonation at C-2 (presumably by Pro-1) leads to an acyl halide or ketene (**10** and **11**, respectively, Scheme 3), which forms a covalent bond with Pro-1 (Scheme 6). Two observations are consistent with the proposed series of events. First, Asp-37 forms a hydrogen bond with the 3-oxo group of the adduct. The hydrogen bonding capability of Asp-37 suggests that it is an acid because it has abstracted a proton from water. Second, Pro-1 is the exclusive site of modification so that the cationic Pro-1 must be deprotonated in the course of the hydration reaction. Deprotonation renders the prolyl nitrogen nucleophilic so that it can attack the acylating agent (**10** or **11** in Scheme 3) and become covalently modified.

A similar scenario can be envisioned for the enzyme-catalyzed hydration of **6** (Scheme 7). Accordingly, Asp-37 activates water and one or both arginine residues could polarize the α,β-unsaturated ketone moiety of **6**. The non-participating arginine could bind the carboxylate group. Pro-1 provides a proton to the C-3 position to produce **9**.

The crystal structures of the (putative) phosphate-bound and inactivated FG41 MSAD clearly show the Arg-76 is almost certainly not directly involved in the decarboxylation of **4** or in the hydration of **7** (leading to inactivation) unless a major conformational change occurs upon substrate/inhibitor binding. In the phosphate-bound FG41 MSAD (Figure 2), interactions are observed between the phosphate oxygen and the side chains of Pro-1, Asp-37, Tyr-123, Gln-73, and Thr-72 (where the closest interactions are those of Gln-73 and Thr-72). Arg-76 is further away (~7.1 Å). In the inactivated structure, interactions are observed between the adduct and the side chains of Asp-37, Tyr-123, Gln-73, and Thr-72. Again, Arg-76 does not appear to interact with the adduct. The fact that the R76A mutant of FG41 MSAD shows no decarboxylase activity might be due to a structural defect or to the removal of a positive charge from the active site. The reason for the complete loss of activity for this mutant is under investigation.

The FG41 MSAD crystal structures suggest that the major changes in the three reaction mechanisms will be the interactions between the active site residues and the carboxylate (**4** and **7**/**8**) or the α-keto carboxylate (**6**) end of the molecule (Scheme 8). Accordingly, Tyr-123, Gln-73, and Thr-72 could interact with the carboxylate moiety of **4** or **7**/**8**. For the decarboxylation of **4**, these interactions could bind and facilitate decarboxylation by orientating the carboxylate group in a conformation favorable for decarboxylation (where the C1-C2 of **4** is parallel to the p-orbitals of the carbonyl group).[43] The side chain of

Tyr-123 could also polarize the 3-keto group of **4** (via hydrogen bonding) along with the presumably cationic Pro-1. The proposed roles are consistent with the substantial loss of decarboxylase activity in the Q73A and Y123F mutants of FG41 MSAD (~8% and 6% of wild type activity, respectively). (Again, because kinetic parameters cannot be measured, precise roles for these residues in catalysis and/or binding cannot be assigned.) Like Pp MSAD, Asp-37 of FG41 MSAD could maintain the pK_a of Pro-1 by its participation in a hydrogen bond network (along with the proline). Disruption of the network might alter the pK_a of Pro-1. Pro-1 could again provide a proton to the enolate to produce acetaldehyde. The position of the carboxylate group of the adduct in the inactivated FG41 MSAD structure (Figure 4) suggests that the carboxylate group of the substrate could face a hydrophobic wall consisting of Trp-114, Phe-116, Leu-128, and Tyr-123, which would facilitate decarboxylation.

These same residues (Tyr-123, Gln-73 and Thr-72) could be involved in binding of the carboxylate group of **7**/**8** as well as the polarization of the α,β-unsaturated acid. Polarization would facilitate the addition of water to C-3 and initiate the cascade of events resulting in the inactivation of the enzyme (protonation at C-2, formation of an acyl halide or ketene, and covalent bond formation, as shown Scheme 3 and 6). Asp-37 is, again, positioned to activate the water for addition to C-3 and the cationic Pro-1 can still provide the proton at C-2.

It's less clear how these three residues might be involved in the FG41 MSAD-catalyzed conversion of **6** to **9** because the low wild-type hydratase activity. However, Tyr-123 might interact with the α-keto moiety while Gln-73 and Thr-72 might interact with the carboxylate group. Changing Tyr-123 or Gln-73 to an alanine has little effect on the hydratase activity (as assessed by the k_{cat}/K_m values) and it is comparable to that of the wild type. The Y123A mutant (of FG41 MSAD) increases the k_{cat} and K_m values. The latter could reflect a role in binding. (The Q73A mutant did not give consistent kinetic parameters and mutations of Thr-72 have not been examined.) The roles of Asp-37 and Pro-1 are the same as described above.

One interesting observation is that the presence of Arg-73 enhances both the decarboxylase and hydratase activities of FG41 MSAD. The Q73R mutant of FG41 MSAD increases the decarboxylase activity (8-fold greater than that observed for the Q73A mutant) and hydratase activity (9-fold greater than that observed for wild type). Superimposing the two active sites of the inactivated enzymes (Figure 4) shows that the side chains of Gln-73 (FG41 MSAD) and Arg-73 (Pp MSAD) overlay one another. Clearly, a positively charged arginine in this position has a major impact on decarboxylation. Likewise, the positive charge (and perhaps hydrogen bonding) (Arg-73) in this position contributes much more to catalysis of the hydration reaction (as assessed by the increase in k_{cat}) than hydrogen bonding alone (Gln-73). These observations also suggest the possibility that FG41 MSAD lacks a robust hydratase activity because the two positively charged arginine residues are not present in the active site (Arg-73 is replaced by a glutamine) or nearby the substrate (Arg-76 is more distant).

The differences between the Pp MSAD and the FG41 MSAD raise two questions: does FG41 MSAD have a different substrate and is there an evolutionary pressure for the observed active site substitutions? In *Coryneform* bacterium strain FG41, the gene coding for MSAD lies upstream of a gene coding for *cis*-CaaD. In between these two genes is a gene coding for a putative aldehyde dehydrogenase similar to a methylmalonate semialdehyde dehydrogenase (unpublished results, G.J. Poelarends, H. Serrano, and C.P. Whitman, 2009). This genomic context is similar to that observed for the gene encoding MSAD in *P. pavonaceae* 170.[7] Hence, FG41 MSAD is likely in a pathway for the

degradation of **2**. Whether the MSAD-catalyzed step is the bottleneck step in either pathway is not known and would require substantial experimentation. The active site substitutions observed in the FG41 MSAD do not affect decarboxylase activity so it is not likely related to the evolutionary optimization of the enzyme.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research reported was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01 GM65324 and by the Robert A. Welch Foundation grants F-1334 (CPW) and F-1219 (MLH). G.J.P. was supported by a VIDI grant from the Division of Chemical Sciences of the Netherlands Organization of Scientific Research (NWO-CW).

ABBREVIATIONS

Ap	ampicillin
***cis*-CaaD and CaaD**	*cis*- and *trans*-3-chloroacrylic acid dehalogenase, respectively
CHMI	5-(carboxymethyl)-2-hydroxymuconate isomerase
ESI-MS	electrospray ionization mass spectrometry
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonate
MALDI-PSD	matrix assisted laser desorption-ionization post-source decay
MALDI-TOF	matrix assisted laser desorption-ionization time-of-flight
MIF	macrophage migration inhibitory factor
Pp MSAD	malonate semialdehyde decarboxylase from *Pseudomonas pavonaceae* 170
FG41 MSAD	malonate semialdehyde decarboxylase from *Coryneform* bacterium strain FG41
NCBI	National Center for Biotechnology Information
NMR	nuclear magnetic resonance
4-OT	4-oxalocrotonate tautomerase
PEG	polyethylene glycol
RMSD	root-mean-square deviation
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis

REFERENCES

1. Hartmans S, Jansen MW, Van der Werf MJ, De Bont JAM. Bacterial metabolism of 3-chloroacrylic acid. J. Gen. Microbiol. 1991; 137:2025–2032. [PubMed: 1720168]
2. van Hylckama Vlieg JET, Janssen DB. Bacterial degradation of 37 chloroacrylic acid and the characterization of *cis*- and *trans*-specific dehalogenases. Biodegradation. 1992; 2:139–150. [PubMed: 1368960]
3. Poelarends GJ, Wilkens M, Larkin MJ, van Elsas JD, Janssen DB. Degradation of 1,3-dichloropropene by *Pseudomonas pavonaceae* 170. Appl. Environ. Microbiol. 1998; 64:2931–2936. [PubMed: 9687453]

4. Poelarends GJ, Saunier R, Janssen DB. *trans*-3-Chloroacrylic acid dehalogenase from *Pseudomonas pavonaceae* 170 shares structural and mechanistic similarities with 4-oxalocrotonate tautomerase. J. Bacteriol. 2001; 183:426974277.
5. Wang SC, Person MD, Johnson WH Jr. Whitman CP. Reactions of *trans*-3-chloroacrylic acid dehalogenase with acetylene substrates: consequences of and evidence for a hydration reaction. Biochemistry. 2003; 42:8762–8773. [PubMed: 12873137]
6. Poelarends GJ, Serrano H, Person MD, Johnson WH Jr. Murzin AG, Whitman CP. Cloning, expression, and characterization of a *cis*-3-chloroacrylic acid dehalogenase: insights into the mechanistic, structural, and evolutionary relationship between isomer-specific 3-chloroacrylic acid dehalogenases. Biochemistry. 2004; 43:759–772. [PubMed: 14730981]
7. Poelarends GJ, Johnson WH Jr. Murzin AG, Whitman CP. Mechanistic characterization of a bacterial malonate semialdehyde decarboxylase: identification of a new activity in the tautomerase superfamily. J. Biol. Chem. 2003; 278:48674–48683. [PubMed: 14506256]
8. Murzin AG. Structural classification of proteins: new superfamilies. Curr. Opin. Struct. Biol. 1996; 6:386–394. [PubMed: 8804825]
9. Whitman CP. The 4-oxalocrotonate tautomerase family of enzymes: how nature makes new enzymes using a β-α-β structural motif. Arch. Biochem. Biophys. 2002; 402:1–13. [PubMed: 12051677]
10. Poelarends GJ, Whitman CP. Evolution of enzymatic activity in the tautomerase superfamily: mechanistic and structural studies of the 1,3-dichloropropene catabolic enzymes. Bioorg. Chem. 2004; 32:376–392. [PubMed: 15381403]
11. Poelarends GJ, Serrano H, Johnson WH Jr. Hoffman DW, Whitman CP. The hydratase activity of malonate semialdehyde decarboxylase: mechanistic and evolutionary implications. J. Am. Chem. Soc. 2004; 126:15658–15659. [PubMed: 15571384]
12. Poelarends GJ, Serrano H, Johnson WH Jr. Whitman CP. Inactivation of malonate semialdehyde decarboxylase by 3-halopropiolates: evidence for the hydratase activity. Biochemistry. 2005; 44:9375–9381. [PubMed: 15982004]
13. Almrud JJ, Poelarends GJ, Johnson WH Jr. Serrano H, Hackert ML, Whitman CP. Crystal structures of the wild-type, P1A mutant, and inactivated malonate semialdehyde decarboxylase: a structural basis for the decarboxylase and hydratase activities. Biochemistry. 2005; 44:14818–14827. [PubMed: 16274229]
14. Johnson WH Jr. Czerwinski RM, Fitzgerald MC, Whitman CP. Inactivation of 4-oxalocrotonate tautomerase by 2-oxo-3-pentynoate. Biochemistry. 1997; 36:15724–15732. [PubMed: 9398301]
15. Strauss F, Kollek L, Heyn W. Uber den ersatz positiven wasserstoffs durch halogen. Chem. Ber. 1930; 63:1868–1899.
16. Andersson K. Additions to propiolic and halogen substituted propiolic acids. Chem. Scripta. 1972; 2:117–120.
17. Sambrook, J.; Fritsch, EF.; Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory; Cold Spring Harbor, NY: 1989.
18. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970; 227:680–685. [PubMed: 5432063]
19. Waddell WJ. A simple ultraviolet spectrophotometric method for the determination of protein. J. Lab. Clin. Med. 1956; 48:311–314. [PubMed: 13346201]
20. Atschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J. Mol. Biol. 1990; 215:403–410. [PubMed: 2231712]
21. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]
22. Pospiech A, Neumann B. A versatile quick-prep of genomic DNA from Gram-positive bacteria. Trends Genet. 1995; 11:217–218. [PubMed: 7638902]
23. Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
24. Matthews BW. Solvent content of protein crystals. J. Mol. Biol. 1968; 33:491–497. [PubMed: 5700707]

25. Vagin A, Teplyakov A. *MOLREP*: an automated program for molecular replacement. J Appl Cryst. 1997; 30:1022–1025.
26. Claude JB, Suhre K, Notredame C, Claverie JM, Abergel C. CaspR: a web server for automated molecular replacement using homology modeling. Nucleic Acids Res. 2004; 32(Web Server issue):W606–609. [PubMed: 15215460]
27. Navaza J. AMoRe: an automated package for molecular replacement. Acta Crystallogr. A. 1994; 50:157–163.
28. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A. 1991; 47:110–119. [PubMed: 2025413]
29. Bruünger AT, Adams PD, Clore GM, Delano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. Sect. D: Biol.Crystallogr. 1998; 54:905–921. [PubMed: 9757107]
30. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect D: Biol. Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
31. Adams PD, Afonine PV, Bunkoczj 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 Crystallogr. Sect. D: Biol. Crystallogr. 2010; 66:213–221. [PubMed: 20124702]
32. Hirel P-H, Schmitter J-M, Dessen P, Fayat G, Blanquet S. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. Proc. Natl. Acad. Sci. USA. 1989; 86:8247–8251. [PubMed: 2682640]
33. DeLano, WL. The PyMol Molecular Graphics System. DeLano Scientific; San Carlos, CA: 2002.
34. Poelarends GJ, Veetil VP, Whitman CP. The chemical versatility of the β–α-β fold: catalytic promiscuity and divergent evolution in the tautomerase superfamily. Cell Mol Life Sci. 2008; 65:3606–3618. [PubMed: 18695941]
35. Gerlt JA, Babbitt PC, Jacobson MP, Almo SC. Divergent Evolution in enolase superfamily for assigning functions. J. Biol. Chem. 2012; 287:29–34. [PubMed: 22069326]
36. Nguyen HH, Wang L, Huang H, Peisach E, Dunaway-Mariano D, Allen KN. Structural determinants of substrate recognition in the HAD superfamily member D-*glycero*-D-*manno*-heptose-1,7-bisphosphate phosphatase (GmhB). Biochemistry. 2010; 49:1082–1092. [PubMed: 20050614]
37. Wang SC, Johnson WH Jr. Czerwinski RM, Stamps SL, Whitman CP. Kinetic and stereochemical analysis of YwhB, a 4-oxalocrotonate tautomerase homologue in *Bacillus subtilis*: mechanistic implications for the YwhB- and 4-oxalocrotonate tautomerase-catalyzed reactions. Biochemistry. 2007; 46:11919–11929. [PubMed: 17902707]
38. Burks EA, Fleming CD, Mesecar AD, Whitman CP, Pegan SD. Kinetic and structural characterization of a heterohexamer 4-oxalocrotonate tautomerase from *Chloroflexus aurantiacus* J-10-fl: implications for functional and structural diversity in the tautomerase superfamily. Biochemistry. 2010; 49:5016–5027. [PubMed: 20465238]
39. Burks EA, Yan W, Johnson WH Jr. Li W, Schroeder GK, Min C, Gerratana B, Zhang Y, Whitman CP. Kinetic, crystallographic, and mechanistic characterization of TomN: elucidation of a function for a 4-oxalocrotonate tautomerase homologue in the tomaymycin biosynthetic pathway. Biochemistry. 2011; 35:7600–7611. [PubMed: 21809870]
40. Poelarends GJ, Serrano H, Person MD, Johnson WH Jr. Whitman CP. Characterization of Cg10062 from *Corynebacterium glutamicum*: implications for the evolution of *cis*-3-chloroacrylic acid dehalogenase activity in the tautomerase superfamily. Biochemistry. 2008; 47:8139–8147. [PubMed: 18598055]
41. Baas BJ, Zandvoort E, Wasiel AA, Quax WJ, Poelarends GJ. Characterization of a newly identified mycobacterial tautomerase with promiscuous dehalogenase and hydratase activities

reveals a functional link to a recently diverged *cis*-3-chloroacrylic acid dehalogenase. Biochemistry. 2011; 50:2889–2899. [PubMed: 21370851]

42. Zhu WW, Wang C, Jipp J, Ferguson L, Lucas SN, Hicks MA, Glasner ME. Residues required for activity in *Escherichia coli o*-succinylbenzoate synthase (OSBS) are not conserved in all OSBS enzymes. Biochemistry. 2012; 51:6171–6181. [PubMed: 22775324]
43. Dunathan HC. Conformation and reaction specificity in pyridoxal phosphate enzymes. Proc. Natl. Acad. Sci. USA. 1966; 55:712–716. [PubMed: 5219675]

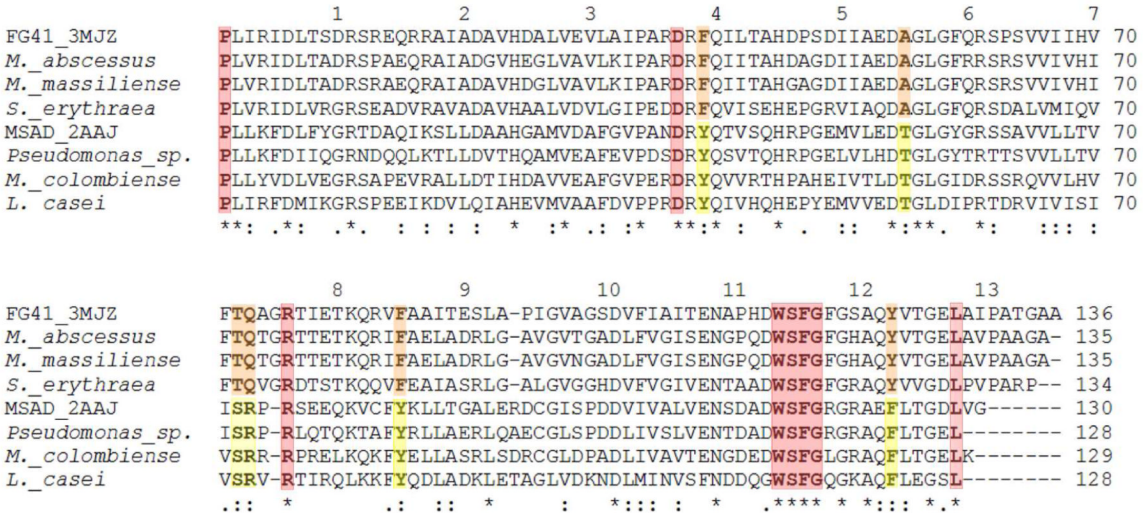


Figure 1.

Sequence alignment of FG41 MSAD and three closely related homologues along with Pp MSAD and three closely related homologues. The first set (FG41 MSAD and homologues from *M. abscessus* ATCC 19977, *M. massiliense* CCUG 48898, and *Saccharopolyspora erythraea* NRRL 2338) shows Pro-1, Asp-37, and Arg-76, and the replacement of Arg-73 with Gln-73. The second set (the *P. pavonaceae* 170 MSAD, and homologues from *Pseudomonas* sp. GM67, *Mycobacterium colombiense* CECT 3035, and *Lactobacillus casei* strain BL23) shows the conservation of Pro-1, Asp-37, Arg-73, and Arg-75. Identical residues are indicated by the single asterisk beneath the sequences. Residues that are similar with respect to hydrophobicity/hydrophilicity or charge are marked with c for lower similarity, and: to indicate higher similarity. The shading indicates that the residue is conserved in both sets (red), conserved in the FG41 MSAD set (orange), or conserved in the Pp MSAD set (yellow). Alignments were obtained using BLAST and CLUSTAL W.[20,21]

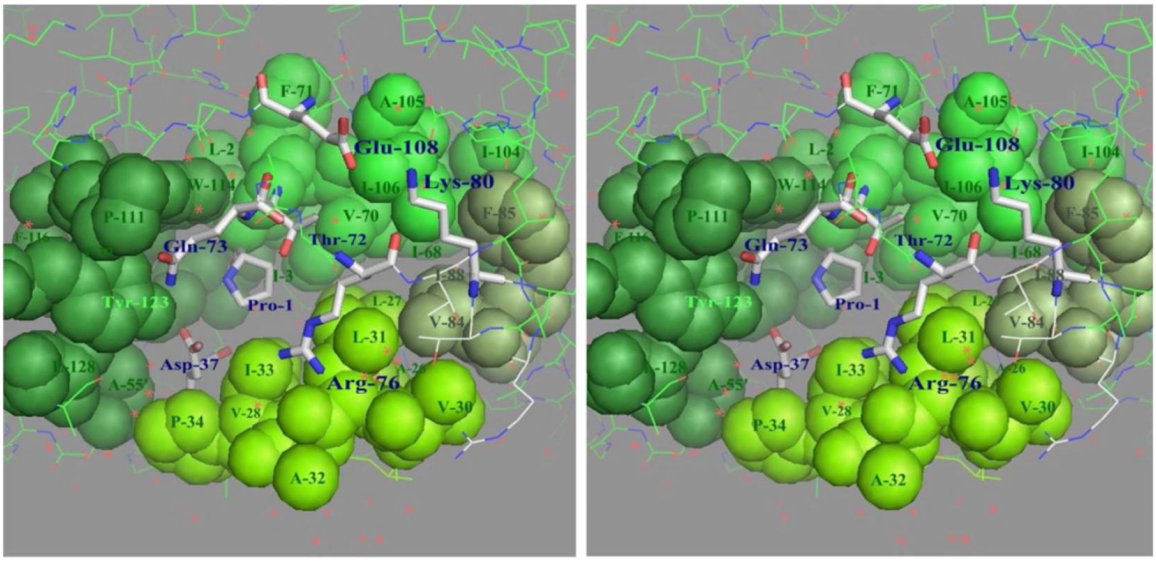


Figure 2.

Stereoview of the active site of FG41 MSAD showing that it consists largely of hydrophobic residues. Looking down on Pro-1, the left side (Forest) is formed by Trp-114, Phe-116, Leu-128, Phe-103′, Ala-55′, Phe-39 and Tyr-123. The right side (Smudge) is formed by Val-84, Phe-85, and Ile-88. The front side (Chartreuse), which is more of a lip, is formed by Ala-26, Leu-27, Val-28, Val-30, Leu-31, Ala-32, Ile-33, and Pro-34. The back wall (Pale Green) is formed by Leu-2, Ile-3, Ile-68, Val-70, Phe-71, Ile-104, Ala-105, and Ile-106. Asp-37 fills a gap between front lip and left wall. The hydroxyl group of Tyr-123 (from the left side) points into the active site cavity. Thr-72, Gln-73 and Arg-76 are part of a string of hydrophilic residues that lie over the top of the cavity with Lys-80 interacting with Glu-108 at the far right end of cavity. (Figure prepared using PyMol).[33]

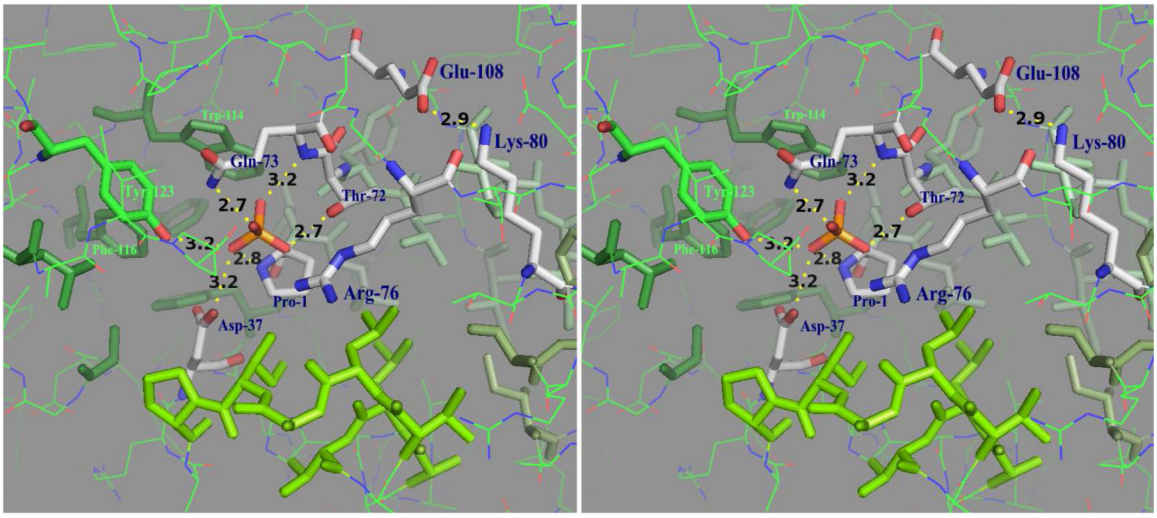


Figure 3.

Stereoview of one of twelve independent active sites of native FG41 MSAD. A large "solvent" peak is found in the active site of native FG41 MSAD and is presumed to be a phosphate ion. There are several potential hydrogen bonding interactions between this putative phosphate ion and active site residues, including the side chains of Tyr-123, Asp-37, Thr-72, the amide and backbone nitrogen of Gln-73, and the prolyl nitrogen of Pro-1. Arg-76 does not interact with the phosphate ion or the residues that interact with it. The hydrophobic residues that form the active site cavity are labeled in Figure 2. Near the active site, there is an interaction between the side chains of Glu-108 and Lys-80 (2.9 Å). (Figure prepared using PyMol).[33]

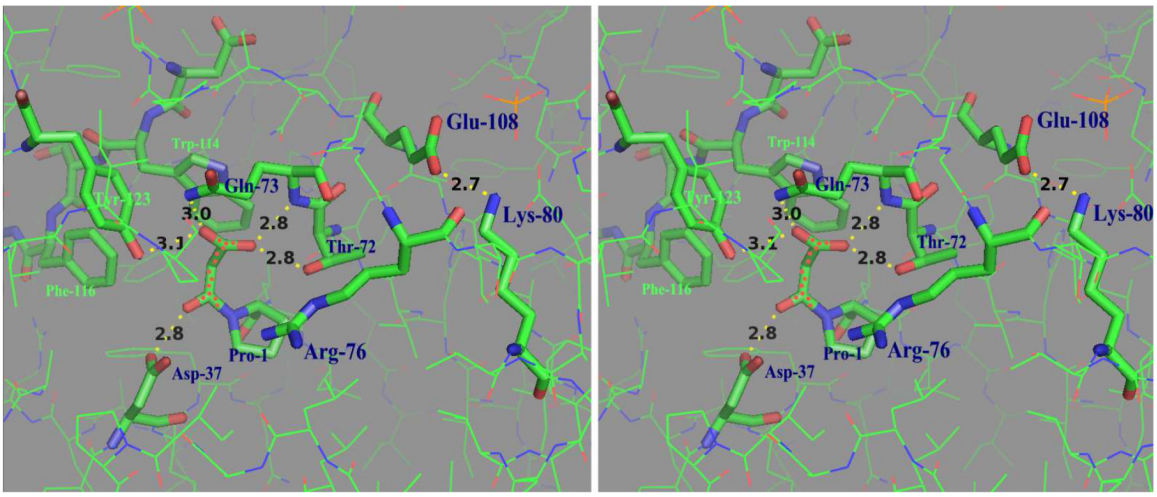


Figure 4.

Stereoview of one of five independent active sites in the inactivated FG41 MSAD. As shown, several interactions are observed between the covalent adduct on Pro-1 (shown in red dots) and Asp-37, Thr-72, Gln-73, Tyr-123, but not Arg-76. One carboxylate oxygen of Asp-37 interacts with the 3-oxo moiety of the adduct. The amide nitrogen of the side chain of Gln-73 and the side chain hydroxyl group of Tyr-123 interact with one carboxylate oxygen of the adduct (3.0 Å and 3.1 Å, respectively). The backbone amide and the side chain of Thr-72 form hydrogen bonds with the second oxygen atom of the carboxylate group of the adduct (both 2.8 Å). However, Arg-76 does not interact with the inhibitor adduct. This suggests that Arg-76 does not play a similar role as Arg-75 in Pp MSAD. (Figure prepared using PyMol).[33]

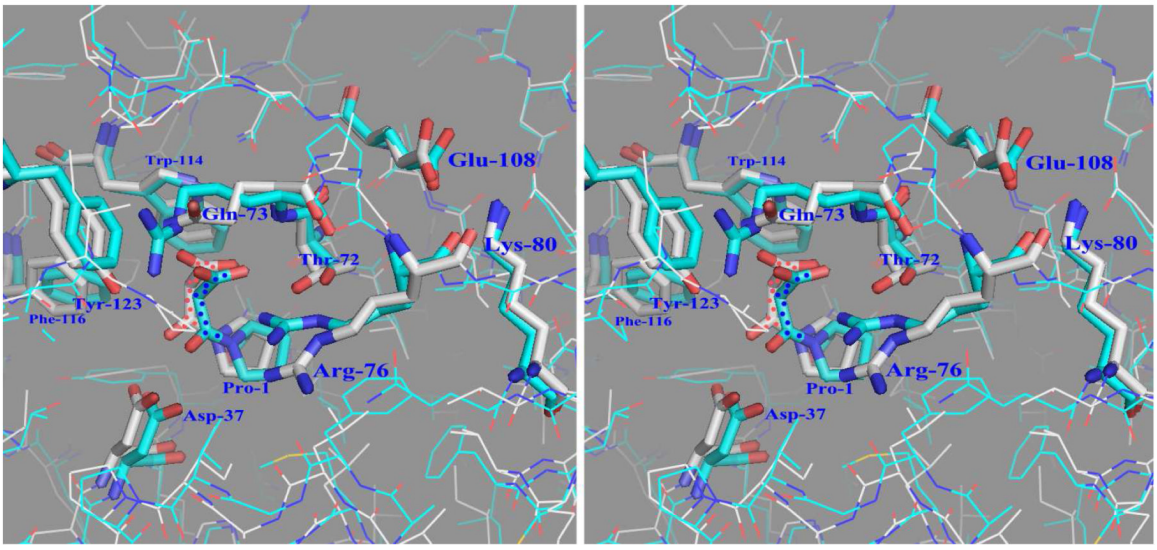
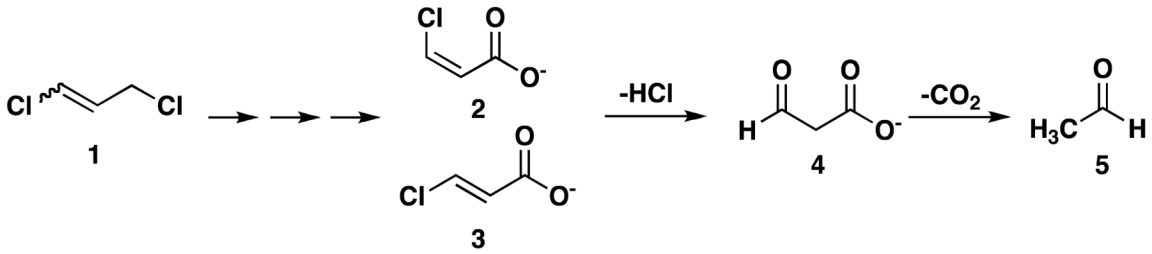
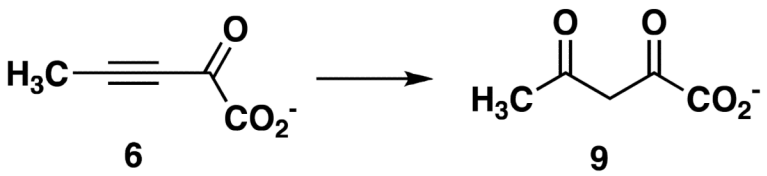


Figure 5.

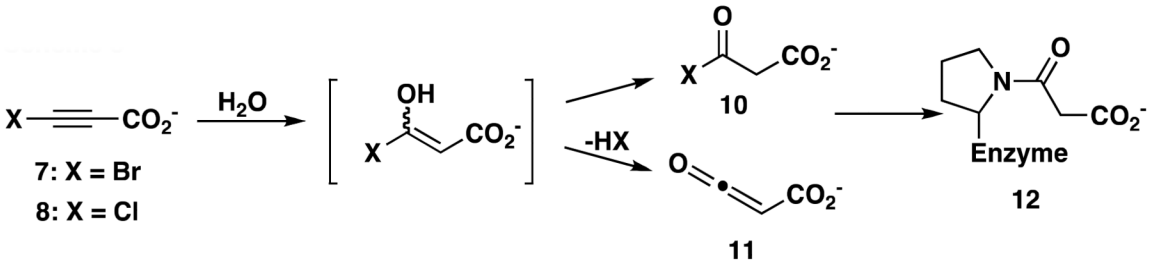
Stereo superimposition of the active sites of inactivated FG41 MSAD and inactivated Pp MSAD. The covalent adduct to Pro-1 of FG41 MSAD is shown in red dots and the covalent adduct to Pro-1 of Pp MSAD is shown in blue dots. There are several differences between the two active sites, with only Pro-1, Asp-37, Phe-116, and Trp-114 being conserved. (Lys-80 and Glu-108 are also conserved, but these residues play structural roles.) In FG41 MSAD, Tyr-123, Thr-72 and Gln-73 replace Phe-123, Ser-72 and Arg-73 of MSAD. (Figure prepared using PyMol).[33]



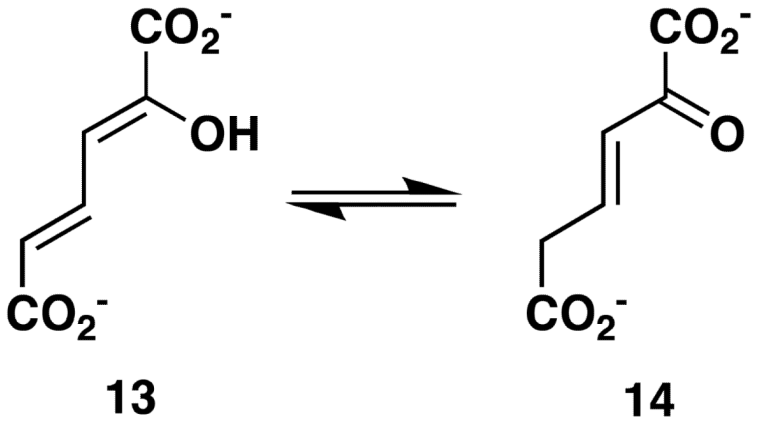
Scheme 1.
The 1,3-dichloropropene catabolic pathway in *Pseudomonas pavonaceae* 170.



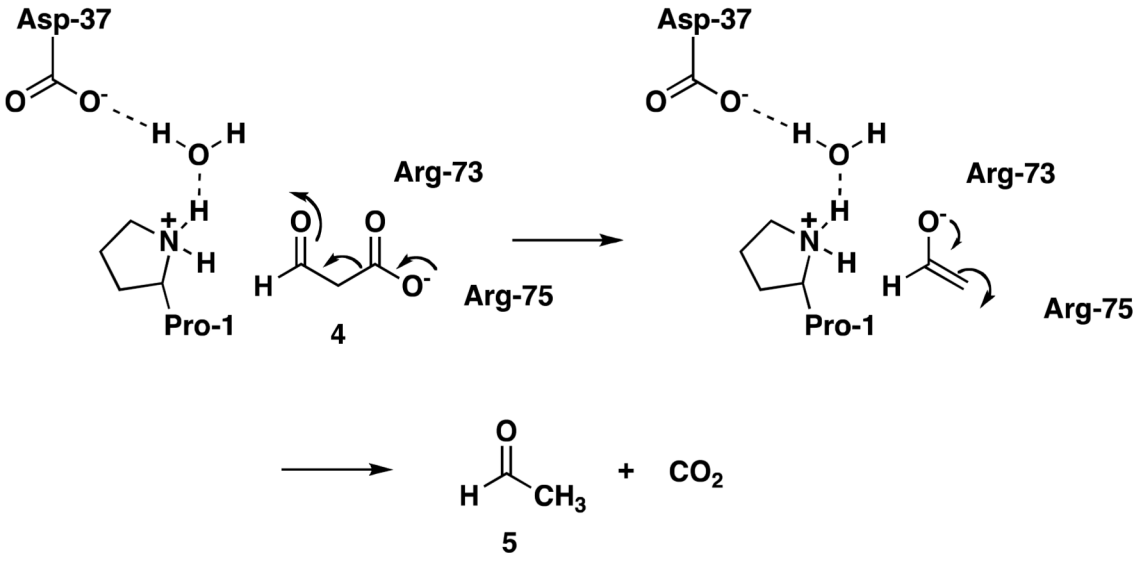
Scheme 2.
The Pp MSAD-catalyzed hydration of 2-oxo-3-pentynoate (**6**) to acetopyruvate (**9**).



Scheme 3.
The reactions of the 3-halopropiolates causing the inactivation of Pp MSAD.

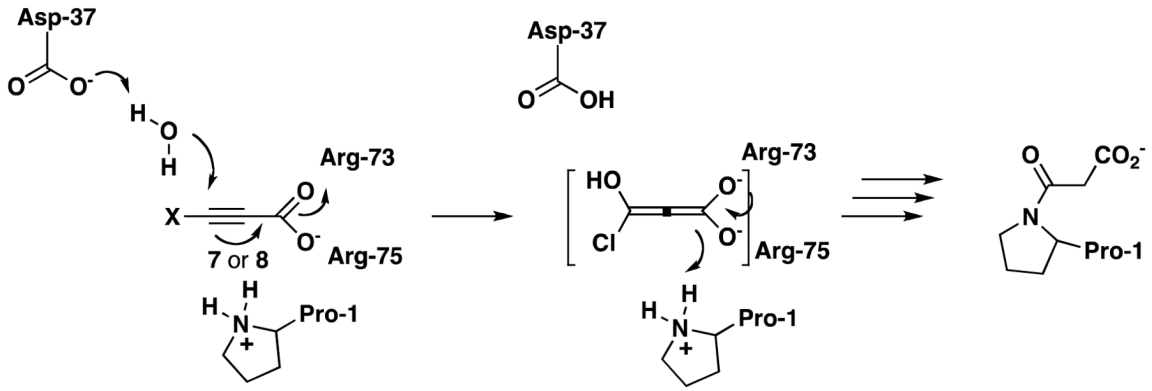


Scheme 4.
The 4-OT-catalyzed reaction.



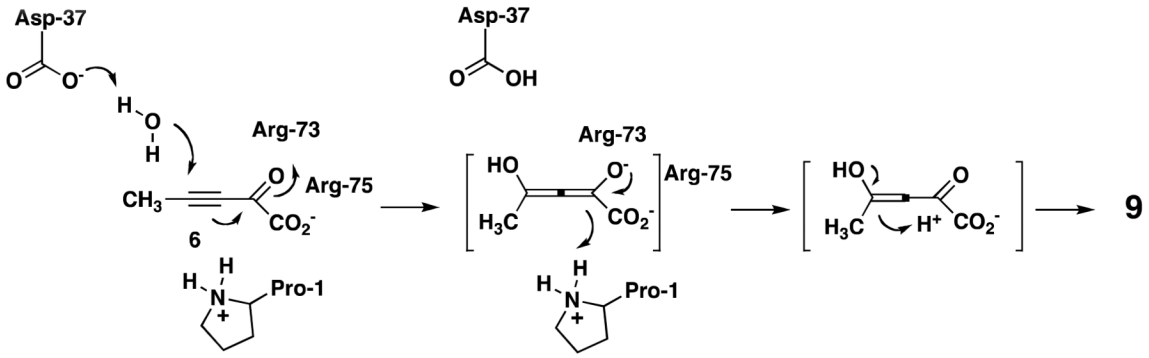
Scheme 5.

The Pp MSAD-catalyzed decarboxylation of **4** with possible roles shown for Pro-1, Asp-37, Arg-73, and Arg-75.

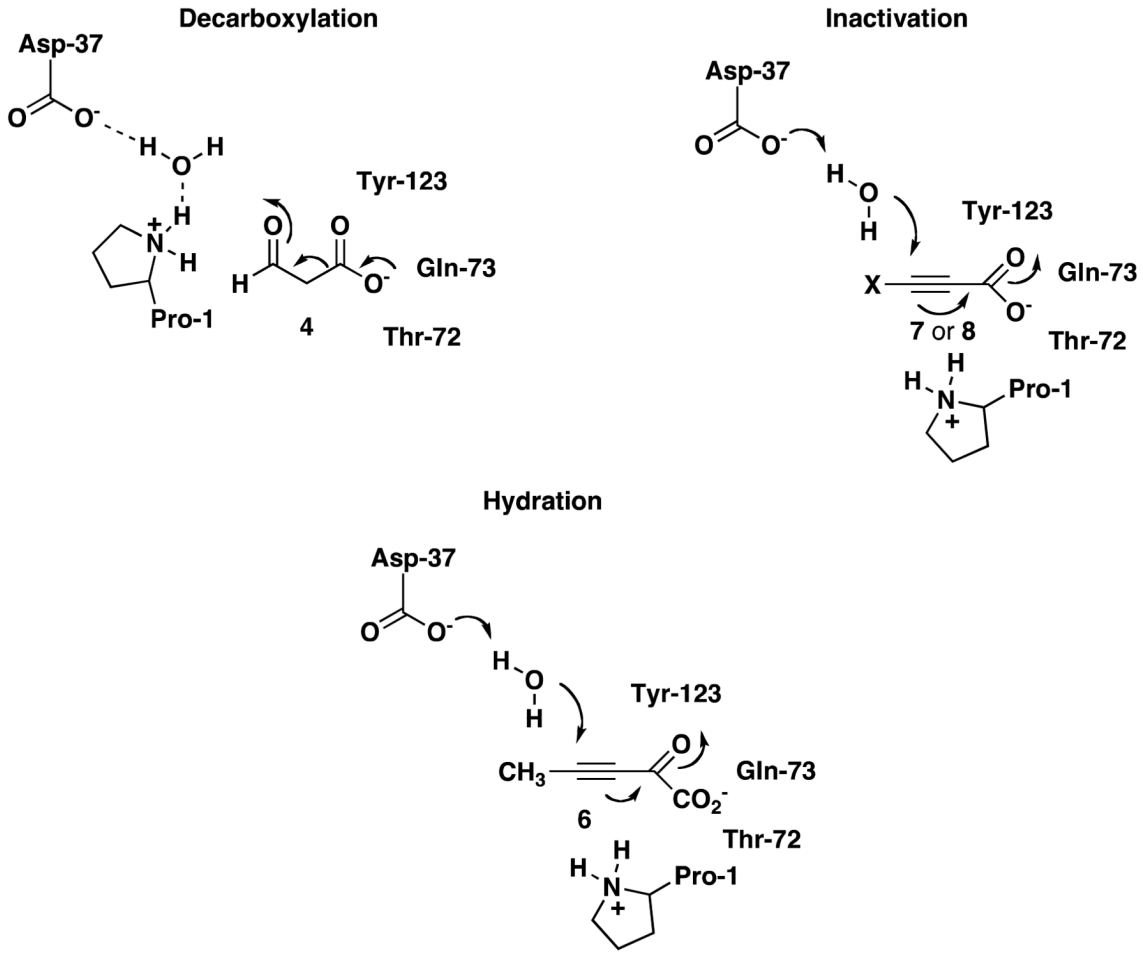


Scheme 6.

The initial steps in the proposed mechanism for the covalent modification of the prolyl-1 nitrogen in Pp MSAD upon incubation with the 3-halopropiolates.



Scheme 7.
The Pp MSAD-catalyzed hydration of **6** with possible roles shown for Pro-1, Asp-37, Arg-73, and Arg-75.



Scheme 8.

The three FG41 MSAD-catalyzed reactions showing proposed roles for Pro-1, Asp-37, Thr-72, Gln-73, and Tyr-123, based on their interactions with the covalent adduct on Pro-1 (i.e., **12** in Scheme 3).

Table 1
Data Collection and Refinement Statistics

	Native FG41 MSAD	Inactivated FG41 MSAD
Data Statistics		
Space Group	$P2_12_12_1$	$P2_13$
No.Chains/Asymmetric Unit	12	5
Unit Cell (Å)	a=88.96, b=94.69, c=190.7	a=144.2
Resolution (Å)	2.0	2.2
Rsym (%) overall (outer shell)	15.4 (67.3)	11.6 (89.4)
Completeness (%) overall (outer shell)	85.8 (86.7)	96.4 (55.9)
I/σ (outer shell)	19.4 (3.8)	32.0 (1.8)
Reflections total (unique)	822,401 (91,098)	91,715 (17,927)
Refinement		
Total non-solvent atoms/waters/ions	11887 / 1148 / 14	5036 / 346 /15
R/Rfree	0.194 / 0.241	0.164 / 0.202
RMSD bonds (Å)/ angles (°)	0.004 / 0.764	0.012/1.295
Ramachandran plot (%)		
(favored / allowed / outliers)	97.28 / 2.72 / 0	98.45 / 1.55 / 0

Table 2

Specific Activities for FG41 MSAD, Pp MSAD, and Mutants[a]

Enzyme	Specific Activity (mU/mg of protein)
FG41 MSAD	25,000-36,000
P1A	N.D.[b]
D37N	N.D.
R76A	N.D.
Q73A	2,700, 8.4%[c]
Q73R	22,000, 61%[d]
Y123F	1,500, 6%[e]
Pp MSAD	36,000-38,000
R73Q-Pp MSAD	700, 2%[f]

c-f Due to the assay variability, the specific activity of the mutant was compared to specific activity of a wild type sample measured at the same time. The specific activities for the wild type enzymes are

[a] The decarboxylase activities were measured by coupling the production of acetaldehyde to the β-NADH-dependent alcohol dehydrogenase-catalyzed reduction of acetaldehyde to ethanol as described in the text and elsewhere.[7,13] The definition of specific activity is provided in the text and elsewhere.[7]

[b] Not detectable above background.

[c] 32,000

[d] 36,000

[e] 25,000

[f] 36,000.

Table 3

Kinetic Parameters for Pp MSAD, FG41 MSAD, and Mutants using 2-Oxo-3-pentynoate[a]

enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
Pp MSAD[b]	3.4 ± 0.8	3000 ±1200	1130
FG41 MSAD	0.17 ± 0.01	7300 ± 800	20
Q73R FG41 MSAD	1.20 ± 0.2	6600 ±1500	180
Y123F FG41 MSAD	0.7 ± 0.1	21000±5000	30
CaaD[c]	0.7 ± 0.02	110 ± 4	6400
cis-CaaD[d]	0.007 ± 0.0005	620 ± 60	11

[a] The steady-state kinetic parameters were determined in 20 mM sodium phosphate buffer (pH 9.0) at 24 °C, as described in the text.

[b] Errors are standard deviations.

[c] The kinetic parameters are taken from reference 5.

[d] The kinetic parameters are taken from reference 6.