Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic Chemistry 32 (2004) 376–392
BIOORGANIC CHEMISTRY
www.elsevier.com/locate/bioorg

Mini-review

Evolution of enzymatic activity in the tautomerase superfamily: mechanistic and structural studies of the 1,3-dichloropropene catabolic enzymes☆

Gerrit J. Poelarends and Christian P. Whitman*

Division of Medicinal Chemistry, College of Pharmacy, The University of Texas, Austin, TX 78712-1074, USA

Received 26 April 2004
Available online 11 June 2004

Abstract

The use of the soil fumigant Telone II, which contains a mixture of *cis*- and *trans*-1, 3-dichloropropene, to control plant-parasitic nematodes is a common agricultural practice for maximizing yields of various crops. The effectiveness of Telone II is limited by the rapid turnover of the dichloropropenes in the soil due to the presence of bacterial catabolic pathways, which may be of recent origin. The characterization of three enzymes in these pathways, *trans*-3-chloroacrylic acid dehalogenase (CaaD), *cis*-3-chloroacrylic acid dehalogenase (*cis*-CaaD), and malonate semialdehyde decarboxylase (MSAD), has uncovered intriguing catalytic mechanisms as well as a fascinating evolutionary lineage for these proteins. Sequence comparisons and mutagenesis studies revealed that all three enzymes belong to the tautomerase superfamily. Tautomerase superfamily members with known structures are characterized by a β-α-β structural fold. Moreover, they have a conserved N-terminal proline, which plays an important catalytic role. Mechanistic, NMR, and pH rate studies of the two dehalogenases, coupled with a crystal structure of CaaD inactivated by 3-bromopropiolate, indicate that they use a general acid/base mechanism to catalyze the conversion of their respective isomer of 3-chloroacrylate to malonate semialdehyde. The reaction is initiated by the conjugate addition of water to the C-2, C-3 double bond and is followed by the loss of HCl. MSAD processes malonate semialdehyde to acetaldehyde, and is the first identified decarboxylase in the tautomerase superfamily.

☆ This research was supported by the National Institutes of Health Grant GM-65324.
* Corresponding author. Fax: 1-512-232-2606.
E-mail address: whitman@mail.utexas.edu (C.P. Whitman).

doi:10.1016/j.bioorg.2004.05.006

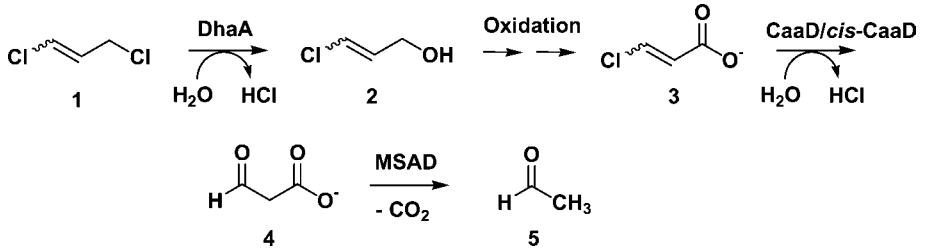
The catalytic mechanism is not well defined but the N-terminal proline plays a prominent role and may function as a general acid catalyst, similar to its role in CaaD and *cis*-CaaD. These are the first structural and mechanistic details for tautomerase superfamily members that catalyze either a hydration or a decarboxylation reaction, rather than a tautomerization reaction, in which Pro-1 serves as a general acid catalyst rather than as a general base catalyst. The available information on the 1,3-dichloropropene catabolic enzymes allows speculation on the possible evolutionary origins of their activities.

Keywords: Tautomerase superfamily; 1,3-Dichloropropene; Nematocide; 3-Chloroacrylic acid dehalogenase; Malonate semialdehyde decarboxylase; Enzyme evolution

1. Introduction

Isomeric 1,3-dichloropropenes are toxic synthetic organohalogens. The lethality of these chemicals has been exploited in the agricultural nematocides Shell D-D and Telone II, where they are the active ingredients. These commercial fumigants are mixed in with the soil to control plant-parasitic nematodes, a common practice to increase the yield of various crops [1]. Because 1,3-dichloropropenes were only introduced into the environment in the 1950s and are not known to be produced naturally in significant concentrations, environmental exposure to these substances has been limited. Despite the toxic properties and short period of exposure, microorganisms have apparently assembled catabolic routes in order to utilize one or both isomers of 1,3-dichloropropene as their sole source(s) of carbon and energy [2–6]. De novo evolution of genes for 1,3-dichloropropene degradation during this short period seems unlikely. Instead, it is more likely that microorganisms assembled functional catabolic pathways by the recruitment and adaptation of enzymes from pathways for naturally occurring compounds. Analysis of the genes and enzymes that make up the 1,3-dichloropropene catabolic pathway makes it possible to distinguish between the two adaptation mechanisms [7–9].

Pseudomonas pavonaceae 170 is the best studied of the 1,3-dichloropropene-degrading bacteria. In this organism, the conversion of 1,3-dichloropropene to acetaldehyde is accomplished by the action of five enzymes (Scheme 1) [5,10,11].



Scheme 1.

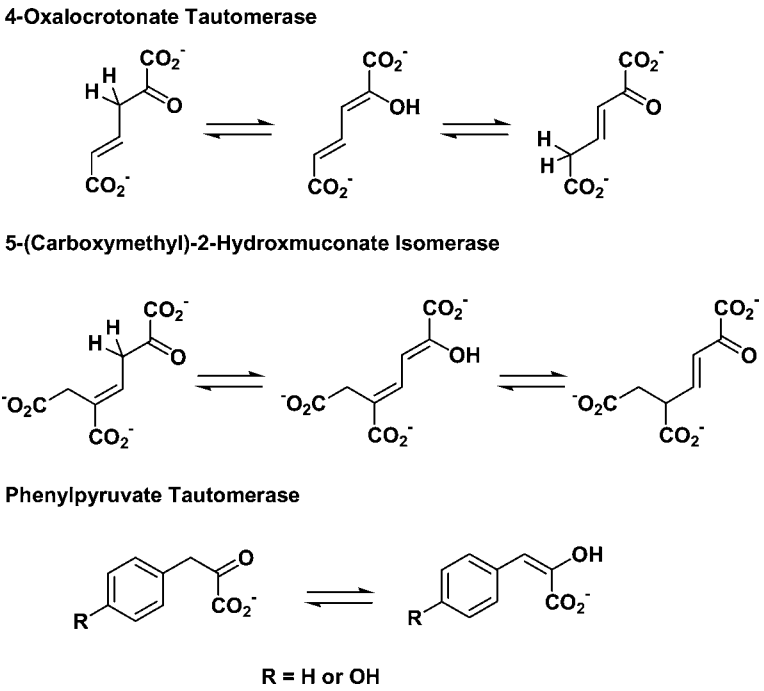
The pathway begins with the conversion of both isomers of 1,3-dichloropropene (**1**) to the corresponding isomers of 3-chloroallyl alcohol (**2**) by a haloalkane dehalogenase (DhaA).[1] The next two steps involve the oxidation of the isomeric 3-chloroallyl alcohol intermediates to *cis*- and *trans*-3-chloroacrylate (**3**). In the fourth step, both isomers of 3-chloroacrylate are converted to malonate semialdehyde (**4**), but by two different dehalogenating enzymes, one specific for the *trans*-isomer and one specific for the *cis*-isomer of 3-chloroacrylate (CaaD and *cis*-CaaD, respectively). To complete the pathway, malonate semialdehyde is converted to acetaldehyde (**5**) by malonate semialdehyde decarboxylase (MSAD).

This review will discuss recent advances in our understanding of the structures and mechanisms of the enzymes responsible for the last two steps of the pathway (i.e., CaaD, *cis*-CaaD, and MSAD). The evidence firmly links these enzymes to the tautomerase superfamily. These enzymes also represent new uses for the β-α-β-fold, which is the structural motif of tautomerase superfamily members. The fold has been assembled into enzymes that catalyze reactions and use mechanisms previously unknown in this superfamily. These studies provide insight into the progenitors of these enzymes, suggest scenarios for the diversification of enzymatic function and mechanism within the tautomerase superfamily, and have implications for the evolution of metabolic pathways and enzymatic activities in general.

2. *trans*-3-Chloroacrylic acid dehalogenase

CaaD catalyzes the cofactor-independent dehalogenation of *trans*-3-chloroacrylate to yield malonate semialdehyde with a k_{cat}/K_m value of $\sim 1.2 \times 10^5\ M^{-1}\ s^{-1}$ [10,13]. The enzyme is a heterohexamer (~50 kDa), composed of three α-subunits (75 amino acids) and three β-subunits (70 amino acids). Sequence analysis uncovered a relationship between the two subunits of CaaD and members of the 4-oxalocrotonate tautomerase (4-OT) family of enzymes [10]. The 4-OT family is one of the major families of the tautomerase superfamily [14–17]. The members of this superfamily are structurally homologous proteins that share a characteristic β-α-β-fold, as well as a catalytic amino-terminal proline [14,15]. The majority of the known reactions catalyzed by members of this superfamily involves a keto–enol tautomerization of a pyruvoyl moiety (Scheme 2). In addition to 4-OT, these reactions are catalyzed by 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI) and phenylpyruvate tautomerase (PPT), also known as macrophage migration inhibitory factor (MIF) [15,18,19]. The catalytic activity of 4-OT, CHMI, and PPT depends on an unusual property of their Pro-1 residues. In these three enzymes, Pro-1 is able to function as a general base because of a low pK_a of ~6.4, which is due primarily to its location in a hydrophobic active site [15,18,20–23]. Significantly, both the α- and β-subunits of CaaD have an amino-terminal proline, but only βPro-1 has been identified as an essential catalytic group [10,13].

[1] A detailed discussion of the structural and mechanistic studies on DhaA can be found elsewhere [12].



Scheme 2.

It was first suggested by Hartmans et al. [24] that CaaD proceeds via a hydrolytic mechanism and may be distinct from other known hydrolytic dehalogenases. These dehalogenases do not belong to the tautomerase superfamily and they are not able to catalyze a cofactor-independent cleavage of a vinylic carbon–halogen bond. In fact, most hydrolytic dehalogenases, such as haloalkane dehalogenase and L-2-haloacid dehalogenase, only displace halogens bound to sp^3-hybridized carbon atoms, with the notable exception of 4-chlorobenzoyl CoA dehalogenase [25,26]. In the latter case, the aromatic substrate, 4-chlorobenzoate, is activated by conversion to a coenzyme A derivative. All three enzymes catalyze the dehalogenation of their respective substrates in two distinct stages [25,26]. In the first stage, the carboxylate group of an active site aspartate displaces the halide, either via a substitution mechanism (haloalkane dehalogenase and L-2-haloacid dehalogenase) or an addition–elimination mechanism (4-chlorobenzoyl CoA dehalogenase), resulting in the formation of a covalent enzyme–substrate complex. Subsequently, this covalent intermediate is hydrolyzed by water to yield product and the free enzyme.

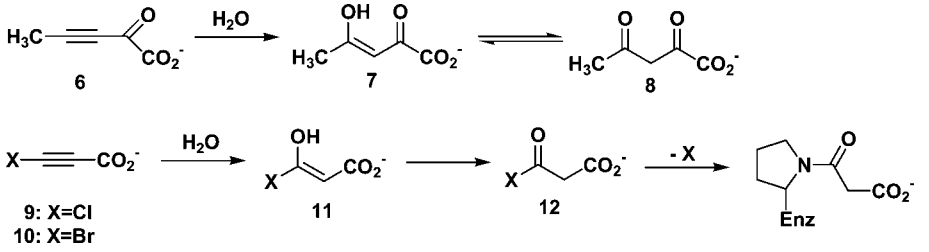
The reaction catalyzed by CaaD is a difficult one. The 3-haloacrylates are stable at room temperature and neutral pH, and non-enzymatic decomposition requires harsher conditions. It has been shown that about 10% of the chloride is removed from 3-chloroacrylate after a 24-h incubation period in 0.5 M aqueous NaOH at 60 °C [27]. A comparison of k_{cat} for CaaD ($3.8\,s^{-1}$) with the pseudo-first order rate constant for the hydroxide-catalyzed dehalogenation of 3-chloroacrylate,

extrapolated to room temperature and pH 9.0, indicates that CaaD accelerates the reaction by a factor of $10^{13.9}$ [28].

With respect to the substrate specificity of the enzyme, only the 3-chloro- and 3-bromoacrylates undergo hydrolytic cleavage. 3-Cyano- and 3-methoxyacrylates are not processed. Structural analogues lacking either the carboxylate group (i.e., the *trans*-isomers of 3-chloroallyl alcohol and 1,3-dichloropropene) or the chlorine substituent (i.e., acrylate) are also not processed. Furthermore, the halogen must be at the C-3 position and the enzyme is highly specific for the *trans*-isomers of 3-haloacrylates [10]. The substrate specificity of the enzyme is readily understood in the context of the structure and mechanism (vide infra).

CaaD could also catalyze one or more reactions unrelated to dechlorination of *trans*-**3**. Such a possibility is plausible in view of the fact that CaaD might represent a recently evolved activity. In this context, several halogenated and non-halogenated unsaturated carboxylic acids were tested as substrates. CaaD was found to catalyze the hydration of 2-oxo-3-pentynoate (**6**, Scheme 3) and 3-chloro- and 3-bromopropiolate (**9** and **10**) in addition to the dehalogenation of *trans*-3-haloacrylates [13]. The halopropiolates are converted into potent irreversible inhibitors of the enzyme upon hydration. Inactivation of CaaD results from the covalent modification of Pro-1 of the β-subunit (Scheme 3). Hydration of 2-oxo-3-pentynoate afforded acetopyruvate (**8**) ($k_{cat}/K_m = 6.4 \times 10^3\ M^{-1}\ s^{-1}$), which does not inactivate the enzyme. Although not physiologically relevant, these hydration reactions are of interest because they are analogous to the hydrolytic dehalogenation of *trans*-3-haloacrylates. Moreover, much insight into the catalytic mechanism of CaaD has come from studies using these acetylene compounds [13].

One striking example is the knowledge gained from the crystal structure of CaaD inactivated by 3-bromopropiolate. This structure identified the active site, suggested interactions that might be responsible for the specific binding of substrate, and led to a working hypothesis for the catalytic mechanism [29]. The enzyme is a barrel-shaped hexamer, which can be viewed as a trimer of heterodimers (Fig. 1A). Each monomer of CaaD displays the β-α-β-fold. The heterodimer is formed by the antiparallel interaction of a two-stranded parallel β-sheet of one monomer with the equivalent β-sheet of the other monomer, forming a four-stranded β-sheet. Most dimer–dimer contacts in the hexamer are mediated by the edges of the β-sheets and are mainly hydrophobic in nature. Additional dimer–dimer interactions are provided by some electrostatic



Scheme 3.

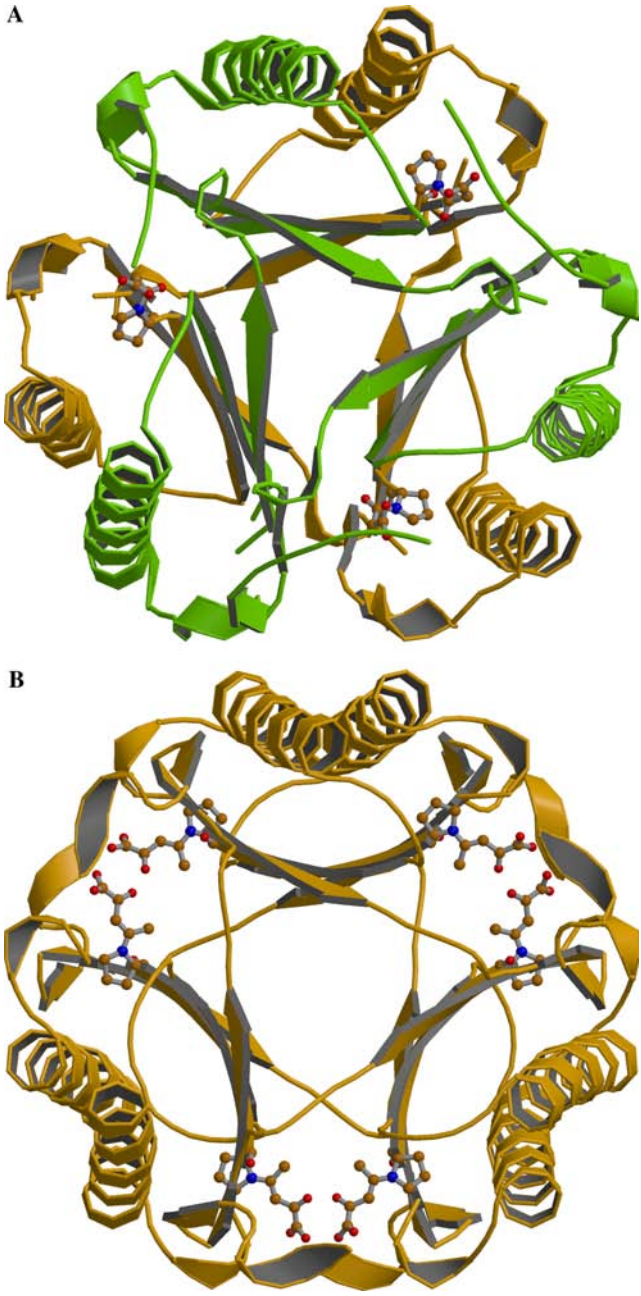


Fig. 1. Ribbon diagrams of the hexameric structures of (A) CaaD and (B) 4-OT. Each diagram shows the catalytically important amino-terminal prolines, modified by either a malonyl group (CaaD) [29], or the adduct resulting from inactivation by 2-oxo-3-pentynoate (4-OT) [30]. For clarity, the α-subunits of CaaD are shown in green, whereas the β-subunits are shown in gold.

interactions between charged residues of each four-stranded β-sheet in the central cavity in the hexamer, and by a β-hairpin at the C-terminus of the α chains. The β-hairpin extends the β-sheet of a neighboring dimer and partly covers the environment of the catalytic βPro-1. As expected, the overall topology of CaaD is similar to that observed in 4-OT (Fig. 1B) [30], further confirming its membership in the tautomerase superfamily.

CaaD contains three active sites related by the threefold rotation symmetry of its $(\alpha\beta)_3$ trimeric structure [29]. The active sites are located at the interface between two (αβ)-heterodimers on one side of the hexamer, each harboring the catalytically important βPro-1. This catalytic proline has several charged residues in its vicinity, among which are αArg-8 and αArg-11 from the same (αβ)-dimer containing the catalytic βPro-1, and αGlu-52 from the same α-chain as αArg-8 and αArg-11 (Fig. 2). In the structure of the inactivated enzyme, βPro-1 forms a covalent bond to the C-3 of a malonyl group, the adduct resulting from the enzyme-catalyzed transformation of 3-bromopropiolate (Scheme 3) [29]. The carboxylate group of the adduct interacts with αArg-8 and αArg-11, suggesting an orientation for substrate in the active site. The charged side chain of αGlu-52 interacts, via a bridging water molecule, with the prolyl nitrogen of βPro-1.

Based on the arrangement of charged residues, a working hypothesis for the mechanism of CaaD was formulated (Scheme 4A) [29]. A key catalytic task for CaaD is to activate a water molecule. This is accomplished by the carboxylate side chain of αGlu-52. The first step in catalysis is the nucleophilic attack of the activated water

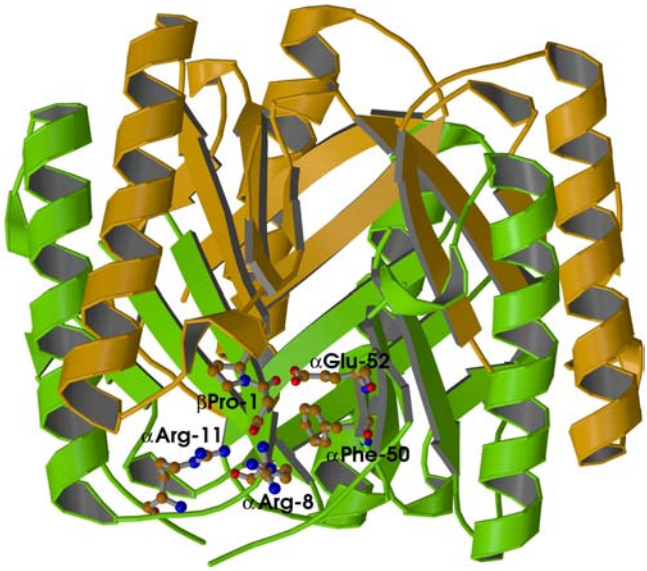
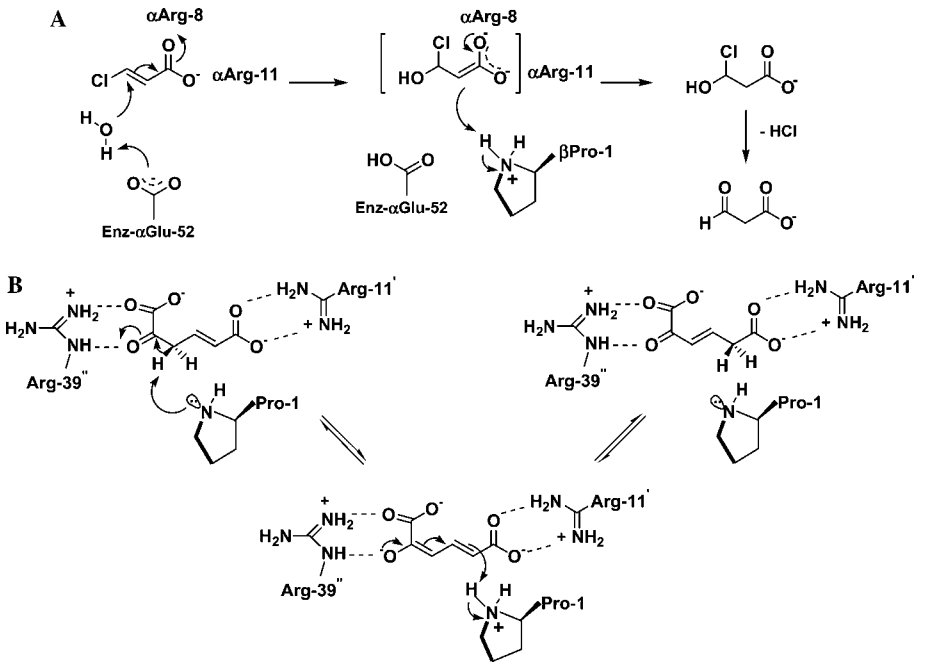


Fig. 2. Ribbon diagram showing a close up of the active site of CaaD at the interface between two (αβ)-heterodimers. βPro-1 forms a covalent bond to the C-3 of a malonyl group, the adduct resulting from the CaaD-catalyzed hydration of 3-bromopropiolate [29]. The carboxylate group of the adduct interacts with αArg-8 and αArg-11 as described in the text. The subunits of CaaD are colored as indicated in Fig. 1A.



Scheme 4.

molecule on C-3 of *trans*-3-chloroacrylate to form an enediolate intermediate. This enediolate is presumably stabilized by interaction with the positively charged side chains of αArg-8 and αArg-11. Ketonization of the enediolate is assisted by βPro-1, which places a proton at the C-2 carbon atom, and generates an unstable chlorohydrin intermediate. Subsequent decomposition yields malonate semialdehyde and HCl [29]. Collapse of the proposed chlorohydrin intermediate could be an enzymatic or a non-enzymatic process.[2]

The proposed substrate binding mode and the catalytic mechanism have been tested by a variety of experiments [28]. The pH dependence of the kinetic parameters for the CaaD-catalyzed reaction provides support for the proposed general acid/base mechanism. The pH rate profile implicates two groups on the free enzyme with pK_a values of 7.6 and 9.2 for optimal activity. ^{15}N NMR titration of the uniformly ^{15}N-labeled CaaD

[2] An alternative mechanism in which the carboxylate group of the active site glutamate (i.e., αGlu-52) displaces the halide via an addition–elimination mechanism and generates a covalent alkenyl–enzyme intermediate has been considered. A single turnover experiment in $H_2{}^{18}O$ is generally used to determine whether such a mechanism is operative. Hydrolysis of the intermediate would release the enzyme and product, and result in the incorporation of the label in the enzyme but not in the product. However, the outcome of this experiment would be ambiguous for the CaaD system due to the rapid formation of the hydrate of malonate semialdehyde, which would produce ^{18}O-labeled product. It should also be noted that this alternate mechanism seems unlikely in view of the fact that CaaD and *cis*-CaaD share no sequence identity with other dehalogenases known to proceed through such a mechanism (i.e., 4-chlorobenzoyl CoA dehalogenases) [26].

and the αP1A-mutant indicated that βPro-1 is responsible for the pK_a value of 9.2. The fact that the catalytically important Pro-1 of CaaD is charged at physiological pH, and not neutral as seen for 4-OT, was an important step in defining the role of this residue in catalysis [28].

The roles of the glutamate and the two arginines in the mechanism were further investigated by analyzing the kinetic properties of six site-specific mutants (αE52Q, αE52D, αR8A, αR8K, αR11A, and αR11K) [13,28,29]. From these studies, it can be concluded that αGlu-52 is essential for catalysis. Replacing this residue with a glutamine or an aspartate results in no detectable activity (i.e., αE52Q) or greatly reduced activity (i.e., αE52D), leading to the conclusion that αGlu-52 is the most likely candidate to activate a water molecule [29]. αArg-8 is also important for catalysis. Both the αR8A and αR8K mutants are not able to process *trans*-3-haloacrylates. However, when analyzed for their ability to hydrate 2-oxo-3-pentynoate (**6**, Scheme 3), there are small effects on K_m and larger decreases in k_{cat}, suggesting an important role for αArg-8 in catalysis.[3] A major part of the loss in activity is likely due to the removal of optimal hydrogen-bonding and/or electrostatic interactions with the enolate intermediate formed at the 2-oxo position of the acetylene substrate, making the nucleophilic attack of an activated water molecule at the C-4 position less favorable (Scheme 3). In contrast, αArg-11 seems to be important for binding and catalysis. Replacement of this residue with an alanine results in a mutant enzyme that has no detectable activity for *trans*-3-haloacrylates and 2-oxo-3-pentynoate [13]. Changing αArg-11 to a lysine produces an active enzyme, but this mutation has pronounced effects on both K_m and k_{cat} for *trans*-3-chloroacrylate and 2-oxo-3-pentynoate, suggesting that αArg-11 plays a significant role in binding the C-1 carboxylate group and in catalysis.[3] The catalytic role of αArg-11 in the hydration mechanism can be ascribed to its ability to act as an electron sink (Scheme 4A). Presumably, αArg-11, together with αArg-8, polarizes the unsaturated acid and thereby facilitates the nucleophilic attack of water at the C-3 position [28,29].

The two arginine residues critical for the activity of CaaD, αArg-8 and αArg-11, were also studied by a series of ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) NMR experiments, using a competitive inhibitor, 3-chloro-2-butenoic acid (3-CBA) [28]. In the presence of a saturating amount of 3-CBA, the NϵH signals for αArg-8 and αArg-11 are observed. In the absence of 3-CBA, these signals (and the NϵH resonances for the other arginine residues) disappear due to base-catalyzed NϵH exchange. The selective protection of the N–H resonances for αArg-8 and αArg-11 in the presence of 3-CBA reflects an interaction between αArg-8 and αArg-11 and 3-CBA. Extrapolation of this observation to substrate binding supports the proposed roles for αArg-8 and αArg-11 in substrate binding and in stabilizing the enediolate intermediate in a Michael addition mechanism [28]. Thus, the NMR studies fully support the working hypothesis for the mechanism of CaaD, gleaned from the crystal structure, as shown in Scheme 4A.

[3] G.J. Poelarends, H. Serrano, W.H. Johnson, Jr., C.P. Whitman, unpublished data (2003).

In the CaaD mechanism, αArg-11 of the Pro/Arg dyad is critical for binding and polarization of the substrate's carboxylate group. This catalytic dyad is found in most, but not all, superfamily members known thus far. The carboxylate base (αGlu-52) and the arginine that assists in stabilizing the *aci*-carboxylate intermediate (αArg-8), as well as the altered catalytic function of Pro-1, are specific adaptations of the superfamily active site template to allow efficient catalysis of a hydration reaction. For comparative purposes, Scheme 4B shows the mechanism of 4-OT, which utilizes the catalytic Pro/Arg dyad and a differently positioned second arginine (Arg-39), to catalyze a proton transfer reaction [15].[4] These two reactions are fascinating examples of how an active site with a particular capability (in this case, binding and polarization of a carboxylate group) can be tailored by altering the catalytic function of a conserved catalytic group (i.e., Pro-1), adding a new catalytic group (i.e., αGlu-52), and repositioning an existing catalytic group (i.e., Arg-8), to allow efficient catalysis of very different transformations.

3. *cis*-3-Chloroacrylic acid dehalogenase

cis-CaaD, like CaaD, catalyzes the hydration of the 3-haloacrylates (in this case the *cis*-isomers), 2-oxo-3-pentynoate, and the 3-halopropiolates (Schemes 1 and 3) [17]. However, its specificity is quite different, which makes a detailed study highly informative. The metabolic function of *cis*-CaaD is to convert *cis*-3-chloroacrylate to malonate semialdehyde, and it does so with a catalytic efficiency (as assessed by the k_{cat}/K_m value of $\sim 3.4 \times 10^5\,M^{-1}\,s^{-1}$) comparable to that of CaaD [17]. The further degradation of malonate semialdehyde in various 3-chloroacrylic acid-utilizing organisms might proceed by two different routes. In several bacteria, such as *P. pavonaceae* 170 and *Pseudomonas cepacia* CAA1, the cofactor-independent decarboxylation to produce acetaldehyde occurs [11,24]. In coryneform bacterium strain FG41, genetic evidence suggests that a coenzyme A- and NAD^+-dependent malonate semialdehyde dehydrogenase is involved in the conversion of malonate semialdehyde to acetyl CoA.[3] The latter pathway for malonate semialdehyde degradation may be similar to those involved in the catabolism of 2-methylmalonate semialdehyde in microbes and higher animals [31,32].

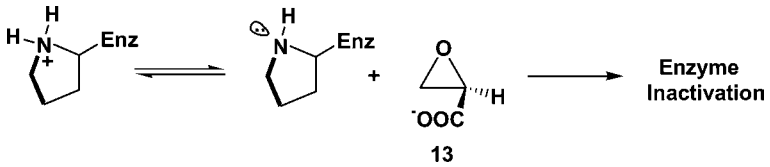
The gene encoding the *cis*-CaaD enzyme from coryneform bacterium strain FG41 was recently cloned and sequenced [17]. The enzyme functions as a homotrimer (~50 kDa), where each subunit is composed of 149 amino acids. As expected on the basis of the presumed structural and mechanistic relationship to CaaD, sequence analysis of *cis*-CaaD revealed that this enzyme is a tautomerase superfamily member. However, the sequence identity between CaaD and *cis*-CaaD is low (~20%). Despite their likely independent origins, CaaD and *cis*-CaaD use similar catalytic strategies to process different isomers of 3-haloacrylates.

[4] The unprimed, primed, and doubly primed residues in Scheme 4B indicate that these residues come from different subunits of the 4-OT homohexamer. Pro-1 and Arg-11′ are within the same dimer but from different monomers. Arg-39″ is from an adjacent dimer.

In the absence of a crystal structure, much of the catalytic mechanism for *cis*-CaaD has been deduced from mutagenesis and chemical studies [17,33]. A plot of the catalytic rate versus pH is bell-shaped, like that of CaaD, and provides support for a general acid/base mechanism [33]. Two groups on *cis*-CaaD with pK_a values of 7.6 and 9.3 are necessary for optimal activity. On the basis of the presumed mechanistic analogy to CaaD, it was proposed that the pK_a of 9.3 corresponds to Pro-1, leading to the tentative conclusion that this residue is the critical general acid catalyst that protonates the C-2 position of the substrate [33]. Further support comes from mutagenesis and chemical modification studies, which indicate that Pro-1 is critical for *cis*-CaaD activity [17]. Substitution of Pro-1 by alanine produces an inactive enzyme. Reaction of *cis*-CaaD with 3-chloropropiolate (**9**), a mechanism-based inhibitor, led to the alkylation of Pro-1 and the concomitant loss of catalytic activity (Scheme 3). The presence of substrate protected *cis*-CaaD from modification and inactivation by 3-chloropropiolate, implicating Pro-1 as an active site residue [17].

The importance of Pro-1 in catalysis was further demonstrated by affinity labeling with (*R*)-oxirane-2-carboxylate (**13**, Scheme 5) [33]. The carboxylate side chain of (*R*)-**13** enables it to bind at the active site of *cis*-CaaD, and this specific binding mode results in the covalent modification of Pro-1. Pro-1 was again determined to be an active site residue by two lines of evidence. First, the affinity-labeling reaction was stereospecific, as the (*S*)-enantiomer of oxirane-2-carboxylate does not alkylate the enzyme. Second, the rate of inactivation was impeded by the presence of *cis*-3-chloroacrylate [33].

Besides the presence of a catalytically important N-terminal proline with a pK_a of ~9.3, the mechanistic relationship between CaaD and *cis*-CaaD is apparent from a common functional sequence motif (Fig. 3). This motif is found in a small set of proteins including *cis*-CaaD and three homologues of unknown function, as well as



13

Scheme 5.

```
CaaD1      1 PMISCDMRYGRTDEQKRALSAGLLRVISEATGEPRE----NIFFVIREGSGINFVEHGEHLPDYVPGNAND-  67
cis-CaaD  63 IFVHGLHREGRSADLKGQLAQRIVDDVSVAAEIDRK----HIWVYFGEMPAQQMVEYGRFLPQPGHEGEWF- 129
Cg10062   63 IWVQATIRSGRTEKQKEELLLRLTQEIALILGIPNE----EVWVYITEIPGSNMTEYGRLLMEPGEEEKWF- 129
OrfX      61 LIINGWVRTGHSDEQTTALVTQVADAATRITGIPAE----RVLVIIGNSPARFAIEGGRILPDPGQELAWL- 127
OrfY      61 SFLGGQIRHGRSVETRQAMLKALRDMWVQTTGQSEA----ELIVGISEVDPRMVLEAGFFMPEPGQEKAWF- 127
```

Fig. 3. Alignment of members of the *cis*-CaaD family with the α-subunit of CaaD. The sequences are those of *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 *Mycobacterium smegmatis*, and CaaD1 corresponding to the α-subunit of CaaD from *P. pavonaceae* 170 (GI: 10637969). Highlighted in grey are arginine residues corresponding to αArg-8 and αArg-11 in CaaD, which are responsible for binding and polarization of the substrate's carboxylate group, and glutamate residues corresponding to αGlu-52 in CaaD, which is the presumed general base that activates a water molecule.

CaaD [17]. Highlighted in grey are the two arginine residues corresponding to αArg-8 and αArg-11 of CaaD, which are involved in binding and polarization of the substrate's carboxylate group, and the glutamate corresponding to αGlu-52 of CaaD, which is the critical active site general base that activates the nucleophilic water molecule. The conservation of these residues in this set of proteins is consistent with the hypothesis that they may all function in a manner similar to that of CaaD. Indeed, mutational analysis of this motif in *cis*-CaaD revealed that Arg-70, Arg-73, and Glu-114 are critical for catalysis [17]. Although the protein sequence databases contain many 4-OT-like sequences, only these five proteins share the characteristic glutamate and the two arginines, suggesting that this variation in the tautomerase superfamily is rare. As indicated above, this variation appears to represent a specific adaptation of the superfamily active site template to catalyze the conjugate addition of water.

The recent characterization of a 149-amino acid protein from *Corynebacterium glutamicum*, designated Cg10062, which is the closest relative of *cis*-CaaD (Fig. 3) lends further support to this hypothesis.[3] This protein catalyzes the conversion of 2-oxo-3-pentynoate (**6**) to acetopyruvate (**8**) ($k_{cat}/K_m = 5.2 \times 10^1\ M^{-1}\ s^{-1}$), consistent with an enzyme-catalyzed hydration reaction as established for CaaD and *cis*-CaaD (Scheme 3). This gram-positive bacterium is not known to degrade 3-chloroacrylate, which raises the interesting possibility that the Cg10062 protein may be able to function as an "accidental" 3-chloroacrylic acid dehalogenase by catalyzing the conjugate addition of water to the double bond. Indeed, this enzyme dehalogenates 3-chloroacrylates, with a notable preference for the *cis*-isomer (*cis*-isomer: $k_{cat}/K_m = 3.2 \times 10^1\ M^{-1}\ s^{-1}$; *trans*-isomer: $k_{cat}/K_m = 2.5 \times 10^{-1}\ M^{-1}\ s^{-1}$).[3] These observations suggest that the dehalogenation of 3-haloacrylates by CaaD and *cis*-CaaD is a hydration reaction, in which the departure of the halide does not require a halide binding pocket but might be a fortuitous side reaction resulting from the very rapid decay of the unstable halohydrin intermediate.

The combined results suggest a mechanism for *cis*-CaaD, which largely parallels that of CaaD (Scheme 4A) [17]. Accordingly, Glu-114 functions as the water-activating base while Pro-1, the general acid, places a proton at the C-2 position of *cis*-3-chloroacrylate. The two arginine residues (Arg-70 and Arg-73) interact with the carboxylate group of the substrate, which would assist the presumed Michael addition of water. There are also subtle mechanistic differences between the two enzymes. One difference involves the activation of the water molecule. For *cis*-CaaD, this process probably involves additional residues because the k_{cat}/K_m value determined for the E114Q mutant is only eightfold less than that measured for wild-type [17]. In CaaD, αGlu-52 appears to be the major player responsible for the activation of water because the αE52Q mutant is inactive (within the limits of detection) [29]. Insight into the structural basis for these mechanistic differences, as well as the differences in isomer specificity, awaits the determination of the three-dimensional structure of *cis*-CaaD.

4. Malonate semialdehyde decarboxylase

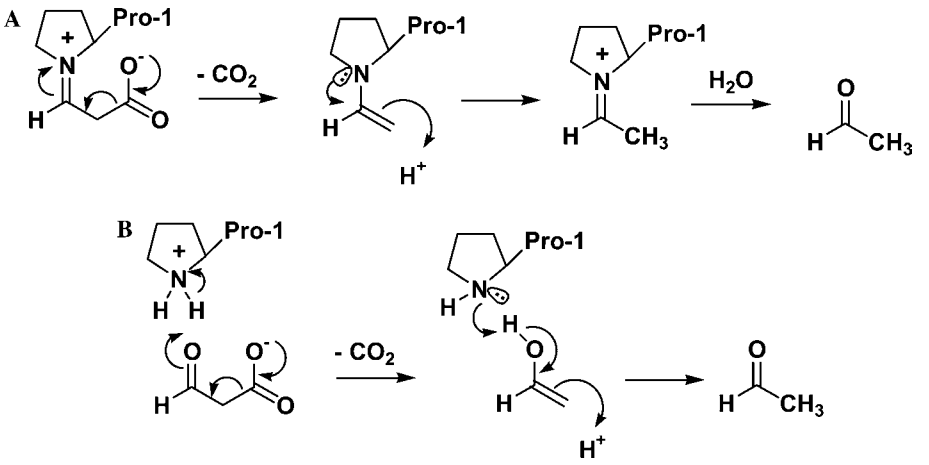
MSAD, the last enzyme in the degradation pathway of 1,3-dichloropropene in *P. pavonaceae* 170, catalyzes the decarboxylation of malonate semialdehyde (**4**) to yield

acetaldehyde (**5**, Scheme 1) [11]. Two lines of evidence suggest that this is the physiological function of the enzyme. First, the gene encoding MSAD is located directly downstream of the genes encoding the two subunits of CaaD thereby linking the two genes of the pathway in a functional and physical sense [11]. Second, MSAD is remarkably efficient in decarboxylating malonate semialdehyde ($k_{cat}/K_m = 2.2 \times 10^7\,M^{-1}\,s^{-1}$) [34]. The mature protein likely functions as a homotrimer (~42 kDa), where each subunit is composed of 129 amino acids. The enzyme is not likely metal dependent because the presence of divalent metal salts or metal chelators does not affect enzyme activity [11].

Sequence analysis of MSAD shows that this enzyme shares an important feature with CaaD and places MSAD in the tautomerase superfamily [11]. This unexpected relationship was initially suggested by the presence of two conserved motifs, (Pro^1–$(X)_9$–Arg^{11}–X–Asp^{13}–X–Gln^{15}) and (Arg^{75}–X-Glu^{77}–X–Gln^{79}), in the N- and C-terminal region of MSAD, respectively. Although MSAD has a low level of overall sequence identity with functionally characterized members of the tautomerase superfamily, the enzyme contains an N-terminal proline that is involved in the decarboxylation reaction. The importance of Pro-1 in catalysis was shown by site-directed mutagenesis and chemical modification experiments [11]. Substitution of Pro-1 by an alanine afforded a mutant, which has only ~1% of the activity observed for wild-type. Reaction of MSAD with malonate semialdehyde in the presence of sodium cyanoborohydride, a reducing agent, led to the alkylation of Pro-1 and the concomitant loss of catalytic activity [11].

The roles of the two arginines (Arg-11 and Arg-75) in the mechanism were investigated by analyzing the kinetic properties of their alanine mutants [11]. Mutation of Arg-11 to an alanine had little effect on enzyme activity, indicating that it does not play a role in the mechanism. In contrast, substitution of Arg-75 by an alanine resulted in a mutant, which has only ~0.2% of the activity observed for wild-type. Hence, Arg-75 of MSAD may interact with the carboxylate group of the substrate, analogous to the roles of the active-site arginines in CaaD and *cis*-CaaD. In addition to assisting in substrate binding, this interaction may place the carboxylate group in a favorable orientation for decarboxylation [35]. Like CaaD and *cis*-CaaD, MSAD retains the superfamily catalytic Pro/Arg dyad although in this case, the dyad consists of Pro-1 and Arg-75.

These combined observations place MSAD in the tautomerase superfamily and point to important roles for the essential Pro/Arg dyad in the mechanism. Two very different catalytic mechanisms have been proposed on the basis of these results [11,34]. At first glance, the results of the chemical modification studies and the apparent non-metal dependent activity seem to point to covalent catalysis involving a Schiff base intermediate (Scheme 6A). To be sufficiently nucleophilic at cellular pH to react with the C-3 carbonyl group, Pro-1 is expected to have a lowered pK_a value such as that observed in 4-OT. Surprisingly, however, the pK_a of Pro-1 was found to be ~9.2 by direct pH titration of the uniformly ^{15}N-labeled enzyme using ^{15}N NMR spectroscopy [34]. Although this finding does not rule out a Schiff base mechanism, it suggests an alternative mechanism in which the C-3 keto group of malonate semialdehyde is polarized by hydrogen bonding and/or electrostatic interactions (Scheme



Scheme 6.

6B), analogous to the mechanism proposed for methylmalonyl CoA decarboxylase [35]. In this mechanism, Pro-1 would function as a general acid catalyst rather than a base. Crystallographic and mechanistic studies are currently underway to delineate the roles of Pro-1 and Arg-75 in the mechanism as well as to identify other mechanistic residues.

A recent study of the substrate and reaction specificity of MSAD shows that this enzyme shares a second feature with CaaD and *cis*-CaaD: it is able to catalyze the hydration of 2-oxo-3-pentynoate (**6**) to yield acetopyruvate (**8**, Scheme 3). The k_{cat}/K_m value for this reaction is $6.0 \times 10^2\ M^{-1}\ s^{-1}$, which is only ~10-fold lower than that measured for CaaD [34]. A likely mechanism for this reaction involves the initial Michael addition of water to the triple bond of 2-oxo-3-pentynoate to form an allenic species [34]. Rearrangement of this species produces an enol intermediate (**7**), which readily ketonizes to acetopyruvate (**8**, Scheme 3).

To investigate the importance of the Pro/Arg dyad to this unexpected hydratase activity of MSAD, the P1A, and R75A mutants were assayed for their ability to convert 2-oxo-3-pentynoate to acetopyruvate. It was found that substitution of these residues with an alanine essentially abolished this enzymatic activity [34]. These findings are consistent with the same active site catalyzing both the decarboxylation and hydration reactions, and show that Pro-1 and Arg-75 are critical for the hydratase activity. The ability of MSAD to catalyze a hydration reaction is further substantiated by the finding that this enzyme is irreversibly inhibited by 3-chloro- and 3-bromopropiolate (**9** and **10**) [34]. Both halopropiolates modify the N-terminal proline as shown in Scheme 3. The presence of 2-oxo-3-pentynoate (**6**) protects MSAD from modification and inactivation by 3-chloropropiolate. This result provides additional evidence for the active site nature of the hydration reaction.

The sum of these observations shows that MSAD is functionally promiscuous. In addition to its physiological decarboxylase activity, it functions as a hydratase when processing 2-oxo-3-pentynoate. The roles for the Pro/Arg dyad of MSAD in this

mechanism may be analogous to the roles of the complementary Pro/Arg dyad in the CaaD- and *cis*-CaaD-catalyzed hydration of 2-oxo-3-pentynoate [13,17]. However, in the absence of a three-dimensional structure, it is not yet known which additional catalytic groups are involved in the hydratase activity of this promiscuous enzyme.

5. Evolutionary implications

The tautomerase superfamily was initially named based on the functions of the prototypical members of the superfamily (e.g., 4-OT, CHMI, and PPT), which all catalyze a keto–enol tautomerization of a pyruvoyl moiety (Scheme 2) [14]. With the discovery of CaaD, *cis*-CaaD, and MSAD, there is now considerable functional, structural, and mechanistic diversity in the tautomerase superfamily (Scheme 1). On the basis of sequence identity and conserved active-site residues, the superfamily can be grouped into five different families represented by 4-OT, CHMI, MIF, MSAD, and *cis*-CaaD [17]. In the 4-OT family, the monomer encodes a single β-α-β motif while in the CHMI and MIF families, the monomers are nearly twice as long and consist of two β-α-β motifs, which resemble the 4-OT dimer [14,18,30,36]. Although structures are not yet available for MSAD and *cis*-CaaD, sequence analysis suggests that their monomers (129 and 149 amino acids, respectively) will also consist of two β-α-β motifs [11,17]. Each ancestor of the latter four families probably resulted from an independent duplication/fusion of a 4-OT-like sequence [17]. Although these proteins are made up of monomers with a common structural fold, they differ in quaternary structure. CHMI, MIF, MSAD, and *cis*-CaaD function as homotrimers, whereas 4-OT and CaaD function as homo- and heterohexamers, respectively [11,17,18,29,36]. The discovery of CaaD, *cis*-CaaD, and MSAD, along with the identification of two new families in the tautomerase superfamily, demonstrates the remarkable versatility of the β-α-β motif for the creation of new enzymatic activities and structures.

Based on the examples discussed here, some conclusions regarding the relationship between structural features and enzymatic function can be made. The active sites in the superfamily are designed to bind and polarize a carboxylate group of the substrate. The primary catalytic activity that is supported by this feature appears to be the protonation of the carbon adjacent to the carboxylate group, involving the general superfamily catalytic Pro/Arg dyad. In CaaD, for example, Arg-11 serves as an electron sink to draw electron density away from C-3 (resulting in activation of this carbon for nucleophilic attack) and toward C-2, thereby facilitating protonation at C-2 by Pro-1 (Scheme 4A) [28,29]. In 4-OT, the Arg-11 may function similarly and draw electron density toward C-5 for protonation by Pro-1 (Scheme 4B) [15]. This phenomenon is also observed in MSAD as part of the hydratase activity further supporting the central mechanistic role of the Pro/Arg dyad [34]. The range of catalytic abilities found in the superfamily is made possible by alterations in the binding pocket and variations in function and position of (additional) catalytic groups. The presence of the catalytic Pro/Arg dyad and the ability to carry out a common chemical step provides compelling evidence that the tautomerase superfamily represents divergent evolution from a common progenitor.

Since 1,3-dichloropropene is a xenobiotic compound, it is interesting to consider when and how the three 1,3-dichloropropene catabolic enzymes arose. Without knowledge of their specific precursor enzymes, it is impossible to reconstruct the evolutionary history of these proteins. It is conceivable, however, that a 4-OT-like precursor could have accommodated the 3-chloroacrylate or malonate semialdehyde substrate and used the conserved Pro/Arg dyad to promote the dehalogenase or decarboxylase reaction. Indeed, the closest relatives of CaaD, 4-OT from *P. putida* mt-2 and its homologue YwhB from *Bacillus subtilis*, have rudimentary *trans*-3-chloroacrylate dehalogenase activities, in addition to their major tautomerase activities [37]. Moreover, 4-OT appears to have a low-level malonate semialdehyde decarboxylase activity.[5] The closest relative of *cis*-CaaD, the *C. glutamicum* protein Cg10062, has a dehalogenating activity towards *cis*-3-chloroacrylate.[3]

The sum of these observations suggests that some 4-OT-like enzymes could fortuitously have favored a rudimentary dehalogenase and/or decarboxylase activity in addition to a primary tautomerase activity. Nature may have recruited the low-level promiscuous activities of such existing enzymes to provide a springboard for the evolution of new dehalogenating or decarboxylating enzymes. It would be interesting to know whether this recruitment and adaptation process occurred recently in response to the anthropogenic introduction of 1,3-dichloropropenes into the environment, or whether it occurred a long time ago in response to some unidentified halogenated natural product. Clearly, it can be argued that the low sequence identities of CaaD, *cis*-CaaD, and MSAD to the currently known tautomerase superfamily members indicate that they diverged a long time ago. However, the enormous enzyme diversity that resides in bacteria in the environment leaves open the possibility that more closely related superfamily members exist that have not yet been isolated.

Acknowledgments

We are grateful to our numerous co-workers whose names are given in the references. In particular, we express our gratitude to Dr. Jeffrey J. Almrud for producing the structure figures and his critical review of the manuscript.

References

[1] R.S. Dungan, S.R. Yates, Vadose Zone J. 2 (2003) 279–286.
[2] H. van Dijk, Pest. Sci. 11 (1980) 625–632.
[3] G. Lebbink, B. Proper, A. Nipshagen, Acta Hortic. 255 (1989) 361–371.
[4] C. Verhagen, E. Smit, D.B. Janssen, J.D. van Elsas, Soil Biol. Biochem. 27 (1995) 1547–1557.
[5] G.J. Poelarends, M. Wilkens, M.J. Larkin, J.D. van Elsas, D.B. Janssen, Appl. Environ. Microbiol. 64 (1998) 2931–2936.
[6] L.-T. Ou, J.E. Thomas, K.-Y. Chung, A.V. Ogram, Biodegradation 12 (2001) 39–47.

[5] W.H. Johnson, Jr., G.J. Poelarends, C.P. Whitman, unpublished data (2004).

[7] G.J. Poelarends, L.A. Kulakov, M.J. Larkin, J.E.T. van Hylckama Vlieg, D.B. Janssen, J. Bacteriol. 182 (2000) 2191–2199.
[8] G.J. Poelarends, M. Zandstra, T. Bosma, L.A. Kulakov, M.J. Larkin, J.R. Marchesi, A.J. Weightman, D.B. Janssen, J. Bacteriol. 182 (2000) 2725–2731.
[9] G.J. Poelarends, Ph.D thesis, University of Groningen, The Netherlands, 2001.
[10] G.J. Poelarends, R. Saunier, D.B. Janssen, J. Bacteriol. 183 (2001) 4269–4277.
[11] G.J. Poelarends, W.H. Johnson Jr., A.G. Murzin, C.P. Whitman, J. Biol. Chem. 278 (2003) 48674–48683.
[12] D.B. Janssen, Curr. Opin. Chem. Biol. 8 (2004) 150–159.
[13] S.C. Wang, M.D. Person, W.H. Johnson Jr., C.P. Whitman, Biochemistry 42 (2003) 8762–8773.
[14] A.G. Murzin, Curr. Opin. Struct. Biol. 6 (1996) 386–394.
[15] C.P. Whitman, Arch. Biochem. Biophys. 402 (2002) 1–13.
[16] J.J. Almrud, A.D. Kern, S.C. Wang, R.M. Czerwinski, W.H. Johnson Jr., A.G. Murzin, M.L. Hackert, C.P. Whitman, Biochemistry 41 (2002) 12010–12024.
[17] G.J. Poelarends, H. Serrano, M.D. Person, W.H. Johnson Jr., A.G. Murzin, C.P. Whitman, Biochemistry 43 (2004) 759–772.
[18] H.S. Subramanya, D.I. Roper, Z. Dauter, E.J. Dodson, G.J. Davies, K.S. Wilson, D.B. Wigley, Biochemistry 35 (1996) 792–802.
[19] E. Rosengren, P. Aman, S. Thelin, C. Hansson, S. Alhfors, P. Bjork, L. Jacobsson, H. Rorsman, FEBS Lett. 417 (1997) 85–88.
[20] J.T. Stivers, C. Abeygunawardana, A.S. Mildvan, G. Hajipour, C.P. Whitman, Biochemistry 35 (1996) 814–823.
[21] M. Swope, H.-W. Sun, P.R. Blake, E. Lolis, EMBO J. 17 (1998) 3534–3541.
[22] S.L. Stamps, M.C. Fitzgerald, C.P. Whitman, Biochemistry 37 (1998) 10195–10202.
[23] R.M. Czerwinski, T.K. Harris, M.A. Massiah, A.S. Mildvan, C.P. Whitman, Biochemistry 40 (2001) 1984–1995.
[24] S. Hartmans, M.W. Jansen, M.J. van der Werf, J.A.M. de Bont, J. Gen. Microbiol. 137 (1991) 2025–2032.
[25] D.B. Janssen, J.E. Oppentocht, G.J. Poelarends, Curr. Opin. Biotechnol. 12 (2001) 254–258.
[26] R.M. de Jong, B.W. Dijkstra, Curr. Opin. Struct. Biol. 13 (2003) 1–9.
[27] S.A. Braddon, C.N. Dence, Tappi 51 (1968) 249–256.
[28] H.F. Azurmendi, S.C. Wang, M.A. Massiah, G.J. Poelarends, C.P. Whitman, A.S. Mildvan, Biochemistry 43 (2004) 4082–4091.
[29] R.M. de Jong, W. Brugman, G.J. Poelarends, C.P. Whitman, B.W. Dijkstra, J. Biol. Chem. 279 (2004) 11546–11552.
[30] A.B. Taylor, R.M. Czerwinski, W.H. Johnson Jr., C.P. Whitman, M.L. Hackert, Biochemistry 37 (1998) 14692–14700.
[31] Y.-X. Zhang, L. Tang, C.R. Hutchinson, J. Bacteriol. 178 (1996) 490–495.
[32] N.Y. Kedishvilli, K.M. Popov, P.M. Rougraff, Y. Zhao, D.W. Crabb, R.A. Harris, J. Biol. Chem. 267 (1992) 19724–19729.
[33] G.J. Poelarends, H. Serrano, W.H. Johnson, Jr., C.P. Whitman, Biochemistry 43 (2004) 7187–7196.
[34] G.J. Poelarends, H. Serrano, W.H. Johnson, Jr., D.W. Hoffman, C.P. Whitman (in preparation).
[35] M.M. Benning, T. Haller, J.A. Gerlt, H.M. Holden, Biochemistry 39 (2000) 4630–4639.
[36] M. Suzuki, H. Sugimoto, A. Nakagawa, I. Tanaka, J. Nishihira, M. Sakai, Nat. Struct. Biol. 3 (1996) 259–266.
[37] S.C. Wang, W.H. Johnson Jr., C.P. Whitman, J. Am. Chem. Soc. 125 (2003) 14282–14283.