Bioorganic Chemistry 39 (2011) 1–9

Contents lists available at ScienceDirect

Bioorganic Chemistry

journal homepage: www.elsevier.com/locate/bioorg

Crystal structures of native and inactivated *cis*-3-chloroacrylic acid dehalogenase: Implications for the catalytic and inactivation mechanisms ☆

Youzhong Guo [a], Hector Serrano [a], William H. Johnson Jr. [a], Stephen Ernst [b], Marvin L. Hackert [b,*], Christian P. Whitman [a,*]

[a] Division of Medicinal Chemistry, College of Pharmacy, University of Texas, Austin, TX 78712, United States
[b] Department of Chemistry and Biochemistry, College of Pharmacy, University of Texas, Austin, TX 78712, United States

ARTICLE INFO

Article history:
Received 19 July 2010
Available online 20 October 2010

Keywords:
Tautomerase superfamily
Catalytic proline
Dehalogenase
β–α–β Motif
Epoxide

ABSTRACT

The isomeric mixture of *cis*- and *trans*-1,3-dichloropropene constitutes the active component of a widely used nematocide known as Telone II®. The mixture is processed by various soil bacteria to acetaldehyde through the 1,3-dichloropropene catabolic pathway. The pathway relies on an isomer-specific hydrolytic dehalogenation reaction catalyzed by *cis*- or *trans*-3-chloroacrylic acid dehalogenase, known respectively as *cis*-CaaD and CaaD. Previous sequence analysis and crystallographic studies of the native and covalently modified enzymes identified Pro-1, His-28, Arg-70, Arg-73, Tyr-103, and Glu-114 as key binding and catalytic residues in *cis*-CaaD. Mutagenesis of these residues confirmed their importance to the dehalogenation reaction. Crystal structures of the native enzyme (2.01 Å resolution) and the enzyme covalently modified at the Pro-1 nitrogen by 2-hydroxypropanoate (1.65 Å resolution) are reported here. Both structures are at a resolution higher than previously reported (2.75 Å and 2.1 Å resolution, respectively). The conformation of the covalent adduct is strikingly different from that previously reported due to its interaction with a 7-residue loop (Thr-32 to Leu-38). The participation of another active site residue, Arg-117, in catalysis and inactivation was also examined. The implications of the combined findings for the mechanisms of catalysis and inactivation are discussed.

1. Introduction

The pollution of the environment by synthetic halogenated organic compounds is the bane of the industrial age [1]. Many compounds, from commercial sources and agricultural applications, are hazardous due to their toxicity and persistence [2]. The repeated exposure of soil bacteria to these compounds may have precipitated the evolution of catabolic pathways for their degradation [3–5]. These pathways are rich in enzyme chemistry, structural diversity, evolutionary puzzles, and have the potential to rid the environment of hazardous compounds [5,6].

Telone II®, a widely used nematocide composed of *cis*- and *trans*-1,3-dichloropropene (**1**, Scheme 1), is degraded by one such pathway [7]. The compounds are converted to acetaldehyde, HCl, and CO_2 in five enzyme-catalyzed steps (Scheme 1). Acetaldehyde is presumably channeled into the Krebs Cycle. In the course of this bacterial catabolic pathway, *cis*- and *trans*-3-chloroacrylic acid (**2** and **3**) are generated and processed to malonate semialdehyde (**4**) by separate dehalogenases, known respectively as *cis*- and *trans*-3-chloroacrylic acid dehalogenase (designated *cis*-CaaD and CaaD) [7–10].

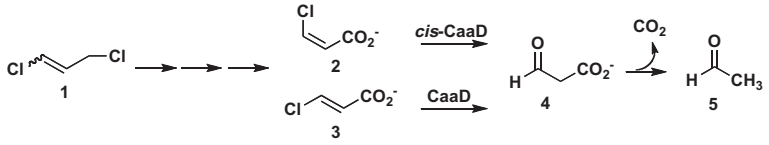
The mechanisms and structures of both enzymes have been extensively studied. The two enzymes represent separate families in the tautomerase superfamily, whose members are characterized by a β–α–β building block and a catalytic amino-terminal proline [11–14]. CaaD functions as a heterohexamer made up of three α- and three β-chains, where the catalytic unit is the α,β-heterodimer [15]. In contrast, *cis*-CaaD is a homotrimer where each polypeptide chain is a fusion product of the α- and β-like chains of CaaD [16]. The enzymes share 8 common residues, Pro-1 (βPro-1 in CaaD), a pair of arginines (αArg-8/αArg-11 in CaaD and Arg-70/Arg-73 in *cis*-CaaD), and a glutamate (αGlu-52 in CaaD and Glu-114 in

Abbreviations: Ap, ampicillin; BICINE, N,N-Bis(2-hydroxyethyl)glycine; CaaD, *trans*-3-chloroacrylic acid dehalogenase; *cis*-CaaD, *cis*-3-chloroacrylic acid dehalogenase; ESI-MS, electrospray ionization mass spectrometry; LB, Luria-Bertani; 4-OT, 4-oxalocrotonate tautomerase; PEG, polyethylene glycol; PDB, Protein Data Bank; rmsd, root-mean-square deviation; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

☆ This research was supported by the National Institutes of Health Grant GM-65324 and the Robert A. Welch Foundation (F-1334 to CPW and F-1219 to MLH). The atomic coordinates and structure factors of the native and inactivated cis-CaaD have been deposited in the Protein Data Bank as entries 3MF8 and 3MF7, respectively.

* Corresponding authors. Fax: +1 512 232 2606 (C.P. Whitman), fax: +1 512 471 8696 (M.L. Hackert).
E-mail addresses: m.hackert@mail.utexas.edu (M.L. Hackert), whitman@mail.utexas.edu (C.P. Whitman).

0045-2068/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bioorg.2010.10.001

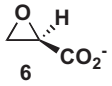


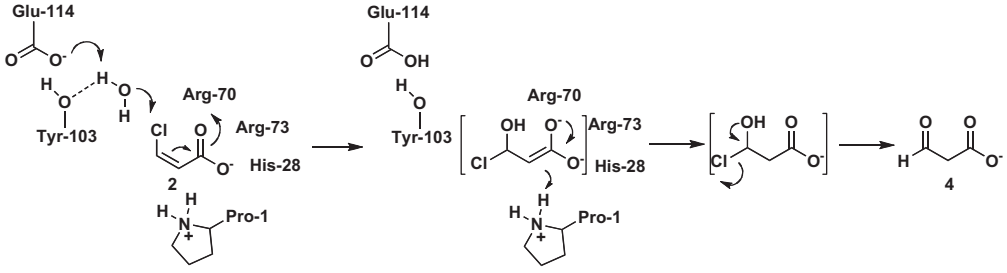
Scheme 1.

cis-CaaD) [7–9,15,16]. His-28 and Tyr-103 assist in *cis*-CaaD catalysis, but have no known counterparts in CaaD [16]. In one proposed mechanism (Scheme 2), the glutamate residue activates a water molecule for nucleophilic attack at C-3 (assisted by Tyr-103 in *cis*-CaaD), the arginine pair binds and polarizes the substrate by interactions with the C-1 carboxylate group (assisted by His-28 in *cis*-CaaD), and Pro-1 completes the addition of water by providing a proton at C-2 [14–16].

The initial studies of CaaD and *cis*-CaaD suggested largely parallel mechanisms [9,10]. However, as more details about the *cis*-CaaD mechanism emerge, it has become apparent that this is not the case. For example, two additional residues (Tyr-103 and His-28) are involved in *cis*-CaaD catalysis [16]. It is also not entirely clear in both enzymes how the catalytic residues interact with the substrate (and each other) as their roles have been largely defined by their positions in crystal structures. For these reasons, we continue to carry out kinetic studies of site-directed mutants coupled with additional crystal structures to better understand the catalytic details.

We report here high resolution crystal structures of both native (2.01 Å resolution) and *cis*-CaaD inactivated by the ring-opened product of (*R*)-oxirane-2-carboxylate (**6**) (1.65 Å resolution). The native structure is much like that of the previously published one with two differences: water molecules are observed in the active site and the C-terminus (residues 118–149) is unstructured [16]. In both structures, a 7-residue loop (Thr-32-Leu-38) is in a conformation defined as the open state. In contrast, the inactivated *cis*-CaaD structure shows a notable difference in the active site because the 2-hydroxypropanoate adduct is in a different conformation. The adduct interacts with residues of the 7-residue loop and places the loop in a different conformation defined as the closed state. Finally, we determined that the R117A mutant has diminished activity but is still inactivated by the (*R*)-oxirane-2-carboxylate. Hence, this active site residue may place a role in catalysis, but not a significant one. The combined findings further delineate the catalytic mechanism of *cis*-CaaD as well as its inactivation by epoxide **6**.





Scheme 2.

2. Materials and methods

2.1. Materials

The sources for the components of Luria–Bertani (LB) media as well as the enzymes and reagents used in the molecular biology procedures are reported elsewhere [9,17]. The Amicon concentrator and the YM10 ultrafiltration membranes were obtained from Millipore Corp. (Bedford, MA). Pre-packed PD-10 Sephadex G-25 columns were purchased from Biosciences AB (Uppsala, Sweden). The QuikChange II Site-Directed Mutagenesis Kit was acquired from Stratagene (La Jolla, CA). Oligonucleotides for DNA amplification and sequencing were synthesized by Genosys (The Woodlands, TX). The synthesis of (*R*)-oxirane-2-carboxylate (**6**) is reported elsewhere [18].

2.2. Bacterial strains, plasmids, and growth conditions

E. coli strain BL21-Gold(DE3) was obtained from Stratagene. The pET3b expression vector (Stratagene) was used for the expression of the R117A mutant gene. The *E. coli* strains were grown at 30 °C in LB medium, supplemented with ampicillin (100 μg/mL).

2.3. General methods

General procedures for cloning and DNA manipulation were performed as described elsewhere [19]. The polymerase chain reaction 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 at The University of Texas at Austin. *cis*-CaaD and the R117A mutant were purified to homogeneity, as assessed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), according to a published procedure [10]. Protein was analyzed by SDS–PAGE on gels containing 15% polyacrylamide [20]. The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the method of Waddell [21]. Absorbance data were obtained on a Hewlett Packard 8452A Diode Array spectrophotometer.

2.4. Crystallization and structure determination of native cis-CaaD

Crystals of native *cis*-CaaD were obtained at room temperature from 6-μL hanging drops consisting of equal amounts of precipitant

solution [0.8 M$(NH_4)_2SO_4$ in 0.1 M BICINE buffer at pH 9.0] and protein solution (14 mg/mL *cis*-CaaD in 10 mM Tris-SO_4 buffer at pH 8.0). Crystals grew to about 1 × 0.25 × 0.25 mm within one week. Diffraction data to 2.0 Å resolution (360 images at 0.5° each) were collected using a Rigaku RU200H rotating anode X-ray source with Cu Kα radiation and equipped with a R-AXIS IV++ image plate detector with a crystal-to-detector distance of 150 mm. The data were integrated and scaled using the HKL-2000 program package [22]. The crystal belongs to the space group $P6_3$ with cell parameters a = 58.99 and c = 58.30 Å. The asymmetric unit contains one monomer with 149 amino acid residues corresponding to a calculated Matthews' coefficient of 1.79 $Å^3$/Da and a solvent content of 31% [23]. Residues 118–149 were not resolved in the electron density map.

Molecular replacement solutions were obtained with the server CASPR available online [24]. A search model was constructed from the atomic coordinates of *cis*-CaaD (PDB code: 2FLT). Rotation and translation functions were calculated using data between 15- and 3-Å resolution, yielding the position and orientation of the one monomer in the asymmetric unit [25]. Refinement of the solution by AMORE gave a correlation coefficient of 0.77 and an initial *R*-factor of 0.31 [26]. After automatic CNS refinement, the new set of coordinates resulted in *R*-test of 0.261 and an *R*-work of 0.214. After several more cycles of refinement with CNS and REFMAC5, manual model building with the program O [27–29], and the addition of water molecules, a final structural model was obtained. A summary of the refinement statistics and geometric quality of the model is given in Table 1.

2.5. Crystallization and structure determination of inactivated *cis*-CaaD

cis-CaaD was covalently modified at Pro-1 by incubating the enzyme with a 100-fold excess of (*R*)-oxirane-2-carboxylate (**6**) and allowing the mixture to sit overnight at room temperature [30]. After removing excess inhibitor by gel filtration chromatography, the inactivated *cis*-CaaD was concentrated to 15.6 mg/mL in 10 mM Tris-SO_4 buffer at pH 8.0. Crystals of inactivated *cis*-CaaD were obtained from 6-μL hanging drops consisting of equal amount of precipitant solution (0.125 M $CaCl_2$, 0.07 M sodium acetate buffer, 12.5% isopropanol, pH 4.6) and the concentrated protein sample. The cubic-shaped crystals grew within one week to ~0.3 × 0.3 × 0.3 mm in size. A diffraction data set to 1.65 Å resolution (82 images at 0.5° each) was collected in-house using a Rigaku RU200H rotating anode X-ray generator (Cu radiation) equipped with a R-AXIS IV++ image plate detector with a crystal-to-detector distance of 100 mm. The data were integrated and scaled using the HKL-2000 program package [22]. The crystals belong to the space group I23 with cell constants a = 96.78 Å. The asymmetric unit contains one monomer of 149 residues, with a calculated Matthews' coefficient of 2.31 $Å^3$/Da corresponding to a solvent content of 47% [23]. Residues 119–149 were not resolved in the electron density map.

Molecular replacement solutions were obtained as described above for the native structure. The molecular replacement yielded the position and orientation of one monomer in the asymmetric unit. Refinement of the solutions by AMORE gave a correlation coefficient of 0.52 and an *R*-factor of 0.43 [26]. After an automatic refinement with CNS a new set of coordinates with an *R*-test of 0.437 and an *R*-work of 0.348 were generated. After several refinement rounds with CNS and REFMAC5, manual model building with the program O [27–29], and the addition of water molecules, a final structural model was obtained. A summary of the refinement statistics and geometric quality of the model is given in Table 1.

Table 1
Data collection and refinement statistics.

	Native *cis*-CaaD	Inactivated *cis*-CaaD
Data statistics		
Space group	$P6_3$	I23
No. chains/asymmetric unit	1	1
Unit cell (Å)	$a = b$ = 58.99 c = 58.30	$a = b = c$ = 96.78
Resolution (Å)	2.01	1.65
R_{sym} (%) overall (outer shell)[a]	6.0 (14.0)	7.0 (25.0)
Completeness (%) overall (outer shell)	98.3 (84.6)	98.2 (99.9)
I/σ (outer shell)	46.6 (57.0)	26.3 (9.7)
Reflections total (unique)	85,402 (7766)	81,955 (17,927)
Refinement		
Total atoms/water	1018/101	1048/120
R/R_{free}[b]	16.5/22.8	19.0/22.4
rmsd bonds (Å)/angles (°)	0.01/1.41	0.01/1.17
rmsd B ($Å^2$) (main chain: bond/angle/side chain: bond/angle)	0.97/1.61	0.67/1.17
	5.30/4.51	1.90/3.00
Ramachandran plot (%) (favored/allowed/generously allowed/disallowed)	97.0/3.0/0/0	94.1/5.0/1.0/0

Note: The outer shell value (i.e., I/σ) of the native *cis*-CaaD is unusually high because diffraction data to a higher resolution were not collected on the high quality crystal.

[a] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity and $\langle I\rangle$ the average intensity.

[b] R based on 95% of the data used in the refinement. R_{free} = R based on 5% of the data withheld from the cross-validation test.

2.6. Construction of the R117A-cis-CaaD mutant

The R117A mutant of *cis*-CaaD was generated using the coding sequence for *cis*-CaaD in plasmid pET(*cis*-CaaD) as the template. The mutant was constructed using the QuikChange mutagenesis kit and the indicated set of primers (*vide infra*) following the manufacturer's instructions. The forward primer was 5′-GGTGGAGTACGGCGCGTTCCTGCCCCAGCCC-3′, and the reverse primer was 5′-GCTTCTCTGTACGCCCCGAAGCAATCGTTGCTTGGACCC-3′. In each set of primers, the mutation is underlined and the remaining bases correspond to the coding sequence (forward primer) or the complementary sequence (reverse primer). DNA sequencing verified that only the intended mutation had been introduced into the mutant genes.

2.7. Production and purification of the R117A-cis-CaaD mutant

The mutant was expressed and purified using a protocol adapted from the one described for the wild-type enzyme [10]. In order to eliminate the possibility of contaminating proteins, the enzyme was purified using disposable hand-packed columns [17]. Typically, in this protocol, cells from 1 L of culture were suspended in ~8 mL of 10 mM Na_2HPO_4 buffer, pH 8.0, (Buffer A), sonicated, and centrifuged. Subsequently, the supernatant was loaded onto a DEAE-Sepharose column (10 × 1.0 cm filled with 8 mL of resin) that had been previously equilibrated with Buffer A. The column was first washed with Buffer A (25 mL) and then the protein was eluted by gravity using a linear Na_2SO_4 gradient (0–0.5 M Na_2SO_4 in Buffer A, 100 mL). The flow rate was estimated to be ~1 mL/min. Fractions (~1.5 mL) were collected and the R117A-mutant of *cis*-CaaD was identified by SDS–PAGE. The mutant eluted 4.5–7.5 min after being loaded onto the column. The appropriate fractions were pooled and made 1 M in $(NH_4)_2SO_4$ by the slow addition of an aliquot of 10 mM Na_2HPO_4 buffer, pH 8.0, containing 2 M $(NH_4)_2SO_4$. After stirring for 1 h, the precipitate was removed by centrifugation (15 min at 20,000*g*), and the

supernatant was filtered and loaded onto a Phenyl-Sepharose column (10 × 1.0 cm filled with 8 mL of resin) that had been previously equilibrated with Buffer A containing 1 M $(NH_4)_2SO_4$. The column was first washed with the loading buffer (25 mL) and then the protein was eluted by gravity using a decreasing linear $(NH_4)_2SO_4$ gradient [1.6–0 M $(NH_4)_2SO_4$ in Buffer A, 100 mL]. The flow rate was estimated to be ~1 mL/min. Fractions (~1.5 mL) were collected and analyzed as described above. The R117A mutant of *cis*-CaaD eluted 31.5–54 min after being loaded onto the column. The purified protein was concentrated to ~5 mg/mL, filtered through a 0.2 mm-pore diameter filter, and stored at 4 °C. A typical yield is ~100 mg of protein purified to homogeneity per liter of culture.

2.8. Kinetic assays

The kinetic assays were performed at 22 °C by following the decrease in absorbance at 224 nm, which corresponds to the hydration of *cis*-3-chloroacrylic acid (**2**, $e = 2900\ M^{-1}\ cm^{-1}$) [10]. An aliquot of the R117A-mutant was diluted into 20 mM Na_2HPO_4 buffer (pH 9.0), yielding a final enzyme concentration of 1.24 mM, 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 a 5 mM 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*-3-chloroacrylic acid (as the free acid) to this buffer adjusted the pH of the stock solution to ~7. The 5 mM stock solution was 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 10 to 50 mM.

2.9. Examination of (R)-oxirane-2-carboxylate (**6**) as an Irreversible Inhibitor of the R117A-mutant of cis-CaaD

The R117A mutant (~1 mg/mL, ~20 mM in monomer concentration) was incubated separately with (*R*)-oxirane-2-carboxylate (0.775 mM) in 10 mM NaH_2PO_4 buffer (pH 8.0) [30]. The mixture (total volume of 1 mL) was incubated for 4 h at 4 °C. An aliquot (10 μL) of the mixture was then withdrawn, diluted into 1 mL of 20 mM Na_2HPO_4 buffer (pH 9.0), and assayed for residual activity as described above (following the decrease in absorbance at 224 nm). The activity assay was initiated by the addition of a small quantity of the substrate to give a final concentration of 200 mM. The R117A mutant is saturated under these conditions. Stock solutions (50 mM) of *cis*-3-chloroacrylic acid were made up in 100 mM Na_2HPO_4 buffer, and the pH was adjusted to 7.3. Stock solutions (10 mM) of the inhibitor were made up in 10 mM NaH_2PO_4 buffer (pH 8.0).

2.10. Mass spectral analysis of the (R)-oxirane-2-carboxylate-treated R117A-mutant of cis-CaaD

The incubation mixture from the previous experiment was loaded onto a PD-10 Sephadex G-25 gel filtration column, which had previously been equilibrated with 100 mM NH_4HCO_3 buffer (pH 8.0) [10,30]. The protein was eluted by gravity flow using the same buffer. Fractions (0.5 mL) were analyzed for the presence of protein by UV absorbance at 215 nm. The appropriate fraction containing the purified protein was analyzed by electrospray ionization mass spectrometry (ESI-MS). In a separate control experiment, the same quantity of the R117A-mutant of *cis*-CaaD was incubated without inhibitor under otherwise identical conditions and analyzed by ESI-MS.

3. Results

3.1. Structure of native cis-CaaD

The native *cis*-CaaD crystal structure was solved to 2.01-Å resolution by molecular replacement methods and refined to *R* and R_{free} values of 16.2% and 22.8%, respectively. The asymmetric unit contains one monomer composed of two repeats of the β−α−β structural motif, which is the characteristic tautomerase superfamily fold [12–14]. In the *cis*-CaaD monomer, Pro-1 to Gln-9 make up a β-strand (β-1), Asp-10 to Thr-13 is in a loop that connects β-1 to the first α-helix (α-1, Pro-14 to Arg-29), and Gly-30 to Ala-39 comprise a loop that connects α-1 to the parallel β-strand (β-2, Gln-40 to Gln-46). The sequence Pro-47 to Thr-62 is in a loop that connects β-2 to the third β-strand (β-3, Ile-63 to Glu-71), which is anti-parallel to β-2, Gly-72 to Ser-74 comprise a loop that connects β-3 to the second α-helix (α-2, Ala-75 to Ala-93), Glu-94 to Lys-98 is in a loop that connects α-2 to the fourth β-strand (β-4, His-99 to Met-107), which is parallel to β-3. Finally, Pro-108 to Arg-117 is in a loop, but the electron densities for the amino acid residues after Arg-117 are not well-defined.

The disordered C-terminus (residues 118–149) is one notable difference between this structure and that reported by de Jong et al. [16]. The de Jong structure was crystallized at low pH (i.e., pH 4.2) whereas this structure was crystallized at higher pH (i.e., pH 9.0). The higher pH likely altered the hydrogen bond and/or electrostatic interactions so that the C-terminus is unstructured and not observed. It should be noted that the inactivated *cis*-CaaD structure reported by de Jong et al. was obtained at pH 7.5 and also shows a disordered C-terminus (residues 117–149) [16].

The structure shows several clusters of polar residues at the monomer–monomer interfaces that help stabilize the trimer (Fig. 1). This includes the pseudo 2-fold related clusters of (Lys-17, Glu-45, Asn-50) and (Lys-78, Glu-106, Gln-111) at each end of the interface, and Tyr-103 and Glu-114 near the active site

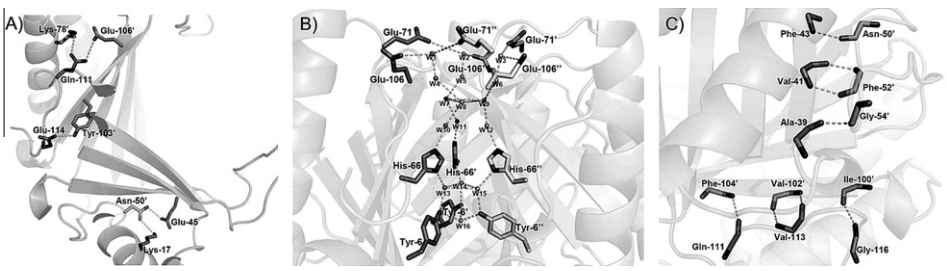


Fig. 1. A number of hydrogen bonding interactions contribute to trimer stability. (A) The interface of two *cis*-CaaD monomers showing selected residues that form a series of hydrogen bonds. (B) The convergence of three monomers showing a well-ordered water network stabilized by amino acid residues near the molecular 3-fold axis. (C) Hydrogen bonds between the C-terminal loop of one monomer with residues from the β-sheet of an adjacent monomer. This Figure and subsequent figures were prepared using PyMol [31].

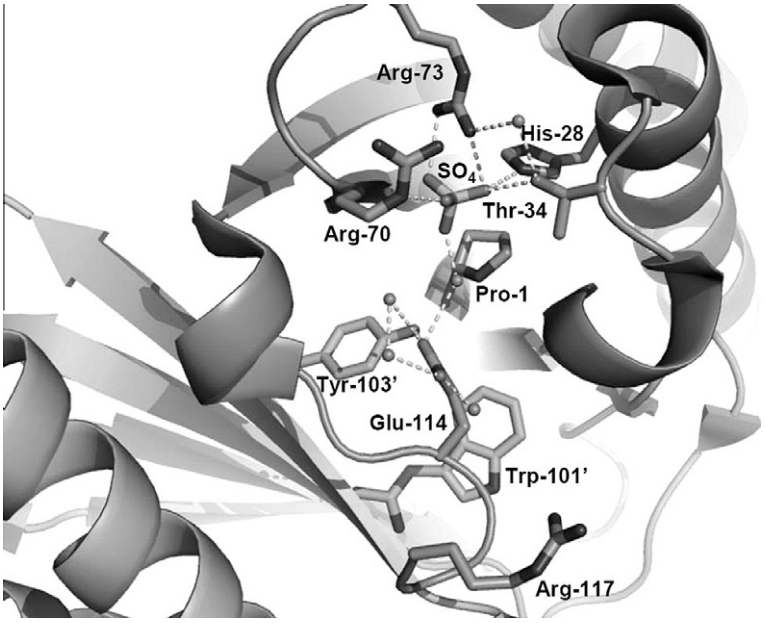


Fig. 2. A close-up view of an active site of native *cis*-CaaD. The key catalytic residues are shown along with Thr-34, Trp-101′, and Arg-117 (Footnote 1).

proline that interact through a series of hydrogen bonds (Fig. 1A).[1] For example, the carbonyl group of Asn-50 forms a hydrogen bond (3.0 Å) with the ε-amino group of Lys-17 and the amide nitrogen forms a hydrogen bond (2.8 Å) with a carboxylate oxygen of Glu-45. Similarly, the carbonyl group of Gln-111 forms a hydrogen bond (2.7 Å) with the ε-amino group of Lys-78 and the amide nitrogen forms a hydrogen bond (2.9 Å) with a carboxylate oxygen of Glu-106. The hydroxyl group of Tyr-103 forms a hydrogen bond (2.7 Å) with a carboxylate oxygen of Glu-114. These interactions play a major role in the stability of the trimer.

Near the crystallographic and inherently inner molecular 3-fold there are well-ordered water molecules that appear to play roles in stabilizing the trimer through a hydrogen bonding interaction network (Fig. 1B). These water molecules are coordinated by the side chains of Tyr-6, His-66, Glu-71, and Glu-106. The same residue is contributed by each monomer (i.e., Tyr-6, Tyr-6, and Tyr-6).

Finally, hydrogen bonds between the C-terminal loop of one monomer and the β-sheet of an adjacent subunit also contribute to the stability of the trimer (Fig. 1C). For example, there are hydrogen bonds between the oxygen atom of the backbone carbonyl group of Gln-111 and the amide proton of the backbone of Phe-104, the oxygen atom of the backbone carbonyl group of Val-113 and the amide proton of the backbone of Val-102, and the amide proton of the backbone of Gly-116 and the oxygen of the backbone carbonyl group of Ile-100. There are also interactions between Phe-43 and Asn-50, Val-41 and Phe-52, and Ala-39 and Gly-54.

3.2. The active site of native cis-CaaD

A visual inspection of the active site shows that it is mostly hydrophilic with several charged residues (both positively and negatively charged) and an extensive water molecule network (Fig. 2). A calculation of the electrostatic surface potential map of the active site illustrates this observation and is consistent with the function of *cis*-CaaD (i.e., hydrolytic dehalogenation) and the assumed pK_a of 9.2 for Pro-1 in *cis*-CaaD [10,30].

cis-CaaD is a trimer with three active sites where an N-terminal proline is centered in each active site [16]. A sulfate or phosphate ion (from the crystallization buffer) is bound in the active site (Fig. 2). de Jong et al. made the same observation [16]. The sulfate ion interacts with the side chains of Thr-34 and Arg-73 and the backbone amide nitrogen of Arg-70. The prolyl nitrogen interacts with the oxygen atom of one of the active site water molecules, which also interacts with the sulfate ion (Fig. 2). Additionally, the carboxylate oxygen of Glu-114 forms a hydrogen bond with the hydroxyl group of Tyr-103′, which is from a second subunit of the *cis*-CaaD trimer. His-28, Arg-70, Trp-101′ and Arg-117, the last residue of the polypeptide chain with defined electron density, are also located in the active site.

The different position of the sulfate ion and crystallization pH value can account for the subtle differences in active site configurations [16]. The de Jong structure was crystallized at low pH, potentially making the side chain of Glu-114 predominately protonated and uncharged [16]. In both structures, one side chain oxygen interacts with the hydroxyl group of Tyr-103′. In the de Jong structure, the side chain carboxylic moiety shifts, placing an oxygen within hydrogen bonding distance of the backbone carbonyl group of Leu-38, reflecting its presumed protonation. In the de Jong structure, the side chains of Arg-70 and Arg-73 also show slightly altered conformations (Fig. 3) [16]. Different rotamers are also observed for the side chains of His-36, Gln-35, and Arg-29, but the backbone elements superimpose.

The loop, which connects the α-1 helix to the parallel β-strand (β-2), is comprised of residues 32–38 (Thr-Gly-Thr-Gln-His-Phe-Leu), adopts the same position in both native structures (designated the open state). With the exception of the different rotamers observed for the side chains of Gln-35 and His-36, both the backbone and side chains of the loop residues superimpose. As explained below, the position of this loop may have implications for the catalytic mechanism (*vide infra*).

3.3. Structure of cis-CaaD inactivated by (R)-**6**

The inactivated complex crystal structure of *cis*-CaaD was solved to 1.65 Å resolution by molecular replacement methods and refined to R and R_{free} values of 19.0% and 22.3%, respectively. The asymmetric unit contains one monomer. The active site shows the proline covalently attached to the expected 2-hydroxypropanoate adduct, generated by the attack of the prolyl nitrogen at C-3 of (*R*)-**6** (Fig. 4 and Scheme 3) [16]. The adduct is in a distinctly

[1] The primed residues refer to different subunits within the *cis*-CaaD trimer.

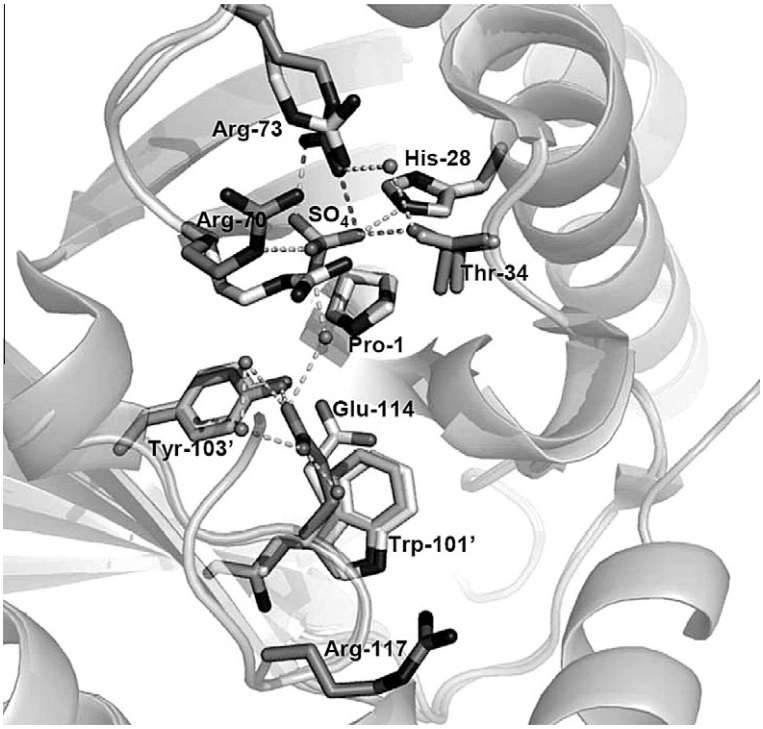


Fig. 3. Superposition of the active site of native *cis*-CaaD reported here (gray) and that reported by de Jong (white) [16].

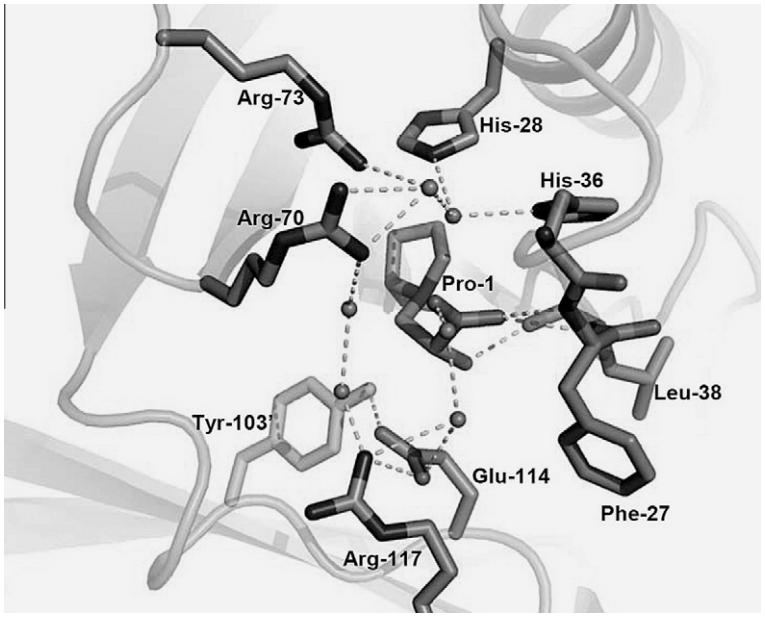
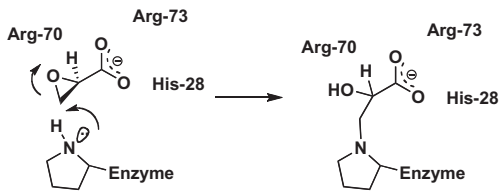


Fig. 4. A close-up view of an active site of inactivated *cis*-CaaD where the Pro-1 is covalently bonded to 2-hydroxypropanoate via the prolyl nitrogen (as shown in Scheme 3).



Scheme 3.

different conformation from that previously reported (Fig. 5) [16]. Moreover, its presence disrupts the hydrogen bond network seen in the native *cis*-CaaD and results in a different hydrogen bond network. The C-2 hydroxyl group of the adduct forms a hydrogen bond with the backbone carbonyl group of the Leu-38 in the 7-membered loop (Thr-32 to Leu-38). One carboxylate oxygen of the adduct forms a hydrogen bond with the backbone amide nitrogens of Phe-37 and Leu-38. The other carboxylate oxygen forms a hydrogen bond with a water molecule that is held in place by the side chains of His-28 and His-36. The interactions between the loop residues and the adduct reposition the loop such that

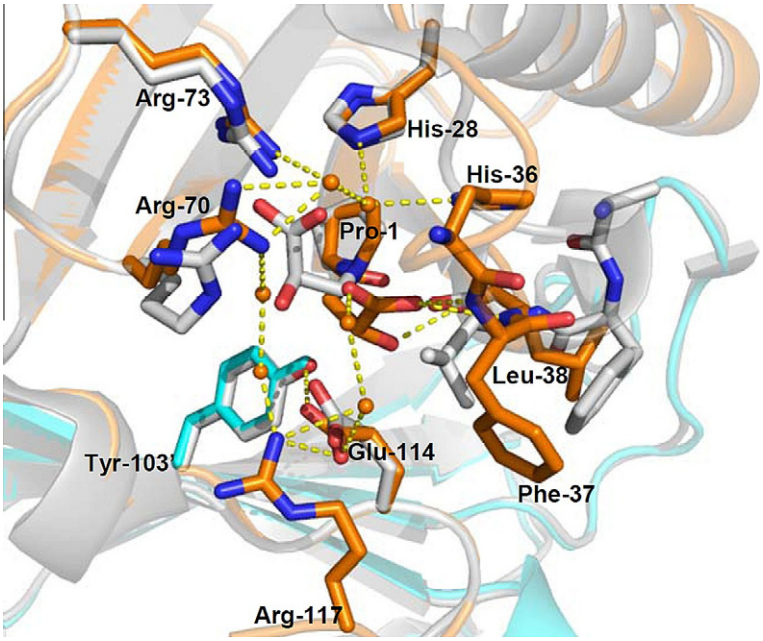


Fig. 5. Superposition of the active sites of inactivated *cis*-CaaD reported here (color) and that reported by de Jong (white) [16]. Note the striking differences in the orientations of the 2-hydroxypropanoate adduct. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

some residues point inward (i.e., the backbone of Phe-37 and Leu-38 and the side chain of His-36) and others point up and away from Pro-1 (i.e., the backbone of Thr-34 and Gln-35) or above Pro-1 (i.e., the side chain of Gln-35 moves towards Arg-70). This conformation is designated the closed state and contrasts with the loop's position in both native structures and the inactivated structure reported by de Jong et al. (Fig. 5) [16].

The residues that interact with the adduct in the inactivated structures reported here and by de Jong et al. make different contacts. Accordingly, one carboxylate oxygen of the Glu-114 forms a hydrogen bond with the phenolic hydroxyl group of Tyr-103′ from a second monomer and the other carboxylate oxygen forms a hydrogen bond with a nearby water molecule and the ηN atom of Arg-117. Arg-70 and Arg-73 form hydrogen bonds with several water molecules. One such water molecule is within hydrogen bond distance with a second water molecule, that is, in turn, near the water molecule positioned by His-28, His-36, and the covalent adduct. All six water molecules together with the adduct and amino acid residues in the active site form a more complex hydrogen bond network than the one in the native structure of *cis*-CaaD.

3.4. Kinetic characterization of the R117A mutant of cis-CaaD

The kinetic parameters for the R117A mutant of *cis*-CaaD were determined using *cis*-3-chloroacrylic acid (**2**) and compared to those measured for the wild type. For the wild-type enzyme, the K_m is 34 ± 8 mM, the k_{cat} is 1.8 ± 0.2 s^{-1}, and the k_{cat}/K_m is 5.3×10^4 M^{-1} s^{-1} [10,17]. For the R117A mutant, the K_m is 15 ± 3 mM (down 2.3-fold), the k_{cat} is 0.4 ± 0.1 s^{-1} (down 4.5-fold), and the k_{cat}/K_m is 2.6×10^4 M^{-1} s^{-1} (down 2-fold). Although the mutation of Arg-117 has some effect on the kinetic parameters, it is not a major one (such as that observed for the Pro-1, His-28, Arg-70, Arg-73, Tyr-103, or Glu-114).

3.5. ESI-MS analysis of the R117A mutant of cis-CaaD treated with (R)-6

In order to determine whether Arg-117 plays a role in the alkylation reaction, the R117A mutant of *cis*-CaaD was treated with (*R*)-**6** for 24 h at 4 °C. Subsequently, the (*R*)-**6**-treated protein sample was purified by gel filtration chromatography and analyzed by ESI-MS. An untreated sample of the R117A mutant showed a signal corresponding to a molecular mass of 16,536 ± 2 Da. Mass spectral analysis of the (*R*)-**6**-treated sample of the R117A mutant indicated that it consists of one major component with an observed molecular mass of 16,624 ± 2 Da (data not shown), which corresponds to the expected molecular mass of the R117A mutant covalently modified by **6** (i.e., a mass increases of 88 Da). Previous work has shown that Pro-1, Arg-70, and Arg-73 are important for the inactivation reaction but Glu-114 is not [30]. Hence, like Glu-114, Arg-117 is not important for the alkylation reaction.

4. Discussion

Sequence analysis, mutagenesis experiments, inhibition studies, and previous crystallographic structures were instrumental in formulating working hypotheses for the catalytic mechanism of *cis*-CaaD as well as its inactivation by the epoxide, (*R*)-**6** [9,10,16,30]. The de Jong crystal structure of *cis*-CaaD inactivated by (*R*)-**6** established that C-3 is the site of attack by the prolyl nitrogen of Pro-1 (Scheme 3) [16]. Furthermore, the alignment of the 2-hydroxyl group and the carboxylate oxygens of the ring-opened 2-hydroxypropanoate adduct with Arg-70, Arg-73, and His-28, respectively, implicated these residues as critical ones in the inactivation mechanism [16]. Based on the observed interactions, it was proposed that Arg-73 and His-28 are primarily involved in binding the carboxylate group of (*R*)-**6** at the active site whereas Arg-70 facilitates ring opening by directly interacting with the epoxide oxygen or by placing a water molecule in position to interact with the oxygen (Scheme 3) [16]. The observations were supported by experiments showing that Pro-1, Arg-70, and Arg-73 are required for the inactivation process, but not Glu-114 [30].

The structure of *cis*-CaaD inactivated by (*R*)-**6** reported here is seemingly at odds with the proposed inactivation mechanism because the ring-opened adduct does not show any of the same interactions except for the covalent bond between C-3 and the prolyl nitrogen of Pro-1. Instead, the adduct interacts with a 7-member

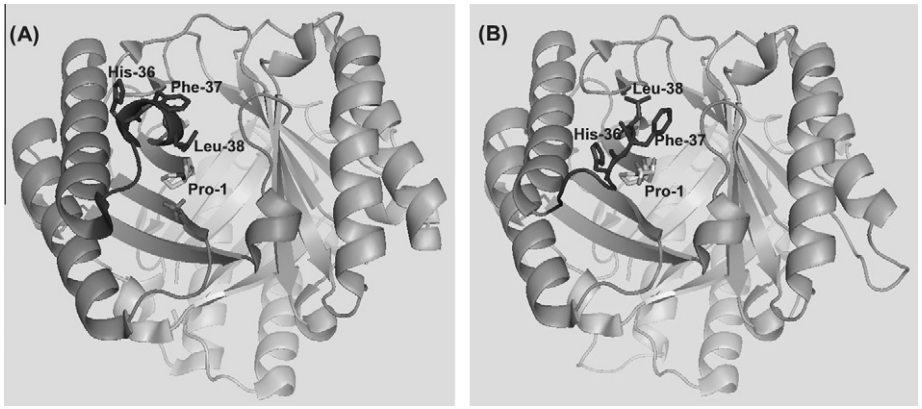


Fig. 6. Substrate/inhibitor binding may trigger a significant conformational change of loop residues 32–38 in *cis*-CaaD from (A) an open state (native *cis*-CaaD) to (B) a closed state (inactivated *cis*-CaaD). Only His-36, Phe-37, and Leu-38 are shown for illustrative purposes.

loop consisting of residues 32–38 (Thr-Gly-Thr-Gln-His-Phe-Leu) where the 2-hydroxyl group forms a hydrogen bond with the backbone carbonyl group of Leu-38 and one C-1 carboxylate oxygen interacts with the backbone amide nitrogens of Phe-37 and Leu-38 (through hydrogen bonding). (The other C-1 carboxylate oxygen interacts with a water bound to His-28 and His-36.) Two observations suggest that this structure does not accurately reflect the inactivation mechanism: it is unlikely that the carbonyl group of Leu-38 assists in ring opening of the epoxide and this crystal structure was obtained at pH 4.6 in the presence of 12.5% isopropanol. Hence, the actual interactions made during the inactivation process might have been distorted by the low pH and isopropanol, and the de Jong structure (obtained at pH 7.5) likely presents a more accurate portrayal of the residues involved in the inactivation process [16].

Nonetheless, the position of the loop in our inactivated structure (compared to its position in the three other known structures) suggests that the loop region is floppy and raises an interesting possibility: the loop may play a role in the catalytic cycle. There are two lines of indirect evidence supporting this possibility. First, a pre-steady state analysis of the *cis*-CaaD reaction shows a change in fluorescence, immediately upon substrate binding. The two fluorescent states might correspond to different states of the active sites, which are designated the open and closed states. Movement of the loop (upon substrate binding) could put the enzyme in the closed state so that catalysis can occur. The second piece of evidence comes from studies of a *cis*-CaaD homologue (designated Cg10062) from *Corynebacterium glutamicum*. Cg10062 has ~53% sequence similarity with *cis*-CaaD and conserves the six active site residues implicated as critical ones in the *cis*-CaaD mechanism and the other residues that define the active site region. However, Cg10062 is not an efficient or highly specific *cis*-CaaD. It has very poor *cis*-CaaD activity (as assessed by k_{cat}/K_m value) and processes both the *cis*- and *trans*-isomers. Interestingly, the 7-member loop in Cg10062 consists of very different residues (Ala-32, His-33, Ala-34, Pro-35, Lys-36, Tyr-37, and Leu-38). It may be that the bulky charged group (His-33), the absent Thr-34 (which may form part of a binding pocket for the chloro group), and the rigid Pro-35 alter the loop properties such that it can no longer adopt the closed conformation. The inability to do this could account for the poor *cis*-CaaD activity of Cg10062.

The position of Arg-117 in the crystal structures reported here and elsewhere suggested that it might also contribute to the mechanism of inactivation and/or catalysis. Notably, in the inactivated structure, the ηN atom of Arg-117 interacts with a water molecule that forms a hydrogen bond with Arg-70, Arg-73, and a carboxylate oxygen of Glu-114. For these reasons, it was changed to an alanine and the properties of the resulting R117A mutant were characterized. Changing Arg-117 to an alanine has little effect on the steady state kinetic parameters and no effect on the alkylation of Pro-1 when incubated with (*R*)-**6**. The combined findings suggest that Arg-117 is not a key residue in either process.

The crystallographic observations reported here add the 7-residue loop as a potential element in the proposed mechanisms for *cis*-CaaD catalysis and its inactivation by (*R*)-**6** (Fig. 6). In catalysis, the loop is initially in the open state (Fig. 6A). Substrate binding may trigger movement of the loop such that it adopts the closed conformation (Fig. 6B). In this closed position, the loop may also complete the active site and put determinants of specificity in place. Once catalysis is complete, the loop may adopt the open conformation and products can dissociate (Fig. 6A). The loop residues are not highly conserved between *cis*-CaaD and CaaD (Thr-32, Gly-33, Thr-34, Gln-35, His-36, Phe-37, and Leu-38 for *cis*-CaaD and Ile-32, Gly-33, Ser-34, Asp-35, Pro-36, Lys-37, and Ile-38 for the β-subunit of CaaD) so that a similar structural element is not part of the CaaD catalytic mechanism. Evidently, epoxide binding also triggers loop movement, but the resulting covalent modification of Pro-1 freezes the enzyme in the closed conformation. A mutagenesis study of the loop residues is underway to determine if the crystallographic observations are representative of catalytic events or if they are artifacts.

Acknowledgments

The authors thank Dr. Gottfried K. Schroeder (The University of Texas at Austin) for helpful comments and discussions.

References

[1] B. Hileman, Chem. Eng. News 71 (1993) 11–20.
[2] S.W. Hooper, C.A. Pettigrew, G.S. Sayler, Environ. Toxicol. Chem. 9 (1990) 655–667.
[3] J.A. Shapiro, Gene 345 (2005) 91–100.
[4] M.M. Benning, K.L. Taylor, R.Q. Liu, G. Yang, H. Xiang, G. Wesenberg, D. Dunaway-Mariano, H.M. Holden, Biochemistry 35 (1996) 8103–8109.
[5] S.D. Copley, in: D. Barton, K. Nakanishi (Eds.), Comprehensive Natural Products Chemistry, vol. 5, Elsevier, Amsterdam, 1999, pp. 401–422.
[6] D.A. Abramowicz, Crit. Rev. Biotechnol. 10 (1990) 241–249.
[7] J.E.T. van Hylckama Vlieg, D.B. Janssen, Biodegradation 2 (1992) 139–150.
[8] G.J. Poelarends, M. Wilkens, M.J. Larkin, J.D. van Elsas, D.B. Janssen, Appl. Environ. Microbiol. 64 (1998) 2931–2936.
[9] S.C. Wang, M.D. Person, W.H. Johnson Jr., C.P. Whitman, Biochemistry 42 (2003) 8762–8773.
[10] G.J. Poelarends, H. Serrano, M.D. Person, W.H. Johnson Jr., A.G. Murzin, C.P. Whitman, Biochemistry 43 (2004) 759–772.
[11] A.G. Murzin, Curr. Opin. Struct. Biol. 6 (1996) 386–394.
[12] C.P. Whitman, Arch. Biochem. Biophys. 402 (2002) 1–13.
[13] G.J. Poelarends, C.P. Whitman, Bioorg. Chem. 32 (2004) 376–392.

[14] G.J. Poelarends, C.P. Whitman, Cell. Mol. Life Sci. 65 (2008) 3606–3618.
[15] R.M. de Jong, W. Brugman, G.J. Poelarends, C.P. Whitman, B.W. Dijkstra, J. Biol. Chem. 279 (2004) 11546–11552.
[16] R.M. de Jong, P. Bazzacco, G.J. Poelarends, W.H. Johnson Jr., Y.J. Kim, E.A. Burks, H. Serrano, A.M. Thunnissen, C.P. Whitman, B.W. Dijkstra, J. Biol. Chem. 282 (2007) 2440–2449.
[17] G.J. Poelarends, H. Serrano, M.D. Person, W.J. Johnson Jr., C.P. Whitman, Biochemistry 47 (2008) 8139–8147.
[18] Y. Petit, M. Larcheveque, Organic Synth. 75 (1998) 37–44.
[19] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.
[20] U.K. Laemmli, Nature 227 (1970) 680–685.
[21] W.J. Waddell, J. Lab. Clin. Med. 48 (1956) 311–314.
[22] Z. Otwinowski, W. Minor, in: C.W. Carter Jr., R.M. Sweet (Eds.), Methods in Enzymology, vol. 276, Academic Press, New York, 1997, pp. 307–326.
[23] B.W. Matthews, J. Mol. Biol. 33 (1968) 491–497.
[24] J-B. Claude, K. Suhre, C. Notredame, J-M. Claverie, C. Abergel, Nucleic Acids Res. 32 (2004) W606–W609. <http://www.igs.cnrs-mrs.fr/Caspr2/index.cgi>.
[25] Collaborative Computational Project, Number 4, Acta Crystallogr. D 50 (1994) 760–763.
[26] J. Navaza, Acta Crystallogr. A 50 (1994) 157–163.
[27] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Acta Crystallogr. D 54 (1998) 905–921.
[28] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Acta Crystallogr. D 53 (1997) 240–255.
[29] T.A. Jones, J.-Y. Zou, S.W. Cowan, M. Kjeldgaard, Acta Crystallogr. A 47 (1991) 110–119.
[30] G.J. Poelarends, H. Serrano, W.J. Johnson Jr., C.P. Whitman, Biochemistry 43 (2004) 7187–7196.
[31] W.L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, 2002, <http://www.pymol.org>.