Cell. Mol. Life Sci. 65 (2008) 3606 – 3618
1420-682X/08/223606-13
DOI 10.1007/s00018-008-8285-x

Cellular and Molecular Life Sciences

Review

The chemical versatility of the β−α−β fold: Catalytic promiscuity and divergent evolution in the tautomerase superfamily

G. J. Poelarends[a] , V. Puthan Veetil[a] and C. P. Whitman[b, *]

[a] Department of Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen (Netherlands)
[b] Division of Medicinal Chemistry, College of Pharmacy, The University of Texas, Austin, TX 78712 – 1074 (USA), Fax: +512-232-2606, e-mail: whitman@mail.utexas.edu

Received 22 May 2008; received after revision 20 June 2008; accepted 02 July 2008
Online First 11 August 2008

Abstract. Tautomerase superfamily members have an amino-terminal proline and a β−α−β fold, and include 4-oxalocrotonate tautomerase (4-OT), 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI), *trans*- and *cis*-3-chloroacrylic acid dehalogenase (CaaD and *cis*-CaaD, respectively), malonate semialdehyde decarboxylase (MSAD), and macrophage migration inhibitory factor (MIF), which exhibits a phenylpyruvate tautomerase (PPT) activity. Pro-1 is a base (4-OT, CHMI, the PPT activity of MIF) or an acid (CaaD, *cis*-CaaD, MSAD). Components of the catalytic machinery have been identified and mechanistic hypotheses formulated. Characterization of new homologues shows that these mechanisms are incomplete. 4-OT, CaaD, *cis*-CaaD, and MSAD also have promiscuous activities with a hydratase activity in CaaD, *cis*-CaaD, and MSAD, PPT activity in CaaD and *cis*-CaaD, and CaaD and *cis*-CaaD activities in 4-OT. The shared promiscuous activities provide evidence for divergent evolution from a common ancestor, give hints about mechanistic relationships, and implicate catalytic promiscuity in the emergence of new enzymes.

Keywords. Tautomerase superfamily, divergent evolution, catalytic promiscuity.

Introduction

The tautomerase superfamily is a group of structurally homologous proteins that are characterized by a β−α−β building block and a catalytic amino-terminal proline (Pro-1) [1–3]. The signature β−α−β fold, as first reported in 4-oxalocrotonate tautomerase (4-OT), begins with Pro-1 at the start of a β-strand (β1), followed by an α-helix (α1) and a 3_{10} helix, which precedes a second parallel β-strand (β2). The fold ends with a β-hairpin near the C-terminus, which is important for the formation of the hexamer (Fig. 1) [2]. Thus far, the superfamily consists of five families represented by 4-OT [1], 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI) [4, 5], macrophage migration inhibitory factor (MIF) [6–9], *cis*-3-chloroacrylic acid dehalogenase (*cis*-CaaD) [10], and malonate semialdehyde decarboxylase (MSAD) [11]. The 4-OT family includes *trans*-3-chloroacrylic acid dehalogenase (CaaD) [3, 12]. 4-OT from *Pseudomonas putida* mt-2 and CHMI from *Escherichia coli* C

* Corresponding author.

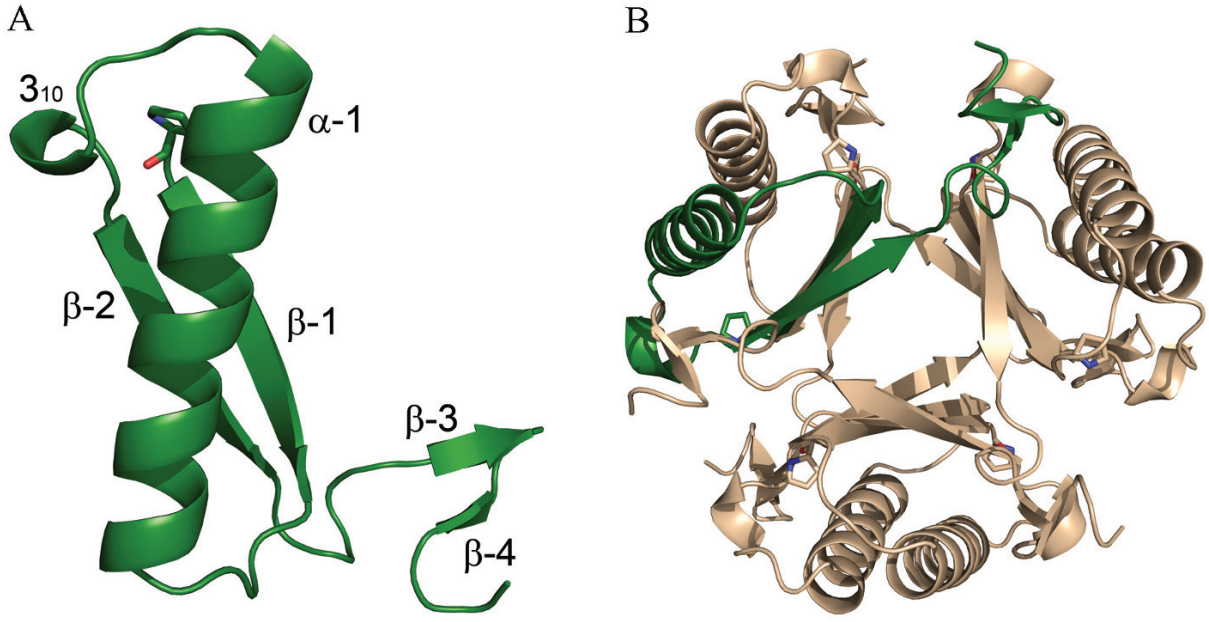
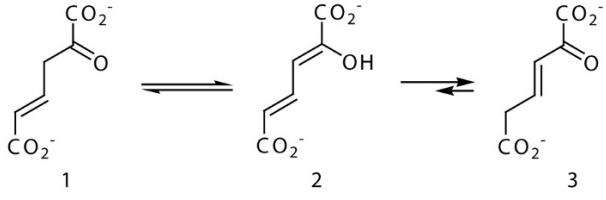


Figure 1. (*A*) The characteristic β-α-β building block of the tautomerase superfamily, as represented by the 4-OT monomer [2]. The key structural elements are labeled and discussed in the text. (*B*) The β-α-β building block highlighted in the 4-OT hexamer [2]. All Figures were prepared with PyMOL [57].

function as tautomerases in degradation pathways for aromatic hydrocarbons and aromatic amino acids, respectively [4, 13, 14]. MIF is a pro-inflammatory cytokine, but also functions as a phenylpyruvate tautomerase (PPT) [8, 9]. CaaD and *cis*-CaaD precede MSAD in a degradation pathway for an isomeric mixture of 1,3-dichloropropene, which is used as a nematocide [3]. CaaD, *cis*-CaaD, and MSAD are found in various gram-positive and gram-negative bacteria including *P. pavonaceae* 170 and the coryneform bacterial strain FG41 [15–17]. The characterized members in the superfamily show rich mechanistic and structural diversity.

In the tautomerase superfamily, Pro-1 can function as a general acid catalyst or a general base catalyst, depending on its protonation state [3]. 4-OT and PPT are two examples where Pro-1 functions as a general base catalyst. The proline functions as a general base because it has a pK_a of ~6.4 (4-OT) [2] or ~6.0 (PPT) [9], which means that it exists largely in the unprotonated state at cellular pH.

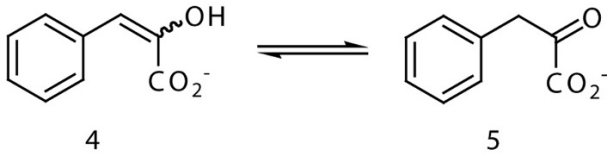
4-OT is found in a bacterial pathway that degrades aromatic hydrocarbons [13, 14]. It exists as a hexamer where each monomer consists of 62 amino acids [18–21]. Initially, it was thought that 4-OT catalyzed a 1,3-allylic rearrangement of 2-oxo-4-hexenedioate (**1**, Scheme 1) to 2-oxo-3-hexenedioate (**3**) through the dienol intermediate known commonly as 2-hydroxymuconate (**2**) [18, 22, 23]. Recent stereochemical findings suggest that 4-OT more likely catalyzes a simple 1,5-keto-enol tautomerization of **2** to **3** where Pro-1 functions as a general base and abstracts the 2-hydroxyl proton and delivers it stereoselectively to the C-5 position [19, 24]. CHMI, a trimer where each monomer is made up of 125 amino acids, functions similarly to 4-OT but uses 5-(carboxymethyl)-2-hydroxymuconate as its substrate [5, 19, 21]. The structural and biological properties of CHMI have been previously reviewed [5].



Scheme 1.

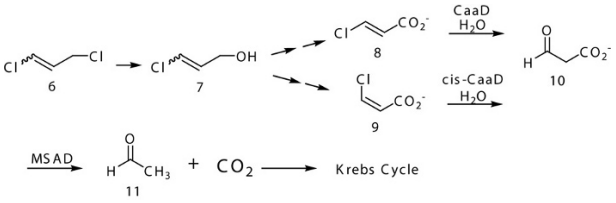
Pro-1 also functions as a general base catalyst in the PPT-catalyzed reaction. PPT converts phenylenolpyruvate (**4**, Scheme 2) to phenylpyruvate (**5,** Scheme 2) [8, 9]. The physiological significance of this reaction is not known, but PPT is also known as macrophage migration inhibitory factor (MIF) [8, 9]. MIF functions primarily as a pro-inflammatory cytokine, but the PPT activity is not believed to play a role in the cytokine activity. PPT is a trimer where each monomer

has 114 amino acids [6, 7]. The monomers are roughly twice as large as those of 4-OT.



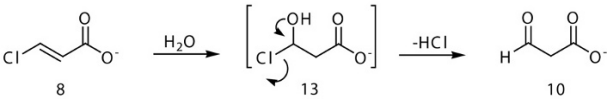
Scheme 2.

The discovery of the 1,3-dichloropropene catabolic pathway (Scheme 3) and the subsequent mechanistic and structural characterization of three enzymes in this pathway showed the versatility of Pro-1 in this superfamily [3]. In CaaD, *cis*-CaaD, and MSAD, Pro-1 acts as a general acid catalyst [3]. In these three enzymes, Pro-1 is cationic and charged at cellular pH because it has a pK_a of 9.2 [3].



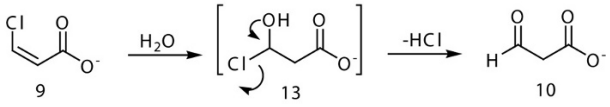
Scheme 3.

Janssen and co-workers first cloned and expressed CaaD and proposed that the enzyme adds water to *trans*-3-chloroacrylic acid (**8**, Scheme 4) to form an unstable halohydrin intermediate (**13**) [25]. This halohydrin then collapses to release HCl and malonate semialdehyde (**10**, Scheme 4). CaaD is a heterohexamer with 3 α-subunits (75 amino acids each) and 3 β-subunits (70 amino acids each). Like all tautomerase superfamily members, CaaD carries out this reaction without any cofactors.



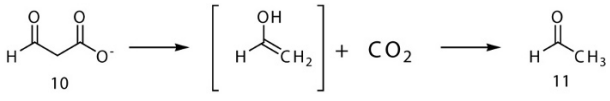
Scheme 4.

cis-CaaD has been cloned and expressed by Whitman and co-workers, and is highly specific for the *cis*-isomer of 3-chloroacrylic acid (**9**, Scheme 5) [10]. It was also proposed that *cis*-CaaD adds water to **9** to form an unstable halohydrin intermediate (**13**). Like the proposed mechanism for CaaD, the halohydrin collapses to form HCl and malonate semialdehyde (**10**, Scheme 5). Unlike CaaD, *cis*-CaaD is a trimer where each monomer consists of 149 amino acids. *cis*-CaaD also carries out its reaction without any cofactors.



Scheme 5.

MSAD has been cloned and expressed by Whitman and co-workers [11]. It was determined to be a metal-ion independent decarboxylase. MSAD catalyzes the decarboxylation of malonate semialdehyde (**10**, Scheme 6) where the cationic Pro-1 may polarize the 3-carbonyl group of **10**, and facilitate decarboxylation. Subsequent ketonization of the resulting enol yields **11.** MSAD is a trimer where each monomer consists of 129 amino acids.



Scheme 6.

All three enzymes are found in a bacterial pathway that degrades *cis*- and *trans*-1,3-dichloropropene (**6**, Scheme 3) [3]. This isomeric mixture is the active ingredient in various nematocides marketed under the tradenames Telone II and Shell D-D [15–17]. These compounds are mixed into the soil to kill nematodes. The 1,3-dichloropropene isomers are rapidly degraded in soil in part due to this pathway. In three enzyme-catalyzed steps, 1,3-dichloropropene (**6**) is transformed to the *cis*- and *trans*-isomers of 3-chloroacrylic acid through the intermediacy of an alcohol (**7,** Scheme 3) [3, 15–17]. CaaD and *cis*-CaaD then process the appropriate isomers of 3-chloroacrylic acid (**8** and **9**, respectively) to malonate semialdehyde (**10**). MSAD converts malonate semialdehyde to acetaldehyde (**11**), which is channeled into the Krebs Cycle.

The mechanistic and structural diversity in the tautomerase superfamily, summarized in Table 1, raise questions about how these enzymes evolved. How enzymes evolve and how new enzymatic activities arise are two questions that have evoked much interest, discussion, and debate [26–36]. There is a considerable body of evidence suggesting that many enzymes evolve from a common ancestor by divergent

Table 1. Characteristics of the five families comprising the tautomerase superfamily.

Family	Enzyme	Quaternary structure	Amino acids per monomer	Key active site residues
4-OT	4-OT	Hexamer	62	Pro-1, Arg-11, Arg-39, Phe-50
	CaaD	Heterohexamer	α-subunit 75 β-subunit 70	βPro-1, αArg-8, αArg-11, αGlu-52
CHMI	CHMI	Trimer	125	Pro-1, Arg-40, Arg-71
MIF	MIF	Trimer	114	Pro-1
cis-CaaD	*cis*-CaaD	Trimer	149	Pro-1, His-28, Arg-70, Arg-73,
	Cg10062	Trimer	149	Tyr-103, Glu-114
MSAD	MSAD	Trimer	129	Pro-1, Asp-37, Arg-73, Arg-75

evolution. The growing number of superfamilies, with shared structural and functional features, provides strong support for divergent evolution [37–42]. The common explanation for divergent evolution involves duplication of the gene for the parent enzyme followed by limited random mutagenesis to produce a new enzymatic function. The evolution of the new activity is driven by a selective advantage for the organism (e.g., frequently growth on alternative substrates). After the acquisition of a new activity, sequences likely diverge rapidly and markedly because there is little sequence identity among superfamily members [37].

In the tautomerase superfamily, Nature has apparently fashioned new activities and structures by "stitching together" various combinations of the same simple structural unit, the β−α−β fold [1–3]. Thus far, there are characterized superfamily members made up of short monomers (4-OT, CaaD) and those made up of longer monomers (PPT, *cis*-CaaD, and MSAD) [1–3]. The lessons learned from studies of these enzymes and how Nature stitched the scaffold together have enormous implications and tremendous potential. If the principles used by Nature to create this diversity can be determined, it may be possible to mimic Nature's strategy, for example, and create new proline-based biocatalysts with designed activities using the β-α-β scaffold [2, 3].

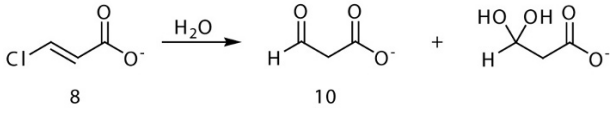
In addition to the fact that the β−α−β motif has been used in different ways to generate diverse reaction types and structures, tautomerase superfamily members have three unique advantages that make them particularly attractive experimental vehicles for study. The small monomer size, the absence of metal ions or coenzymes, and stability make them easily manipulated and amenable to study by a host of diverse techniques. These same features also make tautomerase superfamily members ideal candidates for the laboratory evolution of activities and the resulting insight into divergent evolution. The new constructs are readily purified, their activities determined, and the mechanism and structures delineated by many different techniques. These studies provide insight into the progenitors of these enzymes and suggest scenarios for the diversification of enzymatic function and mechanism within the tautomerase superfamily. This remaining part of this review focuses primarily on our studies of the three 1,3-dichloropropene catabolic enzymes and a recently characterized *cis*-CaaD homologue Cg10062, with an emphasis on the active site motifs and catalytic reaction mechanisms. These studies illustrate how changes in sequence and structure translate into changes in substrate specificity and reaction mechanism and, in turn, lead to the evolution of new enzymatic activities in the tautomerase superfamily. They also address the evolvability of superfamily members by the exploitation of catalytic promiscuity, provide insight into the progenitors of the 1,3-dichloropropene catabolic enzymes, and have implications for the evolution of metabolic pathways and enzymatic activities, in general.

trans-3-chloroacrylic acid dehalogenase (CaaD)

The genes encoding the α- and β-subunits for CaaD were initially cloned from *P. pavonaceae* strain 170, expressed, and the enzyme purified and characterized by Poelarends et al. [25]. CaaD is a heterohexamer, consisting of 3 α-subunits (75 amino acid residues each) and 3 β-subunits (70 amino acid residues each) (Table 1). Sequence analysis revealed similarities between the CaaD α-subunit and various 4-OT sequences (pairwise identities ranging from 23–35 %) and the CaaD β-subunit and the same 4-OT sequences (pairwise identities ranging from 16–25 %). This analysis implicated βPro-1 and αArg-11 as key catalytic residues, which was confirmed by site-directed mutagenesis [25]. Based on the 4-OT mechanism, it was initially thought that Pro-1 might function as a base to activate water and Arg-11 interacted with the C-1 carboxylate group to facilitate the addition of water [25].

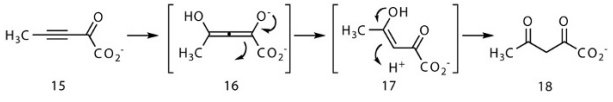
With this information in hand, Whitman and coworkers continued studies on CaaD. An efficient expression system for CaaD and a direct UV assay for

monitoring activity were developed [43]. The identities of the products (i.e., **10** and the hydrate, Scheme 7) were verified by ^{1}H NMR spectroscopy and the behavior of CaaD with three acetylene compounds (**15**, Scheme 8 and **19, 20**, Scheme 9) was examined.



Scheme 7.

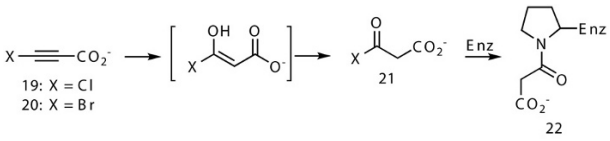
2-Oxo-3-pentynoate (**15**) is a potent active-site-directed irreversible inhibitor of 4-OT that covalently modifies Pro-1 [44,45]. It was anticipated that if βPro-1 of CaaD functioned as a base, it would also be covalently modified by **15.** Instead, it was found that CaaD processed **15** to acetopyruvate (**18**) quite efficiently (k_{cat} = 0.7 s^{-1}, K_m = 110 μM, k_{cat}/K_m = 6.4 ×10^3 M^{-1} s^{-1}) [43]. {By comparison, CaaD processes **8** to **10** with a k_{cat} = 3.8 s^{-1}, a K_m = 31 μM, and a k_{cat}/K_m = 1.2 ×10^5 M^{-1} s^{-1} [43]}. This was the first indication that the protonation states of the catalytic Pro-1 in CaaD and 4-OT were different. It was proposed that CaaD initiates the Michael addition of water to C-4 of **15** to form **16** (Scheme 8) [43]. Protonation at C-3 of **16** leads to **17**, which can, in turn, ketonize to afford acetopyruvate (**18**).



Scheme 8.

The conversion of **15** to **18** suggested that CaaD might also add water to the 3-halopropiolates (**19** and **20**, Scheme 9), transforming them into agents that would function as irreversible inhibitors. Indeed, it was found that **19** and **20** are potent mechanism-based inhibitors of CaaD [43]. In one possible mechanism, hydration of these compounds results in the formation of an unstable enol, which ketonizes to an acyl halide (**21**, Scheme 9). Alternatively, the enol can directly expel halide to form a ketene. Subsequent reaction with an active site nucleophile inactivates the enzyme. Mass spectral analysis showed that the site of inactivation is βPro-1. It should be noted that the initial hydration of **19** or **20** by CaaD would presumably remove a proton from the cationic βPro-1, making it nucleophilic.

A crystal structure of CaaD, inactivated by the species derived from the hydration of **20**, was obtained (Fig. 2)



Scheme 9.

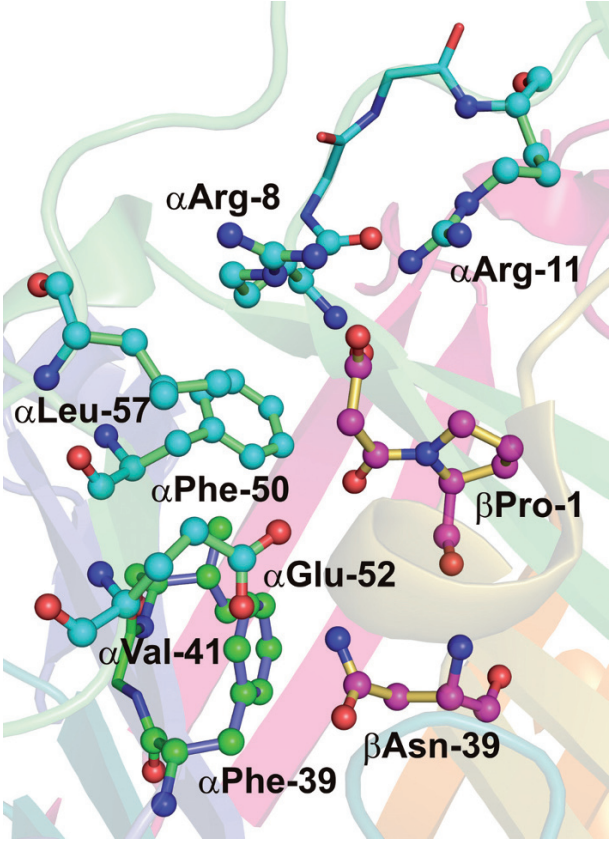
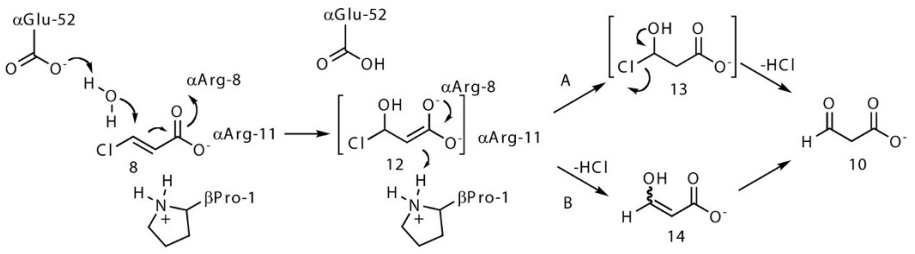
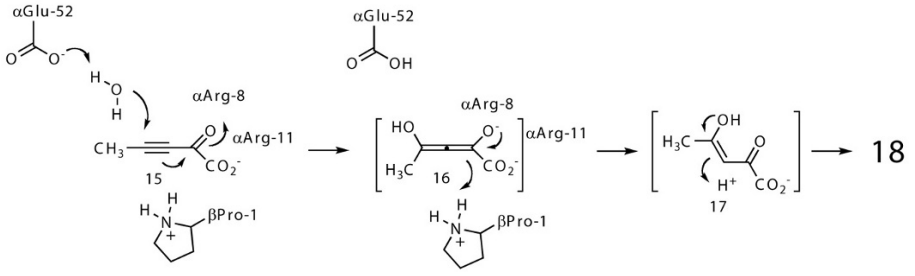


Figure 2. A close-up of the active site of CaaD where βPro-1 is covalently modified by 3-oxopropanoate (see **22** in Scheme 9). The roles of the key active site residues (βPro-1, αArg-8, αArg-11, and αGlu-52) and their interactions are discussed in the text.

[46]. The geometry of the complex identified a new catalytic residue, αGlu-52, and suggested that it activates water for nucleophilic attack while βPro-1 functions as a general acid catalyst (Scheme 10). The proposed mechanism for hydration shows that the 3-carbonyl oxygen in **22** (Scheme 9) is likely derived from water. Moreover, the observed hydrogen bond between the oxygen of αGlu-52 and the backbone carbonyl oxygen of βIle-37 indicates that αGlu-52 is no longer ionized due to the abstraction of a proton from a water molecule [46]. The complex also identified αArg-8 as a second residue which, along with αArg-11, binds to the carboxylate group of **8.** This combined interaction facilitates catalysis by generating a partial positive charge at C-3 and enhancing the electrophilicity of this position. Site-directed muta-



Scheme 10.



Scheme 11.

genesis confirmed the importance of both αGlu-52 and αArg-8 for the activity of CaaD [46]. There was no detectable CaaD activity for the αE52Q mutant even after an extended (24 h) incubation period. The αR8A mutant has a small amount of activity over a 24 h incubation period. The βP1A mutant has no detectable activity within a 24 h incubation period but the αR11A mutant retains a small amount of activity. (The activity observed after these lengthy incubation periods reflects slow catalysis, but is not due to a "resuscitation" of the enzyme.) Finally, the pK_a of the βPro-1 was determined to be ~9.2 by direct titration using ^{15}N NMR spectroscopy, where the chemical shift of the prolyl nitrogen is followed as a function of pH in ^{15}N-labeled CaaD [47]. Hence, CaaD became the first tautomerase superfamily member in which Pro-1 serves as an acid catalyst and not a base catalyst. These studies led to the current working hypothesis for CaaD, which is shown in Scheme 10. In this mechanism, αGlu-52 activates a water molecule for attack at C-3 of **8** while the pair of arginines binds the C-1 carboxylate group, thereby polarizing the substrate. The resulting enediolate (i.e., **12**) can suffer one of two fates: ketonization with protonation at C-2 (route A) or ketonization with loss of chloride (route B). The scenario in route A produces the halohydrin **13**, which can undergo enzyme-catalyzed or chemical collapse to yield HCl and **10.** The scenario in route B produces the enol **14**, which can be ketonized by the enzyme to afford **10.**

This mechanism can be extrapolated to provide a plausible mechanism for the hydration of **15** (Scheme 11). Accordingly, αGlu-52 activates a water molecule for attack at C-4 of **15.** The addition of water is facilitated by the interactions of αArg-8 and αArg-11 with the 2-carbonyl group and the C-1 carboxylate group. The resulting species, **16**, can pick up a proton at C-3 (potentially from βPro-1) to generate **17.** Ketonization of **17** yields the observed product **18.** The linear acetylene molecule is readily accommodated in the active site as indicated by the kinetic parameters determined for the conversion of **15** to **18** (*vide supra*). This observation is not surprising given the extended geometry of the *trans*-substrate.

cis-3-chloroacrylic acid dehalogenase (*cis*-CaaD)

The observation that various bacteria processed both isomers of 1,3-dichloropropene (**6**, Scheme 3) suggested that a *cis*-CaaD activity might be present in these bacteria. Indeed a *cis*-CaaD was isolated from coryneform bacterium strain FG41 in 1992 [16]. The enzyme was purified and a large fragment sequenced. It was reportedly a dimer or trimer consisting of 16.2 kDa subunits [16]. 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.

With this information in hand, the gene for *cis*-CaaD from coryneform bacterium strain FG41 was cloned and expressed and the protein product was purified and characterized [10]. The enzyme is a trimer where each monomer consists of 149 amino acids (Table 1). Like CaaD, it is highly isomer specific. Sequence alignment and structural analysis implicated Pro-1, Arg-70, Arg-73, and Glu-114 as active-site residues (Table 1) [10]. Site-directed mutagenesis experiments confirmed their importance for activity. Once again, the acetylene compounds (**15**, Scheme 8 and **19** and **20**,

Scheme 9) were processed by *cis*-CaaD to products consistent with an enzyme-catalyzed hydration reaction previously established for CaaD [10]. Hydration of **15** afforded **18** (although not as efficiently as CaaD) while the 3-halopropiolates (e.g., **19** and **20**) are transformed into irreversible inhibitors that covalently modify Pro-1. In addition, pH rate profiles suggested that an acid catalyst (pK_a ~9.3) is required for activity [10]. Although the pK_a of Pro-1 has not yet been established by direct NMR titration, the observed kinetic pK_a could be that of Pro-1. These results suggested an initial mechanism for *cis*-CaaD that parallels that of CaaD (Scheme 10) where the residues in CaaD (αGlu-52, αArg-8, αArg-11 and βPro-1 as shown clockwise in Scheme 10) are replaced with the corresponding ones in *cis*-CaaD (Glu-114, Arg-70, Arg-73, and Pro-1) [10].

Mutagenesis results and inhibition studies uncovered two intriguing differences between the CaaD and *cis*-CaaD mechanisms with potential evolutionary implications. The first difference emerged from a comparison of mutagenesis results [10]. The αE52Q mutant of CaaD has no residual activity, whereas the E114Q mutant of *cis*-CaaD retains a significant amount of activity (~10-fold decrease in k_{cat}/K_m). This observation suggested that water activation is more complex (i.e., more residues involved) in *cis*-CaaD than it is in CaaD. The second difference came from inhibition studies using (*R*)- and (*S*)-oxirane-2-carboxylate (**23**, Scheme 12) [48]. Only *cis*-CaaD is inactivated (irreversibly) and only by the *R*-enantiomer (of **23**). It was also found that Pro-1, Arg-70, and Arg-73 are required for covalent modification (i.e., the P1A, R70A, and R73A mutants are not modified), but Glu-114 is not required (i.e., the E114Q mutant is covalently modified). Mass spectral analysis showed (once again) that Pro-1 is the site of covalent modification.

The crystal structures of native *cis*-CaaD and *cis*-CaaD inactivated by (*R*)-**23** were subsequently solved [49]. The crystal structure of *cis*-CaaD inactivated by (*R*)-**23** is shown in Figure 3. The observations gleaned from this structure and that of the native *cis*-CaaD have four major implications. First, the native structure identified Tyr-103 and His-28 as additional catalytic residues. The structures suggest that Tyr-103 assists Glu-114 in the activation of water and His-28 assists Arg-70/Arg-73 in the activation of substrate through an interaction with the carboxylate group. These observations and mutant analysis indicate that water activation is more important to catalysis in CaaD, whereas substrate activation is more important in *cis*-CaaD catalysis. Second, the inactivated structure confirmed the nature of the covalent adduct (e.g., **24**, Scheme 12), established that C3 of **23** is the site of nucleophilic attack by Pro-1 (as the neutral form

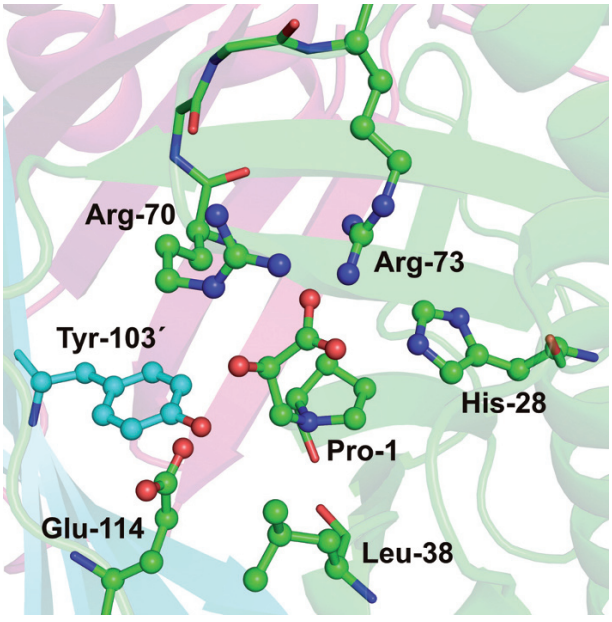
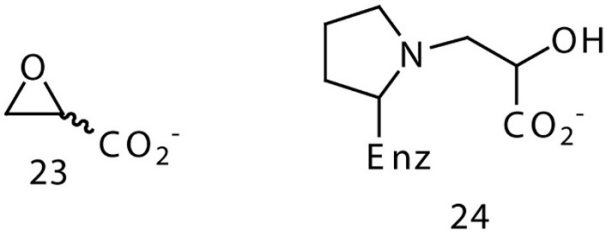


Figure 3. A close-up of the active site of *cis*-CaaD where Pro-1 is covalently modified by (*R*)-**23** (see **24** in Scheme 12). The roles of the key active site residues (Pro-1, His-28, Arg-70, Arg-73, Tyr-103′, and Glu-114) and their interactions are discussed in the text. The prime designation indicates that Tyr-103 comes from an adjacent subunit.

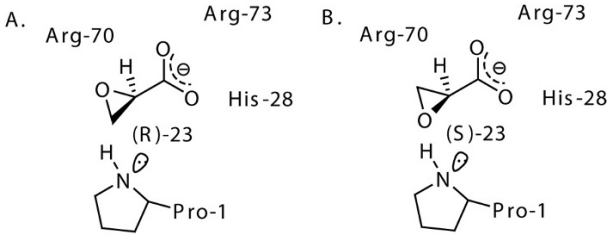
which, under the pH conditions of the inactivation experiments, is present in ~1 %), and implicated His-28 as a major determinant in positioning (*R*)-**23** for alkylation of Pro-1 and subsequent inactivation. Third, His-28 and Tyr-103 appear to be two determinants of the substrate specificity for *cis*-CaaD. Finally, the results are consistent with the proposal that *cis*-CaaD and CaaD form two separate lineages that followed the independent duplication of a 4-OT-like sequence [49].



Scheme 12.

The crystal structure of the inactivated *cis*-CaaD [by (*R*)-**23**] and mutant analysis provided a mechanism for inactivation, a structural basis for the stereospecificity of the reaction [i.e., only the (*R*)-enantiomer inactivates *cis*-CaaD], and a rationale for the observation that CaaD is not inactivated by either enantiomer of **23** [49]. The interactions in the crystal structure

implicate both Arg-73 and His-28 in the binding of the carboxylate of (*R*)-**23** (Scheme 13A). Arg-70 may facilitate ring opening by direct interaction with the epoxide oxygen or by placing a water molecule in position to interact with the epoxide oxygen [49]. If (*S*)-**23** binds with similar interactions between the carboxylate group and Arg-73 and His-28 (Scheme 13B), the C3 carbon is directed away from Pro-1 and towards Arg-70. This binding mode precludes alkylation. CaaD may not be alkylated by either enantiomer of **23** because in the absence of His-28 the carboxylate group might now interact with αArg-8 and αArg-11. This interaction would prevent one of the arginines from functioning as the required proton donor [49].



Scheme 13.

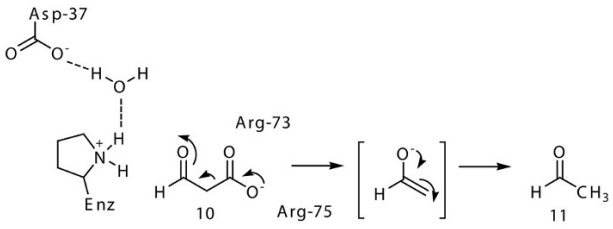
A comparison of the *cis*-CaaD and CaaD structures offers an explanation for the individual substrate specificities [46, 49]. In *cis*-CaaD, the carboxylate group of **9** could be bound by the His-28/Arg-70/Arg-73 cluster, such that the substrate is directed towards the enzyme's surface [49]. In CaaD, the carboxylate group of the substrate interacts with the αArg-8/αArg-11 pair so that the remaining part of the substrate is projected deeper into the active site [46]. The substrate binding pockets fit the "shapes" of their respective substrates, that is, the pocket of *cis*-CaaD is more U-shaped, whereas the pocket of CaaD is more elongated. One of the residues responsible for this shape difference is Tyr-103 (αVal-41 in CaaD). In CaaD, αVal-41 creates a hydrophobic region allowing the 3-chloro moiety of **8** to bind between αPhe-39 and αPhe-50. The presence of the larger Tyr-103 residue in *cis*-CaaD effectively blocks the binding of the 3-chloro group of **8.** Instead, a hydrophobic pocket formed by the side chains of Thr-34, Leu-38 and Leu-119 could "cradle" the 3-chloro moiety of **9.** Thus, His-28 and Tyr-103 might also be partially responsible for the specificity of *cis*-CaaD [49].

Malonate semialdehyde decarboxylase (MSAD)

Poelarends et al. reported that the genes for the CaaD α- and β-subunits are found in a cluster (in *P. pavonaceae* 170) that includes two additional open reading frames located immediately downstream [25]. One, *orf130*, was hypothesized to function as a malonate semialdehyde decarboxylase (MSAD) [25]. The other open reading frame likely encodes a dehydrogenase, based on sequence analysis. Hence, the *orf130* gene was cloned and expressed, and the protein product was purified and characterized [11]. MSAD is a trimer where each subunit consists of 129 amino acids (Table 1). The enzyme was shown to carry out a metal-ion independent decarboxylation reaction (using **10** in Scheme 6, and generating **11** and the hydrate). Sequence analysis placed MSAD in the tautomerase superfamily (but as a representative of yet another new family) and implicated Pro-1 and Arg-75 as potential active site residues. Site-directed mutagenesis confirmed the importance of these residues for activity [11].

A priori, (and in the absence of a cofactor) MSAD could reasonably proceed through a Schiff base mechanism or one in which a cationic Pro-1 polarized the 3-carbonyl group of **10** [50]. A distinguishing feature between the two mechanisms is the ionization state of Pro-1. Accordingly, the behavior of MSAD with **15** (Scheme 8) was determined and the pK_a of Pro-1 was measured by direct titration using ^{15}N NMR spectroscopy and uniformly ^{15}N-labeled enzyme [50]. MSAD converts **15** to **18** (Scheme 8) in a process that also requires Pro-1 and Arg-75. Hence, MSAD is a decarboxylase with a promiscuous hydratase activity. In addition, the pK_a of Pro-1 was determined to be ~9.2 by ^{15}N NMR titration [50]. These observations argued against a Schiff base mechanism and in favor of the mechanism where Pro-1 polarizes the carbonyl oxygen of **10** by hydrogen bonding and/or an electrostatic interaction (Scheme 14). Arg-75 could position and "lock" the carboxylate group in a favorable orientation for decarboxylation. In view of the promiscuous hydratase activity (uncovered by the conversion of **15** to **18**) MSAD was incubated with two potential mechanism-based inhibitors, **19** and **20** (Scheme 9) [51]. Analysis showed that both compounds inactivated MSAD and, once again, Pro-1 was the site of modification [51]. It is again presumed that the initial hydration of **19** or **20** by MSAD removes a proton from Pro-1 and enables it to function as a nucleophile.

Subsequently, crystal structures of native MSAD, the P1A mutant, and an inactivated MSAD were determined [52]. The crystal structure of the inactivated MSAD is shown in Figure 4. The structures confirm that MSAD is metal-ion independent (i.e., there is no metal ion in the active site despite the high concentrations of metal ion in the crystallization buffers). The P1A mutant structure is nearly identical to that of the



Scheme 14.

native structure (except for the mutation) suggesting that the loss of activity associated with this mutant does not result from a structural defect. In the structure of MSAD inactivated by the species derived from the hydration of **19**, the positioning of the 3-oxopropanoate adduct (**22**, Scheme 9) implicated Asp-37 and Arg-73 as additional catalytic residues in the decarboxylation and hydration mechanisms. The importance of these residues to both activities was confirmed by site-directed mutagenesis. These observations and all of the preceding mechanistic work led to the current working hypotheses for the hydration and decarboxylation mechanisms.

The decarboxylation mechanism is summarized in Scheme 14. In the crystal structure, Asp-37 participates in a hydrogen bond network, and in this capacity may be partially responsible for the observed pK_a of Pro-1 (~9.2) [52]. In accord with this supposition, it has recently been determined that the pK_a of Pro-1 is lowered to ~6.4 (as measured by direct titration using ^{15}N NMR spectroscopy) when Asp-37 is replaced with an alanine [Darty J. E., Johnson Jr. W. H., Harris T. K., and Whitman C. P. (2008) unpublished results]. The proposed mechanism for the hydration of **15** parallels that of CaaD (Scheme 11) where Asp-37 replaces αGlu-52, Arg-73/Arg-75 replace αArg-8/αArg-11, and Pro-1 replaces βPro-1. Asp-37 could activate a water molecule for nucleophilic attack at C-4 of **15** to initiate the Michael addition of water. Arg-73 and Arg-75 could polarize the carbonyl oxygen and assist in binding of the carboxylate group. Pro-1 is an obvious candidate to provide a proton at C-3 to complete the Michael addition of water. Ketonization of **17** to **18** could be enzyme-catalyzed or a non-enzymatic process.

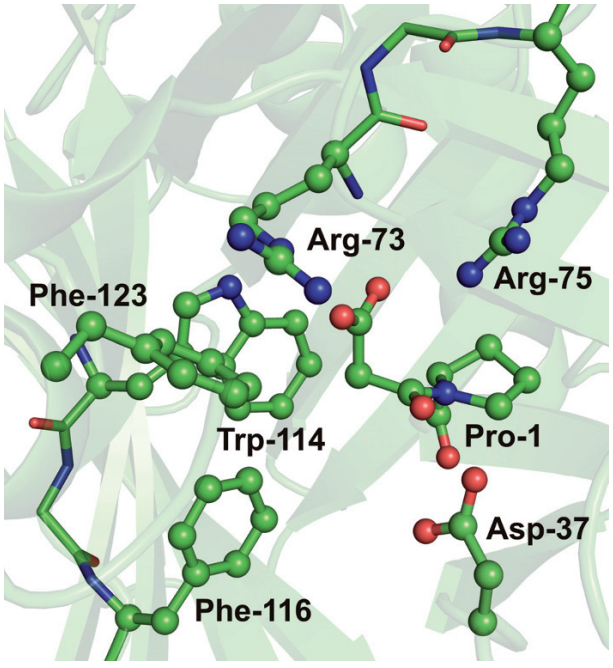


Figure 4. A close-up of the active site of MSAD where Pro-1 is covalently modified by 3-oxopropanoate (see **22** in Scheme 9). The roles of the key active site residues (Pro-1, Asp-37, Arg-73, and Arg-75) and their interactions are discussed in the text.

Evolution of *trans*- and *cis*-3-chloroacrylic acid dehalogenase and malonate semialdehyde decarboxylase

Various scenarios can be envisioned for the evolution of CaaD, *cis*-CaaD, and MSAD. For CaaD and *cis*-CaaD, the conservation of the β−α−β fold and the key functional groups [Pro-1, αArg8/Arg-70, αArg11/Arg-73, and αGlu-52/Glu-114] suggest that the two enzymes are related by divergent evolution from a common ancestor [49]. The different oligomer structures (heterohexamer vs. trimer) and low sequence identity (<20 %) indicate that CaaD and *cis*-CaaD diverged a long time ago. The presence of His-28 and Tyr-103 in the *cis*-CaaD active site (*and* the low sequence identity) invoke a scenario in which *cis*-CaaD evolved from an independent gene duplication event of a small β−α−β gene (e.g., a 4-OT-like gene) followed by gene fusion [49]. CaaD also might have evolved from gene duplication of a small 4-OT-like gene, but the two genes co-evolved to form the two subunits of CaaD.

The functional and mechanistic connection between the two dehalogenases and 4-OT, and the proposed evolutionary link between these enzymes, was further strengthened by the observations that 4-OT from *P. putida* mt-2 and a homologous tautomerase from *Bacillus subtilis*, YwhB, show low-level 3-chloroacrylic acid dehalogenase activity (with a strong preference for the *trans*-isomer) [53]. Furthermore, one point mutation (L8R) in 4-OT results in a mutant enzyme that has an increased CaaD activity approaching 50-fold as defined by the k_{cat}/K_m (0.55 $M^{-1} s^{-1}$) for **8** [54]. Although the CaaD activities of YwhB, 4-OT, and 4-OT-L8R are low, these observations nonetheless suggest that a catalytically promiscuous 4-OT-like enzyme could have been recruited to serve as a CaaD/

cis-CaaD because the enzyme fortuitously had a low but useful level of dehalogenase activity.

To further explore the origin of CaaD, the enzyme was examined for vestigial tautomerase activity. Although CaaD does not catalyze the 4-OT reaction (i.e., the rate of production of 2-oxo-3-hexenedioate (**3**, Scheme 1) cannot be distinguished from the non-enzymatic rate), it exhibits a robust phenylpyruvate tautomerase (PPT) activity, catalyzing the conversion of **4** to **5** (Scheme 2) with a k_{cat}/K_m value that is near the one measured for CaaD-catalyzed dehalogenation of **8** (Scheme 4) [55]. CaaD has a k_{cat}/K_m value of ~7.1×10^4 M^{-1} s^{-1} using **8** and a k_{cat}/K_m value of ~2.3×10^4 M^{-1} s^{-1} using **4** [55]. Moreover, the reaction is stereoselective in D_2O, resulting in the formation of the 3*S*-isomer of [3–^{2}H]phenylpyruvate in a 1.8:1 ratio. The stereoselectivity along with mutagenesis results indicate that the ketonization reaction occurs in the active site. The PPT activity of CaaD is a striking example of catalytic promiscuity and further establishes a functional and mechanistic link between this dehalogenase and the superfamily tautomerases. Taken together, these findings provide strong support for the hypothesis that CaaD may have evolved from a 4-OT-like enzyme by divergent evolution.

The PPT activity of CaaD has an additional mechanistic consequence. The robust activity argues for the presence of the enol intermediate **14** in the dehalogenase reaction (Scheme 10), and is consistent with the α,β-elimination reaction mechanism shown in route B [55]. A mechanism involving the collapse of the halohydrin **13** (route A) cannot be ruled out because the PPT activity could be a reflection of the ketonization of **12**, an enol, which is common to both mechanisms. *cis*-CaaD also exhibits a PPT activity, but a diminished one (k_{cat}/K_m value of ~1.8×10^3 M^{-1} s^{-1}) [55]. The U-shaped active site cavity may not be able to accommodate **4** as well as the elongated active site cavity of CaaD. Nonetheless, this observation could similarly be used to argue that α,β-elimination of HCl follows the addition of water to **9** in the *cis*-CaaD reaction mechanism.

The similarities between CaaD/*cis*-CaaD and MSAD and the fact that they catalyze successive reactions in the 1,3-dichloropropene catabolic pathway suggest that they may have also evolved by a form of "retrograde evolution" in that they diverged from an ancestral enzyme that catalyzed both reactions [26,52]. For CaaD and MSAD, the two enzymes are structurally homologous and key catalytic residues (Pro-1 and the two arginines) are positionally conserved [52]. These observations are significant in view of the functional similarities between the enzymes. Both function as hydratases, converting **15** to **18**, and both use Pro-1 (with a pK_a of ~9.2) and a conserved pair of arginines to catalyze this reaction as well as their physiological reactions [52]. Finally, as a result of the hydration reaction, the 3-halopropiolates (e.g., **19** and **20** in Scheme 9) are converted to reactive species that result in acylation and inactivation of both enzymes.

In one scenario, the ancestral enzyme may have functioned primarily as a hydratase because the hydration of **8,** an α,β-unsaturated acid, is the chemically more difficult reaction [52]. A random encounter of **10** with the cationic Pro-1 could result in decarboxylation and make this ancestral hydratase an accidental decarboxylase. In this scenario, gene duplication (and gene fusion) gave rise to separate enzymes that retained the components for the hydration reaction along with the rudimentary decarboxylase activity. Enhancement of the accidental decarboxylase activity could result from a limited number of mutations that increased the probability of an encounter and optimized the position of **10** with respect to the cationic Pro-1 and introduced additional catalytic elements to facilitate decarboxylation. In this version of events, the hydratase activity of MSAD is a remnant.

Cg10062, a *cis*-3-chloroacrylic acid dehalogenase homologue, from *Corynebacterium glutamicum*

Since the initial discovery of the tautomerase superfamily in 1996, tremendous progress has been made in the identification and characterization of "new" superfamily members. Sequence databases now contain hundreds of tautomerase superfamily members, many of which have unknown catalytic or biological functions. There are currently 10 unique members of the tautomerase superfamily for which high-resolution structures are available. The protein structures, coupled in most cases with a delineation of the catalytic reaction mechanisms, have significantly enhanced our understanding of the breadth of catalytic, structural, and mechanistic diversity in the superfamily. With the sequence and structural data now available, the tautomerase superfamily members can be partitioned into five subgroups (i.e., families), based on sequence identity and variation in the active site residues. The subgroups are named for the first characterized enzyme in each family and are therefore known as the 4-OT, CHMI, MIF, *cis*-CaaD, and MSAD families. Although members of the five families have little overall sequence similarity, the amino-terminal proline is conserved in all sequences. Hence, the conservation of this residue at the N-terminus of small β-α-β proteins (i.e., <150 residues in length) is a primary criterion for identi-

fying new superfamily members in sequence databases.

An interesting example of a recently identified superfamily member is a 149-amino acid *cis*-CaaD homologue from *C. glutamicum,* designated Cg10062 [10]. The physiological function of Cg10062 is unknown and the gene has no obvious genomic context. The protein shares 34% sequence identity (and 53% similarity) with *cis*-CaaD and the residues critical for *cis*-CaaD activity (Pro-1, His-28, Arg-70, Arg-73, Tyr-103, Glu-114) are present in Cg10062 (Table 1) [10,49]. Like *cis*-CaaD, Cg10062 functions as a hydratase [56]. It converts **15** to **18** and is inactivated by the species (e.g., an acyl halide or a ketene in Scheme 9) derived from the hydration of **19** and **20.** However, despite the presence of the core catalytic machinery, Cg10062 is a poor *cis*-CaaD: it has much lower catalytic efficiency and lacks stereospecificity [56]. The enzyme processes both isomers of 3-chloroacrylate at low levels albeit with a clear preference for the *cis*-isomer. The study of Cg10062 shows that all of the determinants responsible for optimal *cis*-CaaD activity and specificity have not yet been identified.

An examination of the *cis*-CaaD crystal structure and the Cg10062 sequence suggests two additional factors. The active site of *cis*-CaaD is defined by Pro-1, His-28, Thr-32, Thr-34, His-69, Arg-70, Arg-73, Tyr-103, Met-112, and Glu-114. Because most of the same residues are found in Cg10062, the active site of Cg10062 is not likely to be very different from that of *cis*-CaaD. There are, however, two intriguing differences. First, His-69 in *cis*-CaaD is replaced with an isoleucine in Cg10062. Second, there are significant differences between some of the residues in a nine-residue loop that connects the α-helix of a β−α−β motif to the second β-strand in the two enzymes.

The effects of these changes on catalysis and specificity are potentially substantial. In *cis*-CaaD, His-69 and His-28 interact with the hydroxyl group of Tyr-3. It is not known if this interaction plays a role in the *cis*-CaaD mechanism, but the uncharged, hydrophobic isoleucine could disrupt a similar interaction in Cg10062. As a result, the position of His-28 could be altered or the properties of the active site could be somewhat modified. The crystal structure of *cis*-CaaD also shows a nine-residue loop connecting the α-helix of a β−α−β motif to the second β-strand. In *cis*-CaaD, Leu-31 and Ala-39 anchor this loop, which consists of Thr-32, Gly-33, Thr-34, Gln-35, His-36, Phe-37, and Leu-38. The Cg10062 sequence shows that Leu-31 and Val-39 anchor the loop, which consists of Ala-32, His-33, Ala-34, Pro-35, Lys-36, Tyr-37, and Leu-38. Modeling studies implicate Thr-34 as part of a binding pocket for the 3-chloro group of the *cis*-substrate [49]. If binding is mediated by a hydrogen bond from Thr-34, this interaction cannot be made in Cg10062. Moreover, Gly-33 of *cis*-CaaD is replaced with a bulky, and presumably charged, histidine, and Gln-35 (in *cis*-CaaD) is replaced with a rigid proline. These changes will almost certainly affect the loop properties, and if this loop contributes to catalysis and/or specificity, the differences could translate into the diminished activity and/or lack of isomer specificity.

On the basis of these observations, it has been suggested that Cg10062 is but a few mutations away from being a highly specific and efficient *cis*-CaaD and, as such, could be representative of a progenitor for *cis*-CaaD [56]. In one scenario, this progenitor, like Cg10062, has the core catalytic machinery for a hydration reaction but lacks specificity for the *cis*-isomer. Mutations in the loop accompanied by the conversion of Ile-69 to a histidine could provide the necessary elements and complete the evolution of Cg10062 to *cis*-CaaD. The results of these and other experiments on Cg10062 will provide a more complete understanding of CaaD and *cis*-CaaD catalysis and how the two enzymes might have diverged to acquire their individual specificities. For these reasons, studies of Cg10062 are being pursued in our laboratories.

Conclusions and future directions

Mechanistic, mutagenesis, and structural studies of CaaD, *cis*-CaaD, and MSAD have identified key active site residues (Table 1) and established minimal mechanisms for each enzyme (Schemes 10 and 14). These studies have also uncovered intriguing similarities, fascinating differences, and tantalizing evolutionary histories for the three enzymes. However, many mechanistic and evolutionary questions remain unanswered. For example, in CaaD and *cis*-CaaD, the actual orientation of the substrate in the active site and the interactions of Pro-1, the arginine pair, and the water-activating glutamate with substrate are really not known. The arrangement shown in Scheme 10 and the analogous one for *cis*-CaaD are inferred from structures of inactivated complexes and may not accurately portray substrate binding. Moreover, it is not clear why two arginines are necessary for substrate activation (i.e., polarization of the carboxylate group of **8** and **9**). The structural basis for the pK_a of Pro-1 is unknown, as is why Pro-1 is so critical for activity. It is also not known if other residues or structural elements are required for efficient CaaD and *cis*-CaaD catalysis. It has only been recently discovered that *cis*-CaaD requires two additional residues (Tyr-103 and His-28) for activity

and may somehow utilize a loop (residues 31–39) for catalysis and/or specificity.

Many of the same questions apply to the MSAD mechanism. For example, the actual substrate orientation in the active site and the interactions of the substrate with active site residues (e.g., Pro-1, Arg-73, and Arg-75) are not known. The mechanism shown in Scheme 14 was also extrapolated from the crystal structure of an inactivated complex. The role of Asp-37 in the mechanism is not completely clear and the residue responsible for C-2 protonation of the enol of **11** (Scheme 6) is not known. Finally, it is not known if other elements are required for efficient catalysis.

Understanding the individual reaction mechanisms of these enzymes will answer one of the fundamental evolutionary questions that has emerged in studies of this superfamily: how do CaaD, *cis*-CaaD, and MSAD use the same core set of residues (i.e., Pro-1, the arginine pair, and a water-activating Glu or Asp) to carry out the three different reactions, isomer-specific dehalogenations and decarboxylation? It is not known how the properties of these core residues have been modulated, how their positions within the active sites have changed, and how the active sites are otherwise modified to accommodate the three different activities. Moreover, it is not entirely clear how the active site properties of this group of enzymes differ from those of the well-characterized tautomerases. We have identified some differentiating elements but not all of them. Our ongoing mechanistic and structural studies will continue to shed light on these unresolved issues.

Tautomerase superfamily members are constructed from a simple β–α–β building block. Our studies show that this scaffold is a versatile and sturdy one, manipulated quite ingeniously to generate structural and mechanistic diversity. Despite the apparent simplicity of the template and our intensive efforts, we continue to be surprised and it appears that we have only scratched the surface of a rich body of enzyme chemistry.

Acknowledgements. This research was supported by the National Institutes of Health Grants GM-65324 and GM-41239 and the Robert A. Welch Foundation (F-1334). G.J.P. was supported by a VIDI grant from the Division of Chemical Sciences of the Netherlands Organization of Scientific Research (NWO-CW).

1 Murzin A. G. (1996) Structural classification of proteins: new superfamilies. Curr. Opin. Struct. Biol. 6, 386–394.
2 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.
3 Poelarends G. J. and Whitman C. P. (2004) Evolution of enzymatic activity in the tautomerase superfamily: mechanistic and structural studies of the 1,3-dichloropropene catabolic enzymes. Bioorg. Chem. 32, 376–392.
4 Jenkins J. R. and Cooper R. A. (1988) Molecular cloning, expression, and analysis of the genes of the homoprotocatechuate catabolic pathway of *Escherichia coli* C. J. Bacteriol. 170, 5317–5324.
5 Whitman, C. P. (1999) Keto-enol tautomerisation in enzymatic reactions. In: Comprehensive Natural Products Chemistry, Barton, D. & Nakanishi, K. (eds.), vol. 5, pp. 31–50, Poulter, C.D. (ed.), Elsevier, Oxford.
6 Suzuki M., Sugimoto H., Nakagawa A., Tanaka I., Nishihira J. and Sakai M. (1996) Crystal structure of the macrophage migration inhibitory factor from rat liver. Nat. Struct. Biol. 3, 259–266.
7 Sun H.-W., Bernhagen J., Bucala R. and Lolis E. (1996) Crystal structure at 2.6 Å resolution of human macrophage migration inhibitory factor. Proc. Natl. Acad. Sci. 93, 5191–5196.
8 Rosengren E., Aman P., Thelin S., Hansson C., Ahlfors S., Bjork P., Jacobsson L. and Rorsman H. (1997) The macrophage migration inhibitory factor MIF is a phenylpyruvate tautomerase. FEBS Lett. 417, 85–88.
9 Lubetsky J. B., Swope M., Dealwis C., Blake P. and Lolis E. (1999) Pro-1 of macrophage migration inhibitory factor functions as a catalytic base in the phenylpyruvate tautomerase activity. Biochemistry 38, 7346–7354.
10 Poelarends G. J., Serrano H., Person M. D., Johnson Jr. W. H., Murzin A. G. and Whitman C.P. (2004) Cloning, expression, and characterization of a *cis*-3-chloroacrylic acid dehalogenase: insights into the mechanistic, structural, and evolutionary relationship between isomer-specific 3-chloroacrylic acid dehalogenases. Biochemistry 43, 759–772.
11 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.
12 Almrud J. J., Kern A. D., Wang S. C., Czerwinski R.M., Johnson Jr. W. H., Murzin A. G., Hackert M. L. and Whitman C. P. (2002) The crystal structure of YdcE, a 4-oxalocrotonate tautomerase homologue from *Escherichia coli*, confirms the structural basis for oligomer diversity. Biochemistry 41, 12010–12024.
13 Harayama S., Lehrbach P. R. and Timmis K. N. (1984) Transposon mutagenesis analysis of *meta*-cleavage operon genes of the TOL plasmid of *Pseudomonas putida* mt-2. J. Bacteriol. 160, 251–255.
14 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.
15 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. Microb. 137, 2025–2032.
16 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.
17 Poelarends G. J., Wilkens M., Larkin M. J., van Elsas J. D. and Janssen D. B. (1998) Degradation of 1,3-dichloropropene by *Pseudomonas cichorii* 170. Appl. Environ. Microbiol. 64, 2931–2936.
18 Whitman C. P., Aird B. A., Gillespie W. R. and Stolowich, N. J. (1991) Chemical and enzymatic ketonization of 2-hydroxymuconate, a conjugated enol. J. Am. Chem. Soc. 113, 3154–3162.
19 Whitman C. P., Hajipour G., Watson R. J., Johnson Jr. W. H., Bembenek M. E. and Stolowich N. J. (1992) Stereospecific ketonization of 2-hydroxymuconate by 4-oxalocrotonate tautomerase and 5-(carboxymethyl)-2-hydroxymuconate isomerase. J. Am. Chem. Soc. 114, 10104–10110.
20 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.
21 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.
22 Stivers J. T., Abeygunawardana C., Mildvan A. S., Hajipour G., Whitman C. P. and Chen L. H. (1996) The catalytic role of the amino-terminal proline in 4-oxalocrotonate tautomerase: affinity labeling and heteronuclear NMR studies. Biochemistry 35, 803–813.
23 Stivers J. T., Abeygunawardana C., Mildvan A. S., Hajipour G. and Whitman C. P. (1996) 4-Oxalocrotonate tautomerase: pH dependence of catalysis and pK_a values of active site residues. Biochemistry 35, 814–823.
24 Wang S. C, Johnson Jr. W. H., Czerwinski R. M., Stamps S. L. and Whitman C. P. (2007) Kinetic and stereochemical analysis of YwhB, a 4-oxalocrotonate tautomerase homologue in *Bacillus subtilis:* mechanistic implications for the YwhB- and 4-oxalocrotonate tautomerase-catalyzed reactions. Biochemistry 46, 11919–11929.
25 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.
26 Horowitz N. H. (1945) On the evolution of biochemical syntheses. Proc. Natl. Acad. Sci. USA, 31, 153–157.
27 Jensen R. A. (1976) Enzyme recruitment in evolution of new function. Ann. Rev. Microbiol. 30, 409–25.
28 Hughes A. L. (1994) The evolution of functionally novel proteins after gene duplication. Proc. R. Soc. London Ser. B 256, 119–124.
29 Hughes A. L. (2005) Gene duplication and the origin of novel proteins. Proc. Natl. Acad. Sci. USA, 102, 8791–8792.
30 Perona J. J. and Craik C. S. (1997) Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. J. Biol. Chem. 272, 29987–29990.
31 Babbitt P. C. and Gerlt J. A. (1997) Understanding enzyme superfamilies. J. Biol. Chem. 272, 30591–30594.
32 Neidhart D. J., Kenyon G. L., Gerlt J. A. and Petsko, G. A. (1990) Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. Nature 347, 692–694.
33 Babbitt P. C. and Gerlt J. A. (2001) New functions from old scaffolds: how nature reengineers enzymes for new functions. Adv. in Protein Chem. 55, 1–28.
34 Holden H. M., Benning M. M., Haller T. and Gerlt J. A. (2001) The crotonase superfamily: divergently related enzymes that catalyze different reactions involving acyl coenzyme A thioesters. Acc. Chem. Res. 34, 145–157.
35 Gerlt J. A. and Raushel F. M. (2003) Evolution of function in $(\beta/\alpha)_8$-barrel enzymes. Curr. Opin. Chem. Biol. 7, 252–264.
36 Gerlt J. A., Babbitt P. C. and Rayment I. (2005) Divergent evolution in the enolase superfamily: the interplay of mechanism and specificity. Arch. Biochem. Biophys. 433, 59–70.
37 O'Brien P. J. and Herschlag D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. Chem. Biol. 6, R91-R105.
38 O'Brien P. J. and Herschlag D. (2001) Functional interrelationships in the alkaline phosphatase superfamily: phosphodiesterase activity of *Escherichia coli* alkaline phosphatase. Biochemistry 40, 5691–5699.
39 Copley S. D. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. Curr. Opin. Chem. Biol. 7, 265–272.
40 Yoshikuni Y., Ferrin T. E. and Keasling J. D. (2006) Designed divergent evolution of enzyme function. Nature 440, 1078–1082.
41 Roodveldt C. and Tawfik D. S. (2005) Shared promiscuous activities and evolutionary features in various members of the amidohydrolase superfamily. Biochemistry 44, 12728–12736.
42 Aharoni A., Gaidukov L., Khersonsky O., McQ Gould S., Roodveldt C. and Tawfik D. S. (2005) The "evolvability" of promiscuous protein functions. Nat. Genet. 37, 73–76.
43 Wang S. C., Person M. D., Johnson Jr. W. H. 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.
44 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.
45 Taylor A. B., Czerwinski R. M., Johnson Jr. 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.
46 de Jong R. M., Brugman W., Poelarends G. J., Whitman C. P. and Dijkstra B. W. (2004) The X-ray structure of *trans*-3-chloroacrylic acid dehalogenase reveals a novel hydration mechanism in the tautomerase superfamily. J. Biol. Chem. 279, 11546–11552.
47 Azurmendi H. F., Wang S. C., Massiah M. A., Poelarends G. J., Whitman C. P. and Mildvan A. S. (2004) The roles of active-site residues in the catalytic mechanism of *trans*-3-chloroacrylic acid dehalogenase: a kinetic, NMR, and mutational analysis. Biochemistry 43, 4082–4091.
48 Poelarends G. J., Serrano H., Johnson Jr. W. H. and Whitman C. P. (2004) Stereospecific alkylation of *cis*-3-chloroacrylic acid dehalogenase by (*R*)-oxirane-2-carboxylate: analysis and mechanistic implications. Biochemistry 43, 7187–7196.
49 de Jong R. M., Bazzacco P., Poelarends G. J., Johnson Jr. W. H., Kim Y.-J., Burks E. A., Serrano H., Whitman C. P. and Dijkstra B. W. (2007) Crystal structures of wild type and inactivated *cis*-3-chloroacrylic acid dehalogenase: the structural basis for substrate specificity and inactivation by (*R*)-oxirane-2-carboxylate. J. Biol. Chem. 282, 2440–2449.
50 Poelarends G. J., Serrano H., Johnson Jr. W. H., Hoffman D. W. and Whitman C. P. (2004) The hydratase activity of malonate semialdehyde decarboxylase: mechanistic and evolutionary implications. J. Am. Chem. Soc. 126, 15658–15659.
51 Poelarends G. J., Serrano H., Johnson Jr. W. H. and Whitman C. P. (2005) Inactivation of malonate semialdehyde decarboxylase by 3-halopropiolates: evidence for hydratase activity. Biochemistry 44, 9375–9381.
52 Almrud J. J., Poelarends G. J., Johnson Jr. W. H., Serrano H., Hackert M. L. and Whitman C. P. (2005) Crystal structures of the wild-type, P1A mutant, and inactivated malonate semialdehyde decarboxylase: a structural basis for the decarboxylase and hydratase activities. Biochemistry 44, 14818–14827.
53 Wang S. C, Johnson Jr. W. H. and Whitman C. P. (2003) The 4-oxalocrotonate tautomerase- and YwhB-catalyzed hydration of 3*E*-haloacrylates: implications for the evolution of new enzymatic activities. J. Am. Chem. Soc. 125, 14282–14283.
54 Poelarends G. J., Almrud J. J., Serrano H., Darty J. E., Johnson Jr. W. H., Hackert M. L. and Whitman C. P. (2006) Evolution of enzymatic activity in the tautomerase superfamily: mechanistic and structural consequences of the L8R mutation in 4-oxalocrotonate tautomerase. Biochemistry 45, 7700–7708.
55 Poelarends G. J., Serrano H., Johnson Jr. W. H. and Whitman, C. P (2007) The phenylpyruvate tautomerase activity of *trans*-3-chloroacrylic acid dehalogenase: evidence for an enol intermediate in the dehalogenase reaction? Biochemistry 46, 9596–9604.
56 Poelarends G. J., Serrano H., Person M. D., Johnson Jr. W. H. and Whitman C. P. (2008) Characterization of Cg10062 from *Corynebacterium glutamicum:* implications for the evolution of *cis*-3-chloroacrylic acid dehalogenase activity in the tautomerase superfamily. Biochemistry 47, 8139–8147.
57 DeLano, W. L. (2002) The PyMOL molecular graphics system, DeLano Scientific, San Carlos, CA, USA. (http://www.pymol.org)