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†

Gerrit J. Poelarends,‡ Hector Serrano,‡ Maria D. Person,‖ William H. Johnson, Jr.,‡ Alexey G. Murzin,§ and Christian P. Whitman*,‡

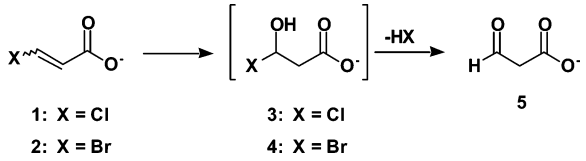
Divisions of Medicinal Chemistry and Pharmacology and Toxicology, College of Pharmacy, The University of Texas, Austin, Texas 78712-1074, and MRC Centre for Protein Engineering, Hills Road, Cambridge CB2 2QH, UK

Received September 4, 2003; Revised Manuscript Received November 17, 2003

ABSTRACT: The gene encoding the *cis*-3-chloroacrylic acid dehalogenase (*cis*-CaaD) from coryneform bacterium strain FG41 has been cloned and overexpressed, and the enzyme has been purified to homogeneity and subjected to kinetic and mechanistic characterization. Kinetic studies show that *cis*-CaaD processes *cis*-3-haloacrylates, but not *trans*-3-haloacrylates, with a turnover number of ∼10 s^{-1}. The product of the reaction is malonate semialdehyde, which was confirmed by its characteristic ^{1}H NMR spectrum. The enzyme shares low but significant sequence similarity with the previously studied *trans*-3-chloroacrylic acid dehalogenase (CaaD) and with other members of the 4-oxalocrotonate tautomerase (4-OT) family. While 4-OT and CaaD function as homo- and heterohexamers, respectively, *cis*-CaaD appears to be a homotrimeric protein as assessed by gel filtration chromatography. On the basis of the known three-dimensional structures and reaction mechanisms of CaaD and 4-OT, a sequence alignment implicated Pro-1, Arg-70, Arg-73, and Glu-114 as important active-site residues in *cis*-CaaD. Subsequent site-directed mutagenesis experiments confirmed these predictions. The acetylene compounds, 2-oxo-3-pentynoate and 3-bromo- and 3-chloropropiolate, were processed by *cis*-CaaD to products consistent with an enzyme-catalyzed hydration reaction previously established for CaaD. Hydration of 2-oxo-3-pentynoate afforded acetopyruvate, while the 3-halopropiolates became irreversible inhibitors that modified Pro-1. The results of this work revealed that *cis*-CaaD and CaaD have different primary and quaternary structures, and display different substrate specificity and catalytic efficiencies, but likely share a highly conserved catalytic mechanism. The mechanism may have evolved independently because sequence analysis indicates that *cis*-CaaD is not a 4-OT family member, but represents the first characterized member of a new family in the tautomerase superfamily that probably resulted from an independent duplication of a 4-OT-like sequence. The discovery of a fifth family of enzymes within this superfamily further demonstrates the diversity of activities and structures that can be created from 4-OT-like sequences.

Over the course of the last century, large quantities of 1,3-dichloropropene and related compounds have been released into the environment due to the large-scale application of the compounds as soil fumigants in agriculture (*1* and refs therein). Strikingly, numerous strains of bacteria have been discovered in both contaminated and pristine soil, which are capable of utilizing 1,3-dichloropropene or the degradation product, 3-chloroacrylate (**1**, Scheme 1), as their sole source of carbon and energy (*1*−*4*). *Pseudomonas pavonaceae* 170 and coryneform bacterium strain FG41 are two such organisms able to grow on one or both isomers of 3-chloroacrylate (*1*, *3*). This capability is due, in part, to the presence of isomer-specific dehalogenases (i.e., *cis*- and *trans*-3-chloroacrylic acid dehalogenase), which are involved in the hydrolytic cleavage of the vinylic carbon−halogen bond in 3-haloacrylates (e.g., **1** and **2**) to afford malonate semialdehyde (**5**). Recently, the genes for the α- and *β*-subunits of *trans*-3-chloroacrylic acid dehalogenase (CaaD)[1] from the 1,3-dichloropropene-degrading bacterium *P. pavonaceae* 170 have been cloned, the enzyme expressed, and the key elements of the mechanism determined (*5*, *6*).

Scheme 1



1: X = Cl; **2**: X = Br — **3**: X = Cl; **4**: X = Br — **5**

The heterohexameric CaaD consists of three α-subunits, each having 75 amino acid residues, and three *β*-subunits, each having 70 amino acid residues. Sequence analysis and results of mutagenesis studies uncovered a relationship between the α- and *β*-subunits of CaaD and members of the 4-oxalocrotonate tautomerase (4-OT) family of enzymes (*5*). The 4-OT family is one of the four major families of the

† This research was supported by the National Institutes of Health Grant GM-65324 and the Robert A. Welch Foundation (F-1334).

* To whom correspondence should be addressed: Tel.: 512-471-6198; fax: 512-232-2606; e-mail: whitman@mail.utexas.edu.

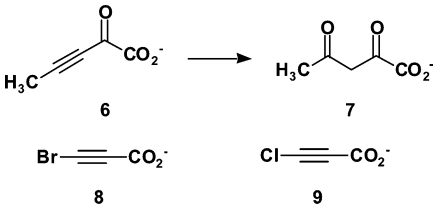
‡ Division of Medicinal Chemistry.

‖ Division of Pharmacology and Toxicology.

§ MRC Centre for Protein Engineering.

10.1021/bi0355948 CCC: $27.50
Published on Web 12/24/2003

Scheme 2



tautomerase superfamily (*7*−*9*). 4-OT family members consist of small subunits (typically 61−81 amino acid residues) and have an amino-terminal proline, which plays a critical mechanistic role (*8*). Both the α- and *β*-subunits of CaaD have an amino-terminal proline, but only *β*-Pro-1 has been identified as an essential catalytic group (*5*).

CaaD was proposed to be a hydrolytic dehalogenase, which was substantiated by the finding that CaaD catalyzes the hydration of 2-oxo-3-pentynoate (**6**, Scheme 2) to afford acetopyruvate (**7**) (*6*). However, a covalent enzyme−substrate intermediate, observed for some haloaliphatic and haloaromatic dehalogenases (*10*, *11*), was not favored on the basis of the sequence analysis. Hence, the proposed mechanism for CaaD likely involves the formation of an unstable halohydrin species (**3**, Scheme 1), which collapses to produce malonate semialdehyde (**5**) (*5*, *6*). Whether or not the collapse of **3** (or **4**) is enzyme-catalyzed remains unknown.

In contrast to CaaD, little is known about the *cis*-specific 3-chloroacrylic acid dehalogenases (*cis*-CaaDs). A *cis*-CaaD that has been isolated from coryneform bacterium strain FG41 is reportedly a dimeric or trimeric protein consisting of 16.2 kDa subunits (*3*). A comparison of its N-terminal amino acid sequence (48 amino acids) with the α- and *β*-subunits of CaaD revealed little sequence identity, but the amino-terminal proline was conserved. The presence of Pro-1, coupled with differences in primary and quaternary structures and the high specificity of CaaD for the *trans*-isomer, raises several exciting and provocative questions about the mechanistic, structural, and evolutionary relationship between these isomer-specific dehalogenases, the relationship between *cis*-CaaD and the 4-OT family of enzymes, and the specificity of *cis*-CaaD as well as the structural basis for the isomer specificity in the two dehalogenating enzymes.

To set the stage for in-depth mechanistic and structural studies that would ultimately address these questions, we cloned and sequenced the gene encoding *cis*-CaaD from coryneform bacterium strain FG41, expressed the enzyme, and carried out kinetic, mechanistic, and inhibition studies on the wild type enzyme. In addition, sequence analysis coupled with site-directed mutagenesis identified Pro-1, Arg-70, Arg-73, and Glu-114 as critical catalytic residues. The combined results delineated similarities as well as intriguing differences between *cis*-CaaD and CaaD. While the hydrolytic nature and the mechanistic elements of the reactions are largely the same, it was surprising to find that *cis*-CaaD and CaaD are not members of the same family. Sequence analysis suggests that *cis*-CaaD belongs to a new family in the tautomerase superfamily that may have evolved from an independent duplication of a 4-OT like sequence. This observation further emphasizes the usefulness of the *β*−α−*β* structural motif, which is the basic building unit for the tautomerase superfamily members, in the creation of new activities and structures.

[1] Abbreviations: Ap, ampicillin; CaaD, *trans*-3-chloroacrylic acid dehalogenase; *cis*-CaaD, *cis*-3-chloroacrylic acid dehalogenase; CHMI, 5-carboxymethyl-2-hydroxymuconate isomerase; DMSO, dimethyl sulfoxide; ESI-MS, electrospray ionization mass spectrometry; HPLC, high-pressure liquid chromatography; LB, Luria-Bertani; 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; MSAD, malonate semialdehyde decarboxylase; NMR, nuclear magnetic resonance; dNTPs, deoxynucleotide triphosphates; 4-OT, 4-oxalocrotonate tautomerase; PCR, polymerase chain reaction; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

MATERIALS AND METHODS

Materials. Chemicals, biochemicals, buffers, 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**), 3-bromo- and 3-chloropropiolate (**8** and **9**, respectively) and the cis- and trans-isomers of 3-bromoacrylate (**2**) (*12*−*14*). Tryptone, yeast extract, and agar were obtained from Becton, Dickerson, and Company (Franklin Lakes, NJ). Enzymes and reagents used for the molecular biology procedures, DNA ladders, protein molecular weight standards, deoxynucleotide triphosphates (dNTPs), the high pure plasmid isolation kit, the high pure PCR product purification kit, and multipurpose agarose were purchased from F. Hoffmann-La Roche, Ltd. (Basel, Switzerland). The Amicon concentrator and the YM10 ultrafiltration membranes were obtained from Millipore Corp. (Bedford, MA). Prepacked PD-10 Sephadex G-25 columns were purchased from Biosciences AB (Uppsala, Sweden). Oligonucleotides for DNA amplification and sequencing were synthesized by Genosys (The Woodlands, TX).

Bacterial Strains, Plasmids, and Growth Conditions. Genomic DNA of coryneform bacterium strain FG41, an organism capable of utilizing *cis*- and *trans*-3-chloroacrylic acid as growth substrates (*3*), was used as the source for the cloning experiments described herein. This strain was a kind gift from Professor Dick B. Janssen (Department of Biochemistry, University of Groningen, The Netherlands). *Escherichia coli* strains JM101 and XL1Blue (Stratagene, La Jolla, CA) and plasmid pGEM-T (Promega Corp., Madison, WI) were used for direct cloning of PCR products. *E. coli* strain BL21(DE3) (Stratagene) was used in combination with the T7 expression system (pET3b vector; Promega Corp.) for overexpression of *cis*-CaaD and the mutant enzymes. Strain FG41 was grown at 30 °C in nutrient broth medium (*3*). *E. coli* strains were grown at 37 °C in Luria-Bertani (LB) medium. When required, Difco agar (15 g/L) and ampicillin (100 *μ*g/mL) were added to the medium. For blue-white color screening, media were supplemented with 5-bromo-4-chloro-3-indolyl-*β*-D-galactopyranoside (X-gal; 40 *μ*g/mL) and isopropyl-*β*-D-thiogalactoside (IPTG; 0.4 mM).

General Methods. General procedures for cloning and DNA manipulation were performed as described elsewhere (*15*). The PCR was carried out in a Perkin-Elmer DNA thermocycler model 480 obtained from Perkin Elmer Inc. (Wellesley, MA). DNA sequencing was performed by the DNA Core Facility in the Institute for Cellular and Molecular

Biology (ICMB) at The University of Texas at Austin. HPLC was performed on a Waters (Milford, MA) 501/510 system using either a TSKgel DEAE-5PW (anion exchange) or a TSKgel Phenyl-5PW (hydrophobic interaction) column (Tosoh Bioscience, Montgomeryville, PA). Protein was analyzed by polyacrylamide gel electrophoresis (PAGE) under either denaturing conditions using sodium dodecyl sulfate (SDS) or native conditions on gels containing 15% polyacrylamide. The gels were stained with Coomassie brilliant blue. Protein concentrations were determined with the Bio-Rad protein assay (Bio-Rad Laboratories, Hercules, CA). The native molecular masses of *cis*-CaaD and mutant enzymes were determined by gel filtration on a Superose 12 column (Pharmacia Biotech AB, Uppsala, Sweden) using the Waters 501/510 HPLC system. Kinetic data were obtained on a Hewlett-Packard 8452A Diode Array spectrophotometer. The cuvettes were mixed using a stirr/add cuvette mixer (Bel-Art Products, Pequannock, NJ). 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. Nuclear magnetic resonance (NMR) spectra were recorded in 100% H_2O on a Varian Unity INOVA-500 spectrometer using selective presaturation of the water signal with a 2-s presaturation interval. The lock signal is dimethyl-d_6 sulfoxide (DMSO-d_6). Chemical shifts are standardized to the DMSO-d_6 signal at 2.49 ppm.

Sequence Analysis. Nucleotide sequence data were analyzed by using the programs LALIGN, TRANSLATE, PROTPARAM, and WEBCUTTER, available on the World Wide Web (http://www.rna.icmb.utexas.edu/linxs/index.html). BLAST and iterative PSI-BLAST searches of the National Center for Biotechnology Information (NCBI) databases were performed using the *cis*-CaaD amino acid sequence as the query sequence (*16*). The databases were searched with the "NR" option (all nonredundant GenBank CDS translations + PDB + SwissProt + PIR + PRF). The conserved domain database search (CD-search), a default option of the NCBI PSI-BLAST search, was also performed. In addition, the NCBI microbial database containing finished and unfinished genome sequences has been searched using the NCBI genomic BLAST server. Amino acid sequences of each family within the tautomerase superfamily were aligned using a version of the CLUSTALW multiple-sequence alignment routines available in the computational tools at the EMBL-EBI Web site (*17*). Different family alignments were combined together manually with the help of the previously constructed structural alignment of known members of the tautomerase superfamily (*7*).

PCR Amplification, Cloning, and Sequencing of the cis-caaD Gene. A 154-bp fragment of the *cis-caaD* gene was obtained by PCR synthesis of a portion of the gene using two degenerate primers that were designed on basis of the N-terminal amino acid sequence of the enzyme isolated from coryneform bacterium strain FG41 (*3*). The forward primer, 5′-GAT**GGATCC**XGTXTAYATGGTXTAYGT-3′, where X = G, A, T, C; Y = T, C, contains a *Bam*HI restriction site (in bold), while the reverse primer, 5′-GCC**GAATTC**-ZACZAGZACYTTYAAXTGZAC-3′, where X = G, A, T, C; Y = T, C; Z = G, A, contains an *Eco*RI restriction site (in bold). Total genomic DNA from strain FG41 was isolated by a phenol extraction procedure described elsewhere (*1*). The amplification reaction mixture contained standard *Taq* amplification buffer, 200 μM of each dNTP, 100 ng of each primer, 100 ng of total genomic DNA, and 2 units of *Taq* DNA polymerase. The cycling parameters were 94 °C for 5 min followed by 30 cycles of 94 °C for 60 s, 55 °C for 60 s, and 72 °C for 90 s, with a final elongation step of 72 °C for 10 min. The amplified DNA was purified, digested with *Bam*HI and *Eco*RI and ligated into pBluescript SK$^-$, which was previously digested with the same restriction enzymes. An aliquot of the ligation mixture was transformed into competent *E. coli* JM101 cells, and transformants were selected on LB/Ap plates at 37 °C. Plasmid DNA was isolated from several transformants and analyzed by restriction analysis for the presence of the insert. One was chosen at random and the nucleotide sequence of the cloned dehalogenase gene fragment was determined.

Two oligonucleotide primers were designed on basis of this newly determined nucleotide sequence of a portion of the *cis-caaD* gene. The forward primer (designated F1) had the sequence 5′-GGACTGACTGGAACCCAGCAC-3′ and the reverse primer (designated R1) had the sequence 5′-GGCCAGGAAGTGCTG-3′. Primer F1 was used in combination with the T7 promoter primer (5′-TAATACGACTCACTATAGGG-3′) to amplify the 3′ end and downstream region of the *cis-caaD* gene from a *Sal*I genomic DNA library made in pBluescript SK$^-$. Primer R1 was used in combination with the T7 promoter primer to amplify the 5′ end and upstream region of the *cis-caaD* gene from a *Bam*HI genomic DNA library made also in pBluescript SK$^-$. The *Sal*I and *Bam*HI genomic DNA libraries were made as follows. Total genomic DNA of strain FG41 and the pBluescript SK$^-$ vector were digested with either *Sal*I or *Bam*HI restriction enzymes, purified, and ligated using T4 DNA ligase. These ligation mixtures were used directly as the template for PCR amplification of the *cis-caaD* gene and its flanking regions. PCRs were carried out in the Perkin-Elmer DNA thermocycler using the synthetic primers, the ligation mixtures (~50 ng of DNA), and the PCR reagents supplied in the Expand High Fidelity PCR system (F. Hoffmann-La Roche, Ltd.) following the manufacturer's instructions. The combination of the F1 and T7 promoter primers yielded a single PCR product of about 2 kb using the *Sal*I genomic library as the template, while the combination of the R1 and T7 promoter primers yielded a single PCR product of about 1 kb using the *Bam*HI genomic library as the template. The two PCR products were purified and directly ligated in the TA-cloning vector pGEM-T. Aliquots of the resulting mixtures were transformed into competent *E. coli* JM101 cells. Transformants were selected at 37 °C on LB/Ap plates that contained IPTG and X-gal for blue-white color screening. Plasmid DNA was isolated from several white colonies and analyzed by restriction analysis for the presence of the insert. The cloning procedure was repeated twice for separate PCRs to identify possible errors that might have been introduced during amplification of the *cis-caaD* gene region. All cloned PCR products were (partially) sequenced and compared with each other to ensure that no mutations had been introduced during the PCR. In this manner, the nucleotide sequence for the entire *cis-caaD* gene was determined.

Construction of the Expression Vector for the Production of cis-CaaD. The *cis-caaD* gene was amplified by the PCR

using two synthetic primers, genomic DNA, and the PCR reagents supplied in the Expand High Fidelity PCR system following the protocol supplied with this system. The forward primer (5′-ATA**CATATG**CCGGTTTATATGGTTTAC-3′) is designated primer F and contains an *Nde*I restriction site (in bold) followed by 18 bases corresponding to the coding sequence of the *cis-caaD* gene. The reverse primer (5′-CAT**GGATCC**CTAGGTGCGAGAGACGTCCACGTT-3′) is designated primer R and contains a *Bam*HI restriction site (in bold) followed by 24 bases corresponding to the complementary sequence of the *cis-caaD* gene. 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. Aliquots of the ligation mixture were transformed into competent *E. coli* BL21(DE3) cells. Transformants were selected at 37 °C on LB/Ap plates, where no NaCl was added to the LB medium. Single colonies were screened for *cis*-CaaD activity by monitoring halide production upon incubation with *cis*-**1** as described below. Plasmid DNA was isolated from a colony showing dehalogenase activity and the presence of the insert was further verified by restriction analysis. The cloned *cis-caaD* gene was sequenced to verify that no mutations had been introduced during the amplification of the gene. The newly constructed expression vector was named pCC5.

Construction of cis-CaaD Mutants. The *cis*-CaaD mutants were constructed using the coding sequence for the dehalogenase in plasmid pCC5 as the template. The P1A mutant was generated by the PCR using the primer 5′-ATA-**CATATG**GCGGTTTATATGGTTTAC-3′ where the *Nde*I site is in bold and the mutated codon is underlined. This primer corresponds to the 5′ end of the wild-type coding sequence and was used in combination with primer R. The R70A, R73A, and E114Q mutants were generated by overlap extension PCR (*18*). Primers F and R were used as the external primers. For the R70A mutant, the internal PCR primers were oligonucleotides 5′-GTTCACGGATTGCAT-**GCG**GAGGGACGTAGC-3′ and 5′-GCTACGTCCCTC-**CGC**ATGCAATCCGTGAAC-3′, where the mutated codon is shown in bold. For the R73A mutant, the internal PCR primers were oligonucleotides 5′-CATAGGGAGGGA**GC-G**AGCGCCGACCTC-3′ and 5′-GAGGTCGGCGCT**CGC**TC-CCTCCCTATG-3′, where the mutated codon is shown in bold. For the E114Q mutant, the internal PCR primers were oligonucleotides 5′-AGCAGATGGTG**CAG**TACGGCCG-GTTC-3′ and 5′-GAACCGGCCGTA**CTG**CACCATCT-GCTG-3′, where the mutated codon is shown in bold. The amplification mixtures contained the appropriate synthetic primers, pCC5 DNA, and the PCR reagents supplied in the Expand High Fidelity PCR system. The restriction sites *Nde*I and *Bam*HI, introduced during the amplification reaction, were used to clone the purified PCR products into plasmid pET3b for overexpression of the *cis*-CaaD mutants. The cloned dehalogenase genes were sequenced to verify that only the desired mutations had been introduced during the PCR.

Expression and Purification of cis-CaaD. The *cis*-CaaD and the mutant enzymes were produced constitutively in *E. coli* BL21(DE3) using the T7 expression system. Fresh BL21(DE3) transformants containing the desired plasmid were collected from a plate by resuspending them in LB medium (1 mL), which was then used to inoculate 1 L of LB/ampicillin medium to a starting OD_{600} of about 0.05. After overnight growth, cells were harvested by centrifugation (10 min at 10000*g*), washed with 0.1 volume of 10 mM Tris-SO_4 buffer, pH 8 (buffer A), and stored at −20 °C until further use.

In a typical purification experiment, cells of two 1 L-cultures were thawed, combined, and suspended in 15 to 20 mL of buffer A. Cells were disrupted by sonication for 30 s per mL of suspension at a 60 W output in a W385 sonicator of Heat systems-ultrasonics, Inc. (Farmingdale, NY), after which unbroken cells and debris were removed by centrifugation (30 min at 20000*g*). The supernatant was filtered through a 0.2 μm-pore diameter filter and applied to the TSKgel DEAE-5PW column (150 × 21.5 mm), which had previously been equilibrated with buffer A. The column was washed with 50 mL of buffer A, and retained proteins were eluted with a 300-mL increasing linear gradient of 0 to 0.5 M Na_2SO_4 in buffer A at a flow rate of 5 mL/min. Fractions (10 mL) that showed the highest *cis*-CaaD activity were pooled and concentrated to about 15 mL using an Amicon stirred cell equipped with a YM10 (10 000 MW cutoff) ultrafiltration membrane. Subsequently, $(NH_4)_2SO_4$ was added to a concentration of 1 M and the resulting solution was stirred for 60 min at 4 °C. After centrifugation (30 min at 20000*g*), the supernatant was loaded onto the TSKgel Phenyl-5PW column (150 × 21.5 mm), which had previously been equilibrated with buffer B (1 M $(NH_4)_2SO_4$ in buffer A). The column was washed with 50 mL of buffer B, and retained proteins were eluted with a 250-mL decreasing linear gradient of 1.0 to 0 M $(NH_4)_2SO_4$ in buffer A at a flow rate of 5 mL/min. Fractions (10 mL) with the highest *cis*-CaaD activity were pooled and concentrated to about 3 mL. The concentrate was loaded onto the Sephadex G-75 column (100 × 2.5 cm), previously equilibrated with buffer A. The protein was eluted with buffer A at a flow rate of 1 mL/min. Fractions (9 mL) with the highest *cis*-CaaD activity were analyzed by SDS−PAGE, and those that contained purified enzyme were pooled and concentrated to a protein concentration of about 10 mg/mL. The purified enzymes were filtered through a 0.2 μm-pore diameter filter and stored at 4 °C. The inactive mutants were identified by their masses on SDS−PAGE gels.

Mass Spectrometric Characterization of cis-CaaD and cis-CaaD Mutants. The masses of *cis*-CaaD and the four mutants were determined using an LCQ electrospray ion trap mass spectrometer (ThermoFinnigan, San Jose, CA), housed in the Analytical Instrumentation Facility Core in the College of Pharmacy at the University of Texas at Austin. The protein samples were made up as described elsewhere (*6*). The observed monomer mass for *cis*-CaaD was 16 622 Da (calc. 16 624 Da). The observed monomer mass for the P1A mutant was 16 596 Da (calc. 16 596 Da), that of the R70A mutant was 16 536 Da (calc. 16 537 Da), that of the R73A mutant was 16 537 Da (calc. 16 537 Da), and that of the E114Q was 16 620 Da (calc. 16 621 Da).

Enzyme Assays and Kinetic Studies. HPLC fractions and single colonies on agar plates were screened for *cis*-CaaD activity by monitoring halide production upon incubation with *cis*-**1**. Accordingly, an appropriate amount of enzyme or cells was incubated in a microtiter plate with *cis*-**1** (150 μL of a 10 mM solution) in 50 mM Tris-SO_4 buffer (pH 8.2). After incubation of the plate at 37 °C for 5−60 min,

100 μL of a 0.25 M solution of $NH_4Fe(SO_4)_2$ in 6 M HNO_3 followed by a drop of saturated $Hg(SCN)_2$ in ethanol were added. A red color indicated the presence of *cis*-CaaD activity.

The kinetic assays were performed at 22 °C by following the decrease in absorbance at 224 nm, which corresponds to the hydration of *cis*-**1** ($\epsilon = 2900$ M^{-1} cm^{-1}) and *cis*-**2** ($\epsilon = 3600$ M^{-1} cm^{-1}). An aliquot of *cis*-CaaD was diluted into 20 mM Na_2HPO_4 buffer (pH 9.0), yielding a final enzyme concentration of 0.05 μM (1 μM for the E114Q mutant), and incubated for 60 min at 22 °C. Subsequently, a 1-mL portion was transferred to a cuvette and the enzyme activity was assayed by the addition of a small quantity of substrate from either a 1, 5, or 50 mM stock solution. The 50 mM stock solution was made up in 100 mM Na_2HPO_4 buffer (pH 9.0). The addition of *cis*-**1** or *cis*-**2** (as the free acid) to this buffer adjusted the pH of the stock solution to about 7. The 1 and 5 mM stock solutions were made up by dilution of an aliquot of the 50 mM stock solution into 20 mM NaH_2PO_4 buffer, pH 7.3. The concentrations of substrate used in the assay ranged from 1 to 200 μM.

*^{1}H NMR Spectroscopic Detection of **5***. A series of ^{1}H NMR spectra monitoring the *cis*-CaaD-catalyzed transformation of *cis*-**1** were recorded as follows. An amount of *cis*-**1** (4 mg, 0.04 mmol) dissolved in DMSO-d_6 (30 μL) was added to 100 mM Na_2HPO_4 buffer (0.6 mL, pH ~9) and placed in an NMR tube. The pH of the reaction mixture was adjusted to 8.5. Subsequently, an aliquot of *cis*-CaaD (50 μL of a 2.4 mg/mL solution made up in 20 mM Na_2HPO_4 buffer, pH 7.3) was added to the reaction mixture. The first ^{1}H NMR spectrum was recorded 4 min after the addition of enzyme and every 3 min thereafter for a total reaction time of 30 min. The final pH of the reaction mixture was 6.67. The signals for malonate semialdehyde (**5**) and the hydrate are reported elsewhere (*6*). In addition, signals are also detected in the later spectra for acetaldehyde (and the hydrate). Acetaldehyde results from the nonenzymatic decarboxylation of **5** (*6*).

*UV and ^{1}H NMR Spectroscopic Detection of Acetopyruvate (**7**) in the cis-CaaD-Catalyzed Hydration of 2-Oxo-3-pentynoate (**6**)*. The hydration of 2-oxo-3-pentynoate (**6**) by *cis*-CaaD was monitored by following the formation of acetopyruvate (**7**) at 294 nm ($\epsilon = 7000$ M^{-1} cm^{-1}) in 20 mM Na_2HPO_4 buffer (pH 9.0) at 22 °C. An aliquot of *cis*-CaaD was diluted into 20 mL of 20 mM Na_2HPO_4 buffer, pH 9.0, yielding a final enzyme concentration of 4 μM. The diluted enzyme was incubated for 60 min at 22 °C. Subsequently, 1-mL aliquots were transferred to a cuvette, and the assay was initiated by the addition of a small quantity (1−8 μL) of **6** from either a 10 or 100 mM stock solution. The 100 mM stock solution was made up by dissolving the appropriate amount of **6** in 100 mM Na_2HPO_4 buffer (pH 9.1). The pH of the solution was adjusted to ~7 using small amounts of a NaOH solution. The 10 mM stock solution was made by diluting an aliquot of the 100 mM stock solution into 100 mM NaH_2PO_4 buffer, pH 7.3. The concentrations of **6** used in the assay ranged from 50 to 800 μM.

For the ^{1}H NMR spectroscopic identification of **7**, it was necessary to incubate small quantities of **6** with *cis*-CaaD over a 26-h interval. A stock solution of **6** (8 mg, 72 mmol) was made up in 100 mM Na_2HPO_4 buffer, pH 9.2 (1.2 mL) and the pH of the solution was adjusted to 7.2 using aliquots of 1 M NaOH. Subsequently, eight individual reaction mixtures were made up containing water (900 μL) and an aliquot of **6** (100 μL) from the stock solution. A quantity of *cis*-CaaD (50 μL of a 20 mg/mL solution) was added and the reaction mixtures were incubated at 23 °C. After 2 h, a second aliquot of *cis*-CaaD (30 μL of a 20 mg/mL solution) was added to the individual mixtures. After an additional 3-h incubation period, a third aliquot (20 μL of a 20 mg/mL solution) was added. After the reaction mixtures were incubated at 23 °C for an additional 21 h, they were combined and the enzyme was removed using an Amicon concentrator equipped with a YM3 ultrafiltration membrane. The resulting effluent was concentrated to ~0.7 mL in vacuo and placed in an NMR tube along with DMSO-d_6 (30 μL). The ^{1}H NMR spectra show the presence of signals corresponding to **7** (*19*), its hydrate at C-2 (*19*), and the enol isomer at C-2 (*19*).

*Irreversible Inactivation of cis-CaaD by **8** or **9***. The irreversibility of *cis*-CaaD inactivation by **8** or **9** was established as follows. The enzyme (1 mg/mL or 20 μM based on the molecular mass of the native enzyme) was incubated with an excess of **8** or **9** (10 mM) in 20 mM NaH_2PO_4 buffer, pH 7.3 (100 μL), for 1 h at 22 °C. In a separate control experiment, the same quantity of *cis*-CaaD was incubated without inhibitor under otherwise identical conditions. The reaction mixtures were dialyzed for 24 h with 20 mM NaH_2PO_4 buffer, pH 9.0. The dialyzed enzymes were assayed for residual activity as described above.

Mass Spectral Analysis of the Modified cis-CaaD and Peptide Mapping. Three samples were made up as follows. The enzyme (20 μM based on the molecular mass of the native enzyme) was incubated with an excess of either **8** or **9** (1 mM) in 0.5 mL of 20 mM NaH_2PO_4 buffer (pH 7.3) for 16 h at 4 °C. In a separate control, the same quantity of enzyme was incubated without inhibitor under otherwise identical conditions. Subsequently, the three samples were loaded onto individual PD-10 Sephadex G-25 gel filtration columns, which had previously been equilibrated with 100 mM $(NH_4)HCO_3$ buffer (pH 8.0). The protein was eluted with the same buffer by gravity flow. Fractions (0.5 mL) were analyzed for the presence of protein by UV absorbance at 214 nm. The purified enzymes were assayed for residual activity as described above. Samples treated with **8** and **9** had no activity while the control sample retained full activity. Subsequently, the samples were analyzed by ESI-MS (*16*) and used in the following peptide mapping experiments.

For the peptide mapping studies, a quantity of unmodified *cis*-CaaD and *cis*-CaaD modified by **8** or **9** (~27 μg in 27 μL of 100 mM $(NH_4)HCO_3$ buffer, pH 8.0) was combined with 3 μL of 10 M guanidine HCl and incubated for 1 h at 37 °C. Subsequently, sequencing grade protease V-8 (2 μL of a 10 mg/mL stock solution made up in water) was used to digest the **8**- and **9**-modified and the unmodified protein for 48 h at 37 °C. The samples were analyzed by MALDI-MS without further purification as described elsewhere (*6*). Selected ions of the sample of *cis*-CaaD modified by **8** were subjected to MALDI−PSD analysis (*6*).

*Kinetics of Irreversible Inhibition of cis-CaaD by **9***. The time- and concentration-dependent inactivation of *cis*-CaaD by **9** was determined as follows. Incubation mixtures (total volume of 100 μL) containing varying amounts of **9** (1−

100 μM) and enzyme (15 μM) in 20 mM NaH_2PO_4 buffer (pH 7.3) at 22 °C were made up in 1.5 mL eppendorf micro test tubes. Aliquots (5 μL) were removed at various time intervals, diluted into 1 mL of 20 mM Na_2HPO_4 buffer (pH 9.0), and assayed for residual *cis*-CaaD activity. The activity assay was initiated by the addition of a small quantity (3 μL) of *cis*-**1** from a 50 mM stock solution made up in 100 mM Na_2HPO_4 buffer. The addition of *cis*-**1** to this buffer adjusted the pH of the stock solution to about 7. Stock solutions of **9** were made up in 100 mM Na_2HPO_4 buffer and the pH of the stock solutions was adjusted to 7.3. The zero time point (i.e., 100% activity) was determined at each inhibitor concentration by removing an aliquot just before the addition of inhibitor.

Protection Studies. Protection against inactivation of *cis*-CaaD by **9** was carried out as described above with the following modifications. The enzyme (15 μM based on the native molecular mass) was incubated with varying concentrations of *cis*-**1** (0−2.5 mM) in 20 mM NaH_2PO_4 buffer (pH 7.3) at 22 °C. After a 30-s interval, a fixed concentration of **9** (60 μM) was added to the mixture. Aliquots (5 μL) were removed at various time intervals, diluted into 1 mL of 20 mM Na_2HPO_4 buffer (pH 9.0), and assayed for residual activity. The data were plotted and analyzed as described above.

RESULTS

Cloning and Characterization of the cis-caaD Gene. The coryneform bacterium strain FG41 has previously been shown to express both a CaaD and a *cis*-CaaD, when grown on either isomer of **1** (*3*). Purification of the *cis*-CaaD to homogeneity and amino-terminal sequencing of the protein identified the first 48 residues of the enzyme (*3*). Using two degenerate primers based on this amino-terminal sequence and the PCR, a 154-bp DNA fragment of the *cis-caaD* gene was isolated from genomic DNA of coryneform bacterium strain FG41. The availability of this DNA fragment enabled the design of two oligonucleotide primers, which were then used in combination with a T7 promoter primer to amplify the 3′ end and downstream region as well as the 5′ end and upstream region of the *cis-caaD* gene from genomic DNA libraries made in pBluescript SK^-. In this manner, a 3-kb DNA segment, which included the *cis-caaD* gene sequence, was isolated from the genomic DNA of FG41. Subsequently, a 1.3-kb fragment containing the open reading frame for *cis*-CaaD was sequenced. The cloning and sequencing procedures were repeated twice for separate PCRs to verify that no mutations had been introduced during amplification of the gene. The *cis-caaD* gene codes for a protein of 150 amino acids with a calculated molecular mass of 16 756 Da, which fits the experimentally determined mass of 16.2 kDa (*3*).[2] The N-terminal amino acid sequence, determined by Edman degradation, was fully conserved in the translated amino acid sequence of the *cis-caaD* gene (*3*).

Expression, Purification, and Characterization of cis-CaaD. The *cis-caaD* gene was amplified from genomic DNA of strain FG41 and fused into the start codon of the expression vector pET3b, resulting in the construct pCC5. Sequencing of the cloned dehalogenase gene verified that no mutations had been introduced during the amplification and cloning procedures. The *cis-caaD* gene in pCC5 is under transcriptional control of a T7 promoter and the dehalogenase was produced constitutively in a soluble and active form in *E. coli* BL21(DE3). The recombinant enzyme was purified by a four-step protocol (using three columns), which typically provides 15 mg of homogeneous enzyme per liter of culture. The purified enzyme was analyzed by ESI-MS and gel filtration chromatography. A sample of *cis*-CaaD generates one major peak in the mass spectrometer that corresponds to a mass of 16 622 ± 1 Da. A comparison of this value to the calculated molecular mass (16 756 Da) of the *cis-caaD* gene product indicates that the initiating methionine is removed during posttranslational processing, resulting in a protein with an N-terminal proline. The native molecular mass of *cis*-CaaD was estimated by gel filtration to be about 44 kDa, suggesting that the native enzyme might be a homotrimeric protein.

Table 1: Kinetic Parameters for the CaaD- and *cis*-CaaD-catalyzed Conversion of **6** to **7**[a]

enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
CaaD[b]	0.7 ± 0.02	110 ± 4	6.4×10^3
cis-CaaD	0.007 ± 0.0005	620 ± 60	0.011×10^3

[a] The steady-state kinetic parameters were determined in 20 mM sodium phosphate buffer (pH 9.0) at 22 °C. Errors are standard deviations. [b] These kinetic parameters were obtained from Wang et al. (*6*).

*^{1}H NMR Spectroscopic Assay for cis-CaaD Activity Using cis-**1**.* A mixture containing *cis*-CaaD and *cis*-**1** was monitored by ^{1}H NMR spectroscopy to verify that the product of the reaction is **5**, as previously reported (*3*). The enzymatic conversion of *cis*-**1** yields **5**, as indicated by a doublet at 3.20 ppm and a triplet at 9.50 ppm (data not shown), which correspond to the protons at C-2 and C-3, respectively. In addition, signals corresponding to the hydrate of **5** are also present. Hence, **5** is the product of the *cis*-CaaD-catalyzed conversion of *cis*-**1**.

*Hydration of 2-Oxo-3-pentynoate (**6**) by cis-CaaD.* It has previously been determined that CaaD functions as a hydratase, and converts 2-oxo-3-pentynoate (**6**) to acetopyruvate (**7**) with a k_{cat} of 0.7 s^{-1} (*6*) (Table 1). This observation prompted us to examine whether *cis*-CaaD catalyzes the hydration of **6**. Interestingly, we observe a small but significant *cis*-CaaD activity with this compound, suggesting that *cis*-CaaD also functions as a hydratase. The kinetic parameters for the conversion of **6** to **7** are summarized in Table 1. The value of k_{cat} for *cis*-CaaD is 100-fold less than that measured for CaaD, while the K_m value is ∼5.6-fold higher. The net effect is a 582-fold difference in the k_{cat}/K_m value. Clearly, **6** is aligned more favorably for the hydration reaction in the active site of CaaD than it is in the active site of *cis*-CaaD.

While **7** has a characteristic λ_{max} of 294 nm, its identity in the incubation mixture described above was further confirmed by ^{1}H NMR spectroscopy. The ^{1}H NMR spectrum showed signals consistent with the structure of **7** (*19*), as well as two additional species, the hydrate of **7** at C-2 (*19*), and the enol of **7** at C-2 (*19*). In the absence of *cis*-CaaD, **6** is stable for several hours in solution and does not decompose

[2] The nucleotide sequence of the *cis-caaD* gene has been deposited in the GenBank databases under accession number AY334362.

Table 2: Kinetic Parameters for CaaD, *cis*-CaaD, and the E114Q Mutant[a]

enzyme	substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
cis-CaaD	*cis*-**1**	10.2 ± 0.2	30 ± 3	3.4×10^5
	cis-**2**	9.3 ± 0.4	28 ± 4	3.3×10^5
E114Q	*cis*-**1**	0.09 ± 0.01	2.1 ± 0.4	4.3×10^4
	cis-**2**	0.11 ± 0.01	1.9 ± 0.3	5.8×10^4
CaaD[b]	*trans*-**1**	3.8 ± 0.1	31 ± 2	1.2×10^5
	trans-**2**	5.1 ± 0.1	37 ± 2	1.4×10^5

[a] The steady-state kinetic parameters were determined in 20 mM sodium phosphate buffer (pH 9.0) at 22 °C. Errors are standard deviations. [b] These kinetic parameters were taken from Wang et al. (*6*).

to **7** (*12*, *20*). At higher concentrations of **6** (60 mM), the enzyme (200 μM based on monomer concentration) is inactivated, with no detectable conversion of **6** to **7**.

Kinetic Properties of cis-CaaD. The rate of hydration of 3-haloacrylates by *cis*-CaaD was measured by following the loss in absorbance at 224 nm at 22 °C. Both *cis*-**1** and *cis*-**2** (Scheme 1) are substrates for the enzyme and the kinetic parameters are summarized in Table 2. The data clearly show that *cis*-CaaD processes both *cis*-3-haloacrylates with similar kinetic parameters and catalytic efficiency. While the K_m values are comparable to those observed for the CaaD-catalyzed hydration of *trans*-**1** and *trans*-**2** (Table 2), the k_{cat} values are slightly higher (2.7-fold for *cis*-**1** and 1.8-fold for *cis*-**2**). As a result, the k_{cat}/K_m values are 2.8- and 2.3-fold higher for *cis*-**1** and *cis*-**2**, respectively.

Under these conditions, activity was not detected for the trans isomers of **1** and **2**. However, prolonged incubation of *cis*-CaaD with *trans*-**1** and *trans*-**2** (in separate reactions) showed a very low amount of activity with *trans*-**2** (i.e., the rate of the reaction is <0.1% that of the rate observed with *cis*-**2**), while no activity was detected with *trans*-**1**. These results indicate that *cis*-CaaD is highly specific for the cis-isomers of **1** and **2**.

Identification of the cis-CaaD Family in the Tautomerase Superfamily. A sequence similarity search with the amino acid sequence of *cis*-CaaD as the query was performed in the NCBI database using the BLASTP program (*16*), as well as in the NCBI microbial database containing finished and unfinished genome sequences using the NCBI genomic BLAST server. These searches yielded three related proteins that shared significant sequence similarity with *cis*-CaaD (Figure 1). The highest sequence identity (34%) was observed between *cis*-CaaD and a hypothetical protein from *Corynebacterium glutamicum* ATCC 13032 (designated Cg10062 in Figure 1). The other two related proteins are hypothetical proteins from *Mycobacterium smegmatis* (designated OrfX and OrfY in Figure 1) that both show ~30% sequence identity with *cis*-CaaD. The relationship between *cis*-CaaD and the tautomerase superfamily is readily detectable by sequence similarity searches. The *cis*-CaaD homologue from *C. glutamicum* is annotated as a 4-OT-like protein. The results of a conserved domain search (CD-search) predict the presence of two 4-OT-like domains in *cis*-CaaD. The results of a PSI-BLAST search further support this prediction. Starting from the second iteration, there are a growing number of hits to the 4-OT family members scoring above the default inclusion threshold of 0.005. The hits are on both halves of the *cis*-CaaD sequence and include the sequences of 4-OT from *Pseudomonas putida* mt-2 (*21*) and the α-subunit of CaaD from *P. pavonaceae* 170 (designated CaaD1 in Figure 1) (*5*). The hit list after four iterations includes many members of the structurally characterized macrophage migration inhibitory factor (MIF) and 5-carboxymethyl-2-hydroxymuconate isomerase (CHMI) families and the functionally characterized malonate semialdehyde decarboxylase (MSAD) family. The MIF, CHMI, and MSAD families belong to the tautomerase superfamily, with their subunits containing two 4-OT-like structural repeats each (*7*−*9*).

The sequences of *cis*-CaaD and its three homologues (Cg10062, OrfX, and OrfY) indicate that they constitute a new family within the tautomerase superfamily. This conclusion is based primarily on the absence of high sequence identity (i.e., <24%) with known members of the other four families. In addition, *cis*-CaaD and the three known homologues have a conserved glutamate residue (corresponding to Glu-114 in *cis*-CaaD) as well as an arginine residue (corresponding to Arg-70 in *cis*-CaaD). These potential catalytic residues (see below) are absent in all members of the other four families except CaaD, where Glu-52 and Arg-8 (both in the α-subunit) are found (Figure 1B). The sequence similarity between the different families of the 4-OT superfamily is generally low so that they cannot be reliably aligned by the sequence-based methods only. Previous structural comparison of 4-OT, CHMI, and MIF allowed an unambiguous structural alignment of their sequences (*7*). Guided by the PSI-BLAST hits, we have added to this alignment the sequences of the α- and β-subunits of CaaD and representatives of the *cis*-CaaD and MSAD families (Figure 1). In the resulting alignment, there are no insertions or deletions in the structurally conserved regions of the three structurally characterized families, with the hydrophobic core sites being occupied by unpolar or, occasionally, neutral residues.

Although the sequence identity between *cis*-CaaD and each subunit of CaaD is low (~20%), the sequence alignments suggest catalytic residues for *cis*-CaaD. Pro-1 in the β-subunit and Glu-52 in the α-subunit likely function as general acid and base catalysts, respectively, in the CaaD-catalyzed hydration reaction (*5*, *6*, *22*). The sequence alignment in Figure 1 indicates that Pro-1 and Glu-114 in *cis*-CaaD may perform analogous roles. Two arginines, αArg-8 and αArg-11, of CaaD may interact with the carboxylate group of the substrate to facilitate both substrate binding and catalysis (*5*, *6*, *22*). The sequence analysis indicates that Arg-70 and Arg-73 of *cis*-CaaD may interact with the carboxylate group of *cis*-**1**.

Characterization of cis-CaaD mutants. On the basis of the preceding sequence analysis, the key catalytic residues of *cis*-CaaD are predicted to be Pro-1, Arg-70, Arg-73, and Glu-114. To investigate the importance of these residues to the mechanism of *cis*-CaaD, four single site-directed mutants were constructed in which Pro-1, Arg-70, and Arg-73 were replaced with an alanine (P1A, R70A, and R73A) and Glu-114 with a glutamine (E114Q). The four mutants were expressed in *E. coli* BL21(DE3) and purified to ~95% homogeneity (as assessed by SDS−PAGE) using the protocol described for the wild-type *cis*-CaaD. The yields (in milligrams of homogeneous protein per liter of cell culture) of the mutants were moderate (~10 mg) in comparison to wild type (~15 mg). Mass spectral analysis of the individual

A

```
                  β1          α                    β2        β3  β4
4-OT          1 PIAQIHILEgrsde-QKETLIREVSEAISRSLDAPLTSVRVIITEMAkgHFGIGGElaskvrr 62
MIF           1 PMFIVNTNVprasvpe--GFLSELTQQLAQATGKPAQYIAVHVVP--dqLMTFSGTndpc-   56
CHMI          1 PHFIVECSDnireeaDLPGLFAKVNPTLAATGIFPLAGIRSRVHWVD--TWQMADGqhdyaf- 60

CaaD2         1 PFIECHIATGLSVA-RKQQLIRDVIDVTNKSIGSDPKIINVLLVEHAEANMSISGRIHGEAASTERTPAVS 70

cis-CaaD      1 PVYMVYVSQDRLTPSAKHAVAKAITDAHRGLTGTQHFLAQVNFQEQPAGNVFLGGVQQGGDT- 62
Cgl0062       1 PTYTCWSQRIRISREAKQRIAEAITDAHHELAHAPKYLVQVIFNEVEPDSYFIAAQSASENH- 62
OrfX          1 PVYTVTMSRGTLNGETKAALAAEITTIHSAVNHVPSTYVNVLFNELAPSNVYTDGKP--AHP- 60
OrfY          1 PLYQIDTVKGRLTPSVKAEIANKVTDIHCQLTGAPDTFVNVVFREYTEGDCFVARKP--EGR- 60

MSAD          1 PLLKFDIFYGRTDA-QIKSLLDAAHGAMVDAFGVPANDRYQTVSQHRPGEMVLEDTGLGYGRSSA- 64
YodA          1 PLLRFDVIEGRDEK-SLQLLLDTAHQAMVEAFGVPERDRYQIVHQHPANELIIQDAGLGFQRTKD- 64
YrdN          1 PLLRFDLIEGRDQS-SLKKLLDVAHNVVVEAFDVPQQDRYQIVHEHPENHMIIEDTGLGFNRTKN- 64
                * u u u           uu  u   u   u     u u u
```

B

```
                  β1          α                    β2        β3  β4
4-OT          1 PIAQIHILEGRSDEQKETLIREVSEAISRSLDAPLT----SVRVIITEMAKGHFGIGGELASkvrr 62
MIF          57 ALCSLHSIGKIGGAQNRNYSKLLCGLLSDRLHISPD----RVYINYYDMNAANVGWNGSTFA    114
CHMI         61 VHMTLKIGAGRSLESRQQAGEMLFELIKTHFAalmesrllALSFEIEELHPTLNFKQnnvhalfk 125

CaaD1         1 PMISCDMRYGRTDEQKRALSAGLLRVISEATGEPRE----NIFFVIREGSGINFVEHGEHLPDYVPGNAND-  67

cis-CaaD     63 IFVHGLHREGRSADLKGQLAQRIVDDVSVAAEIDRK----HIWVYFGEMPAQQMVEYGRFLPQPGHEGEWF- 129
Cgl0062      63 IWVQATIRSGRTEKQKEELLLRLTQEIALILGIPNE----EVWVYITEIPGSNMTEYGRLLMEPGEEEKWF- 129
OrfX         61 LIINGWVRTGHSDEQTTALVTQVADAATRITGIPAE----RVLVIIGNSPARFAIEGGRILPDPGQELAWL- 127
OrfY         61 SFLGGQIRHGRSVETRQAMLKALRDMWVQTTGQSEA----ELIVGISEVDPRMVLEAGFFMPEPGQEKAWF- 127

MSAD         65 VVLLTVISRPRSEEQKVCFYKLLTGALERDCGISPD----DVIVALVENSDADWSFGRGRAEFLTGDLV 129
YodA         65 MVIISMTSKARTESQKEKLYALLAERLEKKCEISPD----DLMVSITENGDADWSFGLGEAQFLNGKL  128
YrdN         65 LVVLSVTSKSRPEEKKQKFYRLLAERLESECGIAST----DLIVSIVENDNADWSFGLGEAQFLTGKL  128
                u u u u  -*-----  uu  u   u   u          u u u
```

FIGURE 1: Structure based alignment of the title enzymes of the 4-OT, MIF, and CHMI families along with *cis*-CaaD and MSAD family members. (A) Alignment of the N-terminal halves of MIF, CHMI, and the *cis*-CaaD and MSAD family members with 4-OT and the *β*-subunit of CaaD. (B) Alignment of the C-terminal halves of MIF, CHMI, and the *cis*-CaaD and MSAD family members with 4-OT and the α-subunit of CaaD. The sequences are those of 4-OT from *P. putida* mt-2 (GI: 4139813), MIF from *Mus musculus* (GI: 5542285), CHMI from *E. coli* C (GI: 1421039), CaaD1 corresponding to the α-subunit of CaaD from *P. pavonaceae* 170 (GI: 10637969), CaaD2 corresponding to the *β*-subunit of CaaD from *P. pavonaceae* 170 (GI: 10637970), *cis*-CaaD from coryneform bacterium strain FG41 (GI: 555299), Cg10062, a *cis*-CaaD homologue from *C. glutamicum* ATCC 13032 (GI: 19551312), OrfX and OrfY, *cis*-CaaD homologues from *M. smegmatis*, MSAD from *P. pavonaceae* 170 (GI: 10637971), and YodA (GI: 16079011) and YrdN (GI: 16079719), MSAD homologues from *B. subtilis*. The determined secondary structure elements of 4-OT are shown above its sequence (*29*). The structurally variable regions in 4-OT, MIF, and CHMI are shown in lower case letters. The absolutely conserved and catalytically important amino-terminal proline (Pro-1) and the putative catalytic arginine (corresponding to Arg-11 in 4-OT) are shown in bold-face letters and indicated by an asterisk. The putative catalytic arginine (corresponding to Arg-8 in CaaD1) and glutamate (corresponding to Glu-52 in CaaD1) residues specific to CaaD and members of the *cis*-CaaD family are boxed and shown in bold-face letters. The region of highest sequence identity among the members of the five different families within the tautomerase superfamily is underlined. The conserved hydrophobic sites are marked with "u". For clarity, the C-termini of CaaD1 and the *cis*-CaaD family have been truncated.

mutants showed one major peak corresponding to the expected molecular mass of a 149-amino acid species. This observation indicates that each of the mutants has undergone posttranslational processing to remove the initiating methionine, resulting in a protein with an N-terminal proline. In addition, the experimentally obtained molecular masses of the purified mutant proteins confirmed the presence of only the intended amino acid substitutions.

The activities of the *cis*-CaaD mutants were assayed using *cis*-**1** and *cis*-**2** as substrates. It was found that substitution of Pro-1, Arg-70, and Arg-73 by an alanine essentially abolished enzymatic activity. Under the conditions of the kinetic assays, no activity could be detected for these mutants. However, prolonged incubation with *cis*-**1** revealed a small amount of activity for the P1A and R70A mutants, while the R73A mutant had no detectable activity. The E114Q mutant, however, processes *cis*-**1** and *cis*-**2**, although at a reduced catalytic efficiency (Table 2). For *cis*-**1**, there is a 14-fold reduction in K_m and a 113-fold reduction in k_{cat}, which results in an 8-fold reduction in k_{cat}/K_m. For *cis*-**2**, there is a 15-fold reduction in K_m and a 84-fold reduction in k_{cat}, which results in a ~6-fold reduction in k_{cat}/K_m. The major effect of this mutation is seen in k_{cat} with the turnover of *cis*-**1** being affected more adversely. To determine whether these reduced enzymatic activities resulted from global changes in protein structure and not from specific effects of the side chains of the new amino acids, the mutants were analyzed by gel filtration chromatography and nondenaturing polyacrylamide gel electrophoresis. The modified enzymes were found to migrate comparably with the wild-type enzyme, which suggests that gross conformational changes are not likely present. Therefore, the proposed catalytic residues, Pro-1, Arg-70, Arg-73, and Glu-114, are indeed important for the *cis*-CaaD-catalyzed hydration reactions.

*Irreversible Inhibition of cis-CaaD by **8** and **9***. It has previously been shown that **8** and **9** (Scheme 2) irreversibly

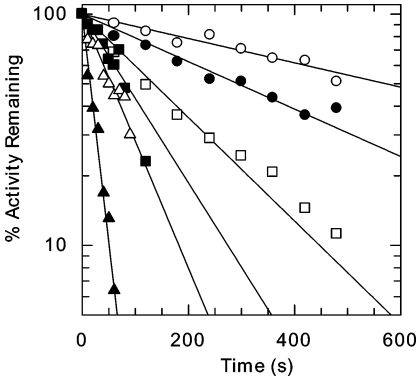


FIGURE 2: Time- and concentration-dependent irreversible inactivation of *cis*-CaaD by **9**. A logarithmic plot showing the percent of *cis*-CaaD activity remaining as a function of the incubation time with varying amounts of **9** (open circles, 2 μM; filled circles, 4 μM; open squares, 8 μM; filled squares, 10 μM; open triangles, 20 μM; filled triangles, 60 μM).

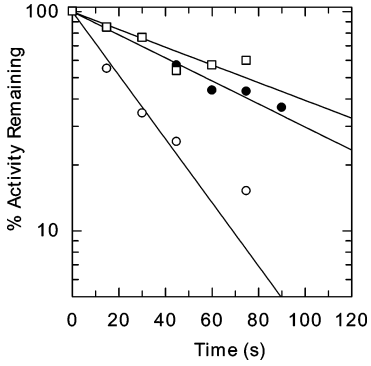


FIGURE 3: Protection of *cis*-CaaD against inactivation by **9** using substrate *cis*-**1**. The *cis*-CaaD was incubated with varying amounts of *cis*-**1** (open circles, 0 mM; filled circles, 1 mM; open squares, 2.5 mM) for 30 s before the addition of **9** (60 μM).

inhibit CaaD (*6*). The inhibition results from the covalent modification of β-Pro-1 by a species formed as a result of the enzyme-catalyzed hydration of the 3-halopropiolic acids. In view of the significantly slower rate of hydration of **6** by *cis*-CaaD, we decided to first examine whether incubation of *cis*-CaaD with either **8** or **9** led to irreversible inhibition before carrying out a more extensive kinetic characterization. After a 1-h incubation period (at 22 °C) with either **8** or **9**, *cis*-CaaD was irreversibly inactivated. Dialysis (24 h) and gel filtration did not result in recovery of enzyme activity. In the absence of **8** or **9**, dialysis and gel filtration have no effect on the activity of *cis*-CaaD. These observations indicate that a covalent bond has formed between *cis*-CaaD and the inhibitor or a species derived from the inhibitor.

*Kinetic Characterization of cis-CaaD Inactivation by **8** and **9***. Both **8** and **9** inactivated *cis*-CaaD in a time- and concentration-dependent manner. A logarithmic plot showing the percent of *cis*-CaaD activity remaining as a function of time using different concentrations of **9** (ranging from 1 to 100 μM) is shown in Figure 2. At all concentrations of **9** used, the decrease in activity was pseudo-first-order in enzymatic activity for at least three half-lives. A plot of the k_{obsd} values obtained from 12 experiments versus inhibitor concentration gave a straight line through the origin, indicative of nonsaturation kinetics (data not shown). Our failure to observe saturation kinetics could be due to a number of factors, but likely stems from the very rapid rates of inactivation at higher inhibitor concentrations. The rapid rates precluded the collection of sufficiently precise data to obtain k_{obsd} values for these concentrations.

To determine whether inactivation occurs at the active site, *cis*-CaaD was incubated with **9** in the presence of *cis*-**1**. The concentration of **9** was maintained at 60 μM (which rapidly inactivated *cis*-CaaD in <90 s), while the concentration of *cis*-**1** was increased from 0 to 2.5 mM. The results clearly show that *cis*-**1** protects the enzyme against inactivation by **9** (Figure 3). At all concentrations examined, the presence of *cis*-**1** significantly extends the lifetime of the enzyme. These observations provide evidence that **9** binds at the active site.

*ESI-MS Analysis of cis-CaaD and cis-CaaD Treated with **8** and **9***. To identify the species resulting in the covalent modification of *cis*-CaaD, the enzyme was incubated with both **8** and **9** (in separate reactions), and the inactivated proteins were isolated, and analyzed by ESI-MS. A control reaction containing only *cis*-CaaD was processed and analyzed similarly. Deconvolution of the ESI mass spectrum of *cis*-CaaD incubated for 16 h without inhibitor shows that the enzyme is highly pure and has an observed molecular mass of 16 622 Da (Figure 4A), which corresponds to the expected mass of the unmodified *cis*-CaaD. Mass spectral analysis of the samples of *cis*-CaaD incubated with **8** or **9** showed that each sample consists of one major component with an observed molecular mass of 16 708 Da (Figure 4, panels C and B, respectively). This mass corresponds to the expected molecular mass of *cis*-CaaD modified by a species derived from **8** or **9** with a mass of 86 Da.

Identification of the Modified Peptide by Mass Spectrometry. To identify the site of modification, the *cis*-CaaD samples treated with **8** or **9** and an unmodified *cis*-CaaD control sample were digested with endoproteinase Glu-C (protease V-8), and the resulting peptide mixtures were analyzed by MALDI-MS. At pH 8.0, in 100 mM (NH_4)-HCO_3 buffer, protease V-8 hydrolyzes peptides at the carboxylate side of glutamate and aspartate residues with a preference for the former (*23*, *24*). There are nine glutamate and 10 aspartate residues in *cis*-CaaD so that complete digestion should result in at least 10 fragments (assuming no hydrolysis at aspartate residues).

Mass spectral analysis of the peptide mixtures revealed that proteolytic cleavage occurred predominantly at Asp-26, Asp-61, Glu-71, Glu-94, Glu-106, Glu-114, Glu-125, Glu-136, and Glu-141 (Table 3). A comparison of the peaks for the *cis*-CaaD samples treated with **8** or **9** to those observed for the unmodified *cis*-CaaD sample revealed a single modification by a species having a mass of 42 on the fragment Pro-1 to Asp-26 (Figure 5 A−C).[3] Analysis of the remaining peaks showed no modification of other fragments (Table 3). These data, along with the ESI-MS analysis of the intact *cis*-CaaD treated with **8** and **9**, indicate that a single site on the enzyme has been modified and that the site of modification is localized to the 26-residue amino-terminal

[3] ESI-MS analysis of *cis*-CaaD treated with **8** or **9** revealed the addition of a species having a molecular mass of 86 Da. However, the peptide masses observed with MALDI-MS (Table 3) only show an increase in mass of 42 Da. This apparent discrepancy is most likely due to the matrix-induced loss of the CO_2 group from the label (Scheme 5). A similar observation has been made previously (*6*).

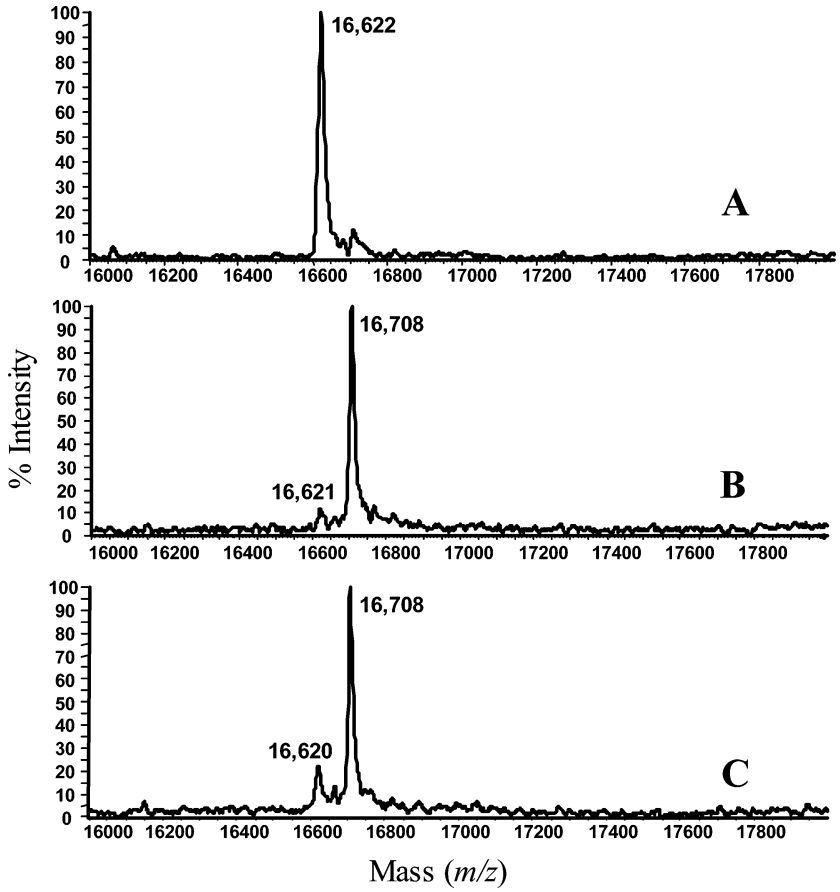


FIGURE 4: ESI-MS spectra of (A) unmodified *cis*-CaaD, (B) *cis*-CaaD treated with **9**, and (C) *cis*-CaaD treated with **8**. The deconvoluted spectra display a singly charged state. The major component in the spectra shown in panels B and C is the modified *cis*-CaaD (*m/z* 16,708 Da) and the minor component is the unmodified protein (*m/z* 16 621 and 16 620 Da, respectively).

Table 3: Assignments of Peptides Produced from Protease V-8 Digestion of Unmodified and **8**- and **9**-Treated *cis*-CaaD

peptide fragment	calculated mass[a]	observed mass unmodified *cis*-CaaD	observed mass *cis*-CaaD treated with **8**	observed mass *cis*-CaaD treated with **9**
^{1}P−^{26}D	2860.50	2860.49	2902.53	2902.53
^{62}T−^{71}E	1208.65	not detected	not detected	1208.64
^{72}G−^{94}E	2398.27	2398.27	2398.27	2398.27
^{95}I−^{106}E	1562.81	1562.81	1562.81	1562.81
^{115}Y−^{125}E	1300.64	1300.64	1300.64	1300.64
^{137}R−^{141}E	653.30	653.28	653.28	653.28
^{142}T−^{149}T	891.45	not detected	not detected	891.46

[a] The monoisotopic singly charged masses are predicted from analysis of the translated amino acid sequence of the *cis-caaD* gene (corresponding to the unmodified *cis*-CaaD).

fragment (Pro-1 to Asp-26) that harbors the potential catalytic proline (Pro-1).

There are, however, several other potential targets for modification within the 26-residue amino-terminal fragment (PVYMVYVSQDRLTPSAKHAVAKAITD). To determine the actual site of the single modification, selected peaks observed in the V8-digested control sample and in the V8-digested sample treated with **8** were subjected to MALDI−post-source decay (PSD) fragmentation analysis (*25*). The expected b-ion masses for the unmodified peptide and the peptide labeled at the N-terminal proline were calculated (Figure 6A) and compared to the data (Figure 6B,C). The PSD spectrum of the precursor ion (*m/z* 2860.5, $(M + H)^+$) of the unlabeled peptide displays the characteristic b_2, b_3, b_4, and b_5 fragment ions (Figure 6B). PSD analysis of the precursor ion (*m/z* 2902.5, $(M + H)^+$) of the **8**-treated peptide revealed an increase in the masses by 42 Da of the b-ions starting at b_1 (Figure 6C). Because the b_1 ion is shifted by a mass of +42 Da, the modification has occurred on Pro-1. The analysis also excludes modification of any other residue within the 26-residue amino-terminal fragment.

DISCUSSION

Halogenated compounds are utilized by several microorganisms as sources of carbon and energy. One of the steps encountered by the microorganism in the catabolism of these compounds is the cleavage of the carbon−halogen bond. In many microorganisms, specific dehalogenases have evolved to catalyze the cleavage of these bonds. In other microorganisms, an enzyme-catalyzed reaction generates an unstable product, which decomposes to release the halide. These dehalogenases and so-called "accidental dehalogenases" display a variety of catalytic strategies (*10*, *11*).

One strategy that has not been observed for microbial dehalogenases is the hydrolytic dehalogenation of vinylic halides via noncovalent catalysis (*6*, *22*). In this regard, the mechanism used by CaaD (and now *cis*-CaaD) is unique. Sequence analysis along with site-directed mutagenesis linked CaaD to the 4-OT family (*5*), thereby arguing against covalent catalysis (i.e., formation of an alkyl enzyme intermediate) and implicating another catalytic strategy. A recent study provided experimental evidence that the hydrolytic dehalogenation by CaaD is mechanistically a hydra-

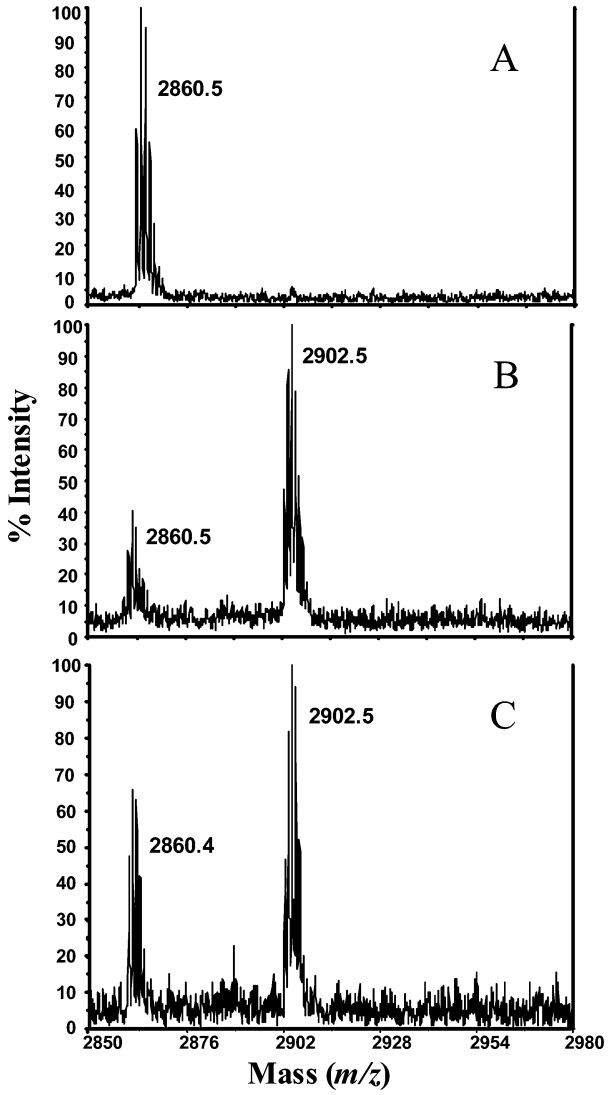


FIGURE 5: MALDI-MS spectra displaying the 26-residue amino-terminal peptide of protease V8 digested *cis*-CaaD samples. (A) Unmodified *cis*-CaaD, (B) *cis*-CaaD treated with **9**, and (C) *cis*-CaaD treated with **8**. Note that the spectra shown in panels B and C, but not panel A, display a major species with an observed mass of 2902.5 that corresponds to the mass of the 26-residue amino-terminal peptide that has been covalently modified by a species having a mass of 42 Da.

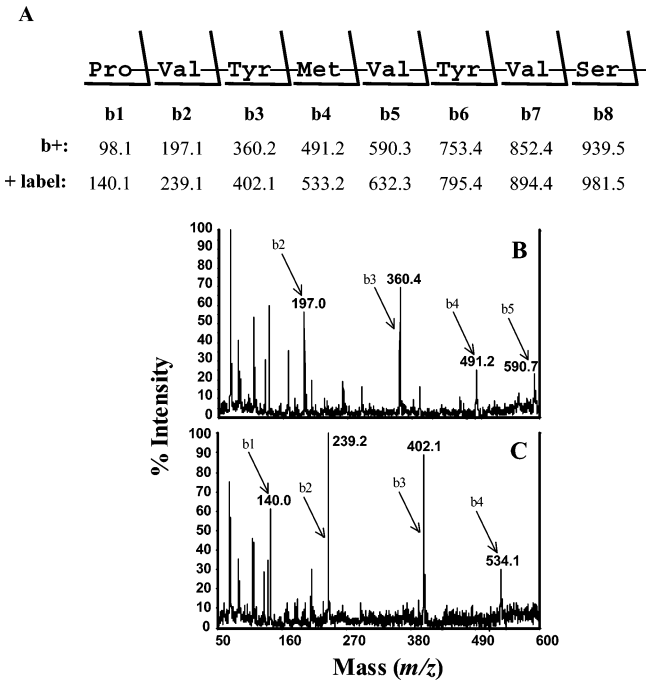
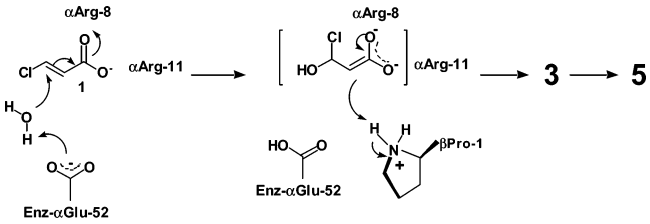


FIGURE 6: Post-source decay analysis of the 26-residue amino-terminal peptide Pro-1-Asp-26. (A) A portion of the peptide sequence with the calculated monoisotopic masses for the PSD fragment b ions of the unlabeled peptide and the peptide modified by an adduct having a mass of 42 Da. (B) PSD spectrum of the precursor ion *m*/*z* 2860.5 obtained with peptide Pro-1-Asp-26 of the unlabeled enzyme. (C) PSD spectrum of the precursor ion *m*/*z* 2902.5 obtained with the same peptide derived from the V8 digestion of *cis*-CaaD treated with **8**.

Scheme 3



tion reaction (*6*). It was found that CaaD hydrates **6** to afford **7** (with a k_{cat} of ~0.7 s^{-1}), and it converts **8** and **9** into potent irreversible inhibitors, presumably through an acyl halide intermediate.

On the basis of mechanistic (*6*), crystallographic (*22*), and mutagenesis studies (*5*, *22*), a mechanism involving the general base-catalyzed attack of water has emerged. Sequence analysis and mutagenesis first identified Pro-1 (in the β-subunit) and Arg-11 (in the α-subunit) as critical mechanistic residues. Replacement of either residue by alanine resulted in greatly reduced activity or no detectable activity (*5*, *6*). Subsequently, crystal structures of the native CaaD and CaaD inactivated by **8** showed that two additional residues, αGlu-52 and αArg-8, were correctly positioned to function in catalysis (*22*). Mutagenesis of these residues confirmed their importance in the mechanism: there was no detectable activity for the αE52Q mutant and a strongly reduced activity for the αR8A mutant (*22*). The positions of αGlu-52 and βPro-1 in the crystal structures (*22*) coupled with the preliminary determination of a pK_a value for βPro-1 of ~9.2 (*26*) implicated αGlu-52 as the general base catalyst that activates a water molecule for attack at C-3 (of *trans*-**1**) and βPro-1 as the general acid catalyst that provides a proton at C-2 (Scheme 3). The two arginines likely interact with the carboxylate group to align the substrate and draw electron density away from C-3. Such an interaction would make C-3 partially positively charged and facilitate the Michael addition of water to generate the halohydrin intermediate (e.g., **3**). The enzyme-catalyzed or the nonenzymatic collapse of **3** produces **5**.

The discovery of a *cis*-CaaD and its preliminary characterization raised significant mechanistic and evolutionary questions. The fundamental mechanistic questions concerned whether the enzyme catalyzes a hydration reaction resulting in the dehalogenation of the *cis*-3-haloacrylates such that it functions as an "accidental dehalogenase", and whether the mechanism also involves a general base-catalyzed attack of water. The major evolutionary question concerned how two bacterial 3-chloroacrylic acid dehalogenases evolved with such different quaternary structures and isomer specificities.

To address these questions, it was necessary to clone the *cis-caaD* gene and express the enzyme in large quantities, and establish the basic aspects of the reaction and the mechanism.

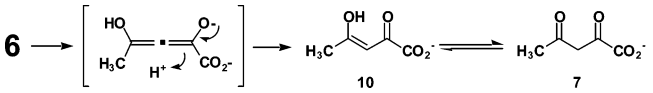
Characterization of *cis*-CaaD revealed primary and quaternary structure differences between CaaD and *cis*-CaaD as well as rate differences in the CaaD- and *cis*-CaaD-catalyzed reactions. CaaD functions as a heterohexamer, composed of relatively small α-subunits (75 amino acids) and *β*-subunits (70 amino acids). In contrast, *cis*-CaaD functions as a homotrimer, where each monomer is composed of 149 amino acids. Moreover, CaaD is less efficient in the turnover of substrate when compared with *cis*-CaaD. The k_{cat}/K_m values determined for the CaaD-catalyzed hydration of *trans*-**1** and *trans*-**2** are 35 and 42% (respectively) of those determined for the *cis*-CaaD-catalyzed hydration of *cis*-**1** and *cis*-**2** (*6*). Sequence analysis and site-directed mutagenesis produced a working hypothesis for the mechanism of *cis*-CaaD, which parallels that of CaaD. Accordingly, Glu-114 might function as the water-activating base, while Pro-1 places a proton at C-2. The two arginines (Arg-70 and Arg-73) are likely to interact with the carboxylate group of the substrate, which would assist the presumed Michael addition of water.

It is interesting to note that the E114Q mutant of *cis*-CaaD retains a significant amount of activity (∼8-fold decrease in k_{cat}/K_m value when using *cis*-**1**). This observation stands in contrast to CaaD, where the αE52Q mutant shows no residual activity. The presence of activity for the E114Q mutant of *cis*-CaaD has at least two possible explanations. The first explanation is that the remaining interactions of Pro-1 and the two arginines create a favorable environment for the addition of water across the double bond by activating C-3 of the substrate. The two arginines could polarize the α,*β*-unsaturated acid to generate an enediolate intermediate, resulting in a positive charge at C-3 (or partial positive charge). Assuming that the active site of *cis*-CaaD is hydrophilic (as would be expected for a hydration reaction), a nearby water molecule could add to the now activated 3-haloacrylate. The presence of Pro-1 could facilitate the ketonization of the enediolate intermediate and "push" the equilibrium toward product formation. In a second explanation, the enzyme may use other residues along with Glu-114 to activate the water molecule. The observation that the αE52Q mutant of CaaD shows no residual activity argues in favor of the second explanation as one would expect some activation of *trans*-**1** by the presence of *β*Pro-1, αArg-8, and αArg-11. The structural basis for the residual activity in the E114Q mutant of *cis*-CaaD is under investigation.

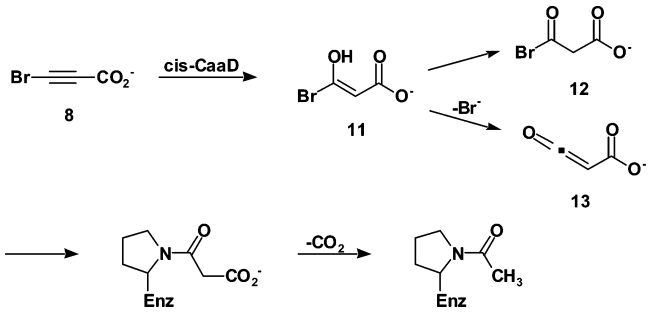
The E114Q mutant of *cis*-CaaD also displays a lower K_m value for the substrate. A structural basis for this observation is not yet known but one possible explanation may stem from the fact that this mutation removes a negative charge. It is possible that removal of the negative charge would decrease the hydrophilic nature of the active site and increase the strength of the electrostatic interaction between the substrate's carboxylate group and the two arginines.

This study also shows that *cis*-CaaD functions as a hydratase when processing the three acetylene compounds (**6**, **8**, and **9**). The *cis*-CaaD-catalyzed conversion of **6** to **7** likely involves the initial Michael addition of water to the triple bond to form an allenic species. Rearrangement of this species produces **10**, which readily ketonizes to form **7** (Scheme 4). One or both the arginine residues (Arg-70 or Arg-73) could make the addition of water across the triple bond more favorable by interacting with the 2-carbonyl group. Such an interaction would polarize the carbonyl group. Interestingly, while *cis*-CaaD processes the 3-haloacrylates faster than CaaD, it converts **6** to **7** at a much slower rate (582-fold as assessed by the k_{cat}/K_m values) than does CaaD. The slower rate likely reflects the different topologies of the active sites. There are significant differences in the K_m value of *cis*-CaaD for **6** (∼5.6-fold) when compared with the K_m value measured for **6** and CaaD. There are even more significant differences in the K_m value of *cis*-CaaD for **6** (∼21-fold lower) when compared to the K_m values measured for the 3-haloacrylates and *cis*-CaaD. Taken together, these observations suggest that **6** is not optimally bound in the active site of *cis*-CaaD to undergo a hydration reaction.

Scheme 4



Scheme 5



Mass spectral analysis indicates that both 3-halopropiolates result in the modification of a single residue of *cis*-CaaD, identified as Pro-1, with an adduct having a mass of 86 Da. Crystallographic analysis shows that indubation of CaaD with **8** results in the attachment of a malonyl species to Pro-1 (of the *β*-subunit) (*22*). This observation supported an earlier mass spectral analysis of CaaD inactivated by **8** (or **9**), which showed the attachment of a species to the *β*-subunit of CaaD, having a mass of 85 Da (*6*). This species is likely derived from the hydration of **8** followed by rearrangement to an acyl halide (Scheme 5). In such a mechanism, hydration initially generates the unstable species designated **11** in Scheme 5. While **11** can rearrange to an acyl halide (**12**) or a ketene (**13**), the acyl halide route is presumed to be the predominant one as proton transfer is likely a faster reaction than bromide elimination. The presence of an acyl halide in the active site will cause rapid alkylation of nearby nucleophiles. Presumably, the same mechanism operates upon incubation of *cis*-CaaD with **8** or **9**. The initial hydration of **8**, by the combined efforts of Glu-114 and Pro-1, will leave Pro-1 without a proton and render it nucleophilic. Thus, Pro-1 becomes an ideal candidate for alkylation. Interestingly, the enzyme processes **9** at a slower rate than does CaaD (*6*), which may again be a reflection of the different topologies of the two active sites. These studies also confirm that Pro-1 is an active site residue in *cis*-CaaD, thereby providing further support for its categorization in the tautomerase superfamily.

The similarities between CaaD and *cis*-CaaD are somewhat reminiscent of those found between two bacterial isomerases, 4-OT and CHMI. 4-OT is found in a plasmid-encoded pathway in *P. putida* mt-2 and is responsible for the degradation of aromatic hydrocarbons such as benzene, toluene, and xylenes (*27*). CHMI is found in *E. coli* C as part of a pathway responsible for the degradation of phenylalanine and tyrosine (*28*). The two enzymes are structurally homologous (where a 4-OT dimer is equivalent to a CHMI monomer) and carry out the same mechanism (*8*, *29*), but the substrates differ by a carboxymethyl group at the C-5 position: the 4-OT substrate has a hydrogen at C-5, while the CHMI substrate has a carboxymethyl group. However, 4-OT and CHMI belong to separate families in the tautomerase superfamily where the CHMI family ancestor has probably resulted from the duplication of a 4-OT-like sequence. Despite their likely independent origins, 4-OT and CHMI use similar catalytic strategies to process differently substituted molecules (*7*, *8*).

Along these lines, sequence analysis of *cis*-CaaD indicates that it cannot be categorized into one of the four known families of the tautomerase superfamily, a group of structurally homologous proteins characterized by a conserved $\beta-\alpha-\beta$ structural motif and a catalytically important amino-terminal proline (*7*). 4-OT and CaaD are the two best-characterized members of the 4-OT family of enzymes with assigned functions. The other three families are represented by CHMI, MIF, a mammalian cytokine with a phenylpyruvate tautomerase activity (*30*), and MSAD, which is responsible for the decarboxylation of **5** in the 1,3-dichloropropene catabolic pathway in *P. pavonaceae* 170 (*9*). In the 4-OT family, the monomer encodes a single $\beta-\alpha-\beta$ motif while in the CHMI and MIF families, the monomers are nearly twice as long and consist of two $\beta-\alpha-\beta$ motifs, which resemble the 4-OT dimer (*8*). Although a structure is not yet available for MSAD, sequence analysis suggests that its monomer (129 amino acids) will also consist of two $\beta-\alpha-\beta$ motifs (*9*).

The *cis*-CaaD family represents a new family in the tautomerase superfamily, which probably resulted from an independent duplication of a 4-OT-like sequence, as has been proposed for the MIF, MSAD, and CHMI families (*7*−*9*). There are currently three additional members of this family: the homologue from *C. glutamicum* and two homologues from *Mycobacterium smegmatis*. The discovery of the *cis*-CaaD along with the identification of a new family in the tautomerase superfamily demonstrates the remarkable versatility of the $\beta-\alpha-\beta$ motif for the creation of new enzymatic activities.

While this investigation has shown that CaaD and *cis*-CaaD use similar catalytic strategies to process the different isomers of 3-haloacrylates, they have apparently evolved as a result of divergent evolution and show modest rate differences in substrate turnover. This last observation raises an immediate question of why *cis*-CaaD is apparently more efficient than CaaD. This observation and the differential behavior of the αE52Q and E114Q mutants suggest subtle differences in the mechanisms. With the enzyme now available in large quantities and the fundamental characterization in hand, the stage is set for a structural analysis of *cis*-CaaD to determine the basis for the differences in catalytic efficiencies and substrate specificity between the two isomer-specific 3-chloroacrylic acid dehalogenases.

ACKNOWLEDGMENT

We thank Professor Dick B. Janssen (Department of Biochemistry, University of Groningen, The Netherlands) for the kind gift of coryneform bacterium strain FG41 and for his stimulating discussions. We gratefully acknowledge Dr. Rick Rink (BioMade, Groningen, The Netherlands) for his assistance in the design and use of the degenerate primers. We are also grateful to Dr. Rene M. de Jong and Professor Bauke W. Dijkstra (Laboratory of Biophysical Chemistry, University of Groningen, The Netherlands) for communication of their crystallographic results with CaaD prior to publication. Finally, we thank Steve D. Sorey (Department of Chemistry, University of Texas at Austin) for his expert assistance in acquiring the NMR spectra. Preliminary sequence data were obtained from the Institute for Genomic Research Web site (http://www.tigr.org). Mass spectrometry was performed by the Analytical Instrumentation Facility Core (College of Pharmacy, The University of Texas at Austin) supported by the NIEHS Center Grant ES07784.

REFERENCES

1. Poelarends, G. J., Wilkens, M., Larkin, M. J., van Elsas, J. D., and Janssen, D. B. (1998) Degradation of 1,3-dichloropropene by *Pseudomonas pavonaceae* 170. *Appl. Environ. Microbiol. 64*, 2931−2936.
2. Poelarends, G. J., van Hylckama Vlieg, J. E. T., Bosma, T., and Janssen, D. B. (2002) The haloalkane dehalogenase genes *dhlA* and *dhaA* are globally distributed and highly conserved, in *Biotechnology for the Environment: Strategy and Fundamentals* (Agathos, S. N., and Reineke, W., Eds.) pp 59−66, Kluwer Academic Publishers, Dordrecht, The Netherlands.
3. van Hylckama Vlieg, J. E. T., and Janssen, D. B. (1992) Bacterial degradation of 3-chloroacrylic acid and the characterization of *cis*- and *trans*-specific dehalogenases. *Biodegradation 2*, 139−150.
4. Hartmans, S., Jansen, M. W., van der Werf, M. J., and de Bont, J. A. M. (1991) Bacterial metabolism of 3-chloroacrylic acid. *J. Gen. Microbiol. 137*, 2025−2032.
5. Poelarends, G. J., Saunier, R., and Janssen, D. B. (2001) *trans*-3-Chloroacrylic acid dehalogenase from *Pseudomonas pavonaceae* 170 shares structural and mechanistic similarities with 4-oxalocrotonate tautomerase. *J. Bacteriol. 183*, 4269−4277.
6. Wang, S. C., Person, M. D., Johnson, W. H., Jr., and Whitman, C. P. (2003) Reactions of *trans*-3-chloroacrylic acid dehalogenase with acetylene substrates: consequences of and evidence for a hydration reaction. *Biochemistry 42*, 8762−8773.
7. Murzin, A. G. (1996) Structural classification of proteins: new superfamilies. *Curr. Opin. Struct. Biol. 6*, 386−394.
8. Whitman, C. P. (2002) The 4-oxalocrotonate tautomerase family of enzymes: how nature makes new enzymes using a β-α-β structural motif. *Arch. Biochem. Biophys. 402*, 1−13.
9. Poelarends, G. J., Johnson Jr., W. H., Murzin, A. G., and Whitman, C. P. (2003) Mechanistic characterization of a bacterial malonate semialdehyde decarboxylase: identification of a new activity in the tautomerase superfamily. *J. Biol. Chem. 278*, 48674−48683.
10. Janssen, D. B., Oppentocht, J. E., and Poelarends, G. J. (2003) Bacterial growth on halogenated aliphatic hydrocarbons: genetics and biochemistry, in *Dehalogenation: Microbial Processes and Environmental Applications* (Häggblom, M. M., and Bossert, I. D., Eds.) pp 207−226, Kluwer Academic Publishers, Dordrecht, The Netherlands.
11. Copley, S. D. (2003) Aromatic dehalogenases: insights into structures, mechanisms, and evolutionary origins in *Dehalogenation: Microbial Processes and Environmental Applications* (Häggblom, M. M., and Bossert, I. D., Eds.) pp 227−259, Kluwer Academic Publishers, Dordrecht, The Netherlands.

12. Johnson, Jr., W. H., Czerwinski, R. M., Fitzgerald, M. C., and Whitman, C. P. (1997) Inactivation of 4-oxalocrotonate tautomerase by 2-oxo-3-pentynoate. *Biochemistry 36*, 15724−15732.
13. Strauss, F., Kollek, L., and Heyn, W. (1930) Uber den ersatz positiven wasserstoffs durch halogen. *Chem. Ber. 63*, 1868−1899.
14. Andersson, K. (1972) Additions to propiolic and halogen substituted propiolic acids. *Chem. Scr. 2*, 117−120.
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
16. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI−BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389−3402.
17. 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.
18. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*, 51−59.
19. Guthrie, J. P. (1972) Acetopyruvic acid. Rate and equilibrium constants for hydration and enolization. *J. Am. Chem. Soc. 94*, 7020−7024.
20. Taylor, A. B., Czerwinski, R. M., Johnson, W. H., Whitman, C. P., and Hackert M. L. (1998) Crystal structure of 4-oxalocrotonate tautomerase inactivated by 2-oxo-3-pentynoate at 2.4 Å resolution: analysis and implications for the mechanism of inactivation and catalysis. *Biochemistry 37*, 14692−14700.
21. Chen L. H., Kenyon G. L., Curtin F., Harayama S., Bembenek M. E., Hajipour G., and Whitman C. P. (1992) 4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per monomer. *J. Biol. Chem. 267*, 17716−17721.
22. de Jong, R. M., and Dijkstra, B. W. (2003) Structure and mechanism of bacterial dehalogenases: different ways to cleave a carbon-halogen bond. *Curr. Opin. Struct. Biol. 13*, 722−730.
23. Sorensen, S. B., Sørenson, T. L., and Breddam, K. (1991) Fragmentation of proteins by *S. aureus* strain V8 protease. Ammonium bicarbonate strongly inhibits the enzyme but does not improve the selectivity for glutamic acid. *FEBS 294*, 195−197.
24. Houmard, J., and Drapeau, G. R. (1972) Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds. *Proc. Natl. Acad. Sci. U.S.A. 69*, 3506−3509.
25. Person, M. D., Monks, T. J., and Lau, S. S. (2003) An integrated approach to identifying chemically induced posttranslational modifications using comparative MALDI-MS and targeted HPLC-ESI-MS/MS. *Chem. Res. Toxicol. 16*, 598−608.
26. Azurmendi, H. F., Wang, S. C., Whitman, C. P., and Mildvan, A. S. (2003) New role for the amino-terminal proline in the tautomerase superfamily: *trans*-3-chloroacrylic acid dehalogenase. *Biochemistry 42*, 8619, Abstr. No. 113.
27. Harayama, S., Rekik, M., Ngai, K.-L., and Ornston, L. N. (1989) Physically associated enzymes produce and metabolize 2-hydroxy-2,4-dienoate, a chemically unstable intermediate formed in catechol metabolism via meta cleavage in *Pseudomonas putida*. *J. Bacteriol. 171*, 6251−6258.
28. Sparnins, V. L., Chapman, P. J., and Dagley, S. (1974) Bacterial degradation of 4-hydroxyphenylacetic acid and homoprotocatechuic acid. *J. Bacteriol. 120*, 159−167.
29. Subramanya, H. S., Roper, D. I., Dauter, Z., Dodson, E. J., Davies, G. J., Wilson, K. S., and Wigley, D. B. (1996) Enzymatic ketonization of 2-hydroxymuconate: specificity and mechanism investigated by the crystal structures of two isomerases. *Biochemistry 35*, 792−802.
30. Taylor, A. B., Johnson, W. H., Jr., Czerwinski, R. M., Li, H.-S., Hackert, M. L., and Whitman, C. P. (1999) Crystal structure of macrophage migration inhibitory factor complexed with (*E*)-2-fluoro-p-hydroxycinnamate at 1.8 Å resolution: implications for enzymatic catalysis and inhibition. *Biochemistry 38*, 7444−7452.

BI0355948