Exploring chemical versatility within the tautomerase superfamily
Baas, Jan

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2014

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Characterization of a newly identified Mycobacterial tautomerase with promiscuous dehalogenase and hydratase activities reveals a functional link to a recently diverged cis-3-chloroacrylic acid dehalogenase.

Bert-Jan Baas, Ellen Zandvoort, Anna A. Wasiel, Wim J. Quax, and Gerrit J. Poelarends

Department of Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands

Published in *Biochemistry* (2011) 50, 2889-2899.
Abstract

The enzyme *cis*-3-chloroacrylic acid dehalogenase (*cis*-CaaD) is found in a bacterial pathway that degrades a synthetic nematocide, *cis*-1,3-dichloropropene, introduced in the 20th century. The previously determined crystal structure of *cis*-CaaD and its promiscuous phenylpyruvate tautomerase (PPT) activity link this dehalogenase to the tautomerase superfamily, a group of homologous proteins that are characterized by a catalytic amino-terminal proline and a β-α-β structural fold. The low-level PPT activity of *cis*-CaaD, which may be a vestige of the function of its progenitor, prompted us to search the databases for a homologue of *cis*-CaaD that was annotated as a putative tautomerase and test both its PPT and *cis*-CaaD activity. We identified a mycobacterial *cis*-CaaD homologue (designated MsCCH2) that shares key sequence and active site features with *cis*-CaaD. Kinetic and $^1$H NMR spectroscopic studies show that MsCCH2 functions as an efficient PPT and exhibits low-level promiscuous dehalogenase activity, processing both *cis-* and *trans*-3-chloroacrylic acid. To further probe the active site of MsCCH2, the enzyme was incubated with 2-oxo-3-pentynoate (2-OP). At pH 8.5, MsCCH2 is inactivated by 2-OP due to the covalent modification of Pro-1, suggesting that Pro-1 functions as a nucleophile at pH 8.5 and attacks 2-OP in a Michael-type reaction. At pH 6.5, however, MsCCH2 exhibits hydratase activity and converts 2-OP to acetopyruvate, which implies that Pro-1 is cationic at pH 6.5 and not functioning as a nucleophile. At pH 7.5, the hydratase and inactivation reactions occur simultaneously. From these results, it can be inferred that Pro-1 of MsCCH2 has a $pK_a$ value that lies in between that of a typical tautomerase ($pK_a$ of Pro-1 ~6) and that of *cis*-CaaD ($pK_a$ of Pro-1 ~9). The shared activities and structural features, coupled with the intermediate $pK_a$ of Pro-1, suggest that MsCCH2 could be characteristic of an evolutionary intermediate along the past route for the divergence of *cis*-CaaD from an unknown superfamily tautomerase. This makes MsCCH2 an ideal candidate for laboratory evolution of its promiscuous dehalogenase activity, which could identify additional features necessary for a fully active *cis*-CaaD. Such results will provide insight into pathways that could lead to the rapid divergent evolution of an efficient *cis*-CaaD enzyme.
Introduction

The bacterial enzymes trans-3-chloroacrylic acid dehalogenase (CaaD)\(^1\) and cis-3-chloroacrylic acid dehalogenase (cis-CaaD) catalyze the hydrolytic dehalogenation of the trans- and cis-isomers of 3-chloroacrylate (2 and 3, respectively) to yield malonate semialdehyde (4) and HCl (Scheme 1) (1-3). These reactions represent key steps in the degradation of the synthetic nematocide 1,3-dichloropropene (1), which was introduced into the environment in the 20th century (4,5). Both dehalogenases belong to the tautomerase superfamily and as such share a β-α-β structural fold and a catalytic amino-terminal proline (6-8). However, they are not members of the same family and likely have evolved independently from respective family progenitors (3,8). This is quite surprising given the fact that CaaD and cis-CaaD use similar catalytic mechanisms to process the different isomers of 3-chloroacrylate (9-14). Despite their most recent emergence, the evolutionary origins of CaaD and cis-CaaD remain elusive.

Scheme 1.

CaaD is a heterohexamer consisting of three α-subunits (75 amino acid residues each) and three β-subunits (70 amino acid residues each) (1). Similarities in sequence and tertiary and quaternary structures connect CaaD to members of the 4-oxalocrotonate tautomerase (4-OT) family within the tautomerase superfamily (1,9,15). 4-OT is part of a bacterial degradation pathway for aromatic hydrocarbons such as toluene (5) and converts 2-hydroxy-2,4-hexadienedioate (6) to 2-oxo-3-hexenedioate (7) (Scheme 2) (15-19). The evolutionary link between CaaD and members of the 4-OT family is strengthened by the observations that 4-OT from Pseudomonas putida mt-2 and YwhB, a 4-OT homologue from Bacillus subtilis, show low-level CaaD activity (20). This suggests that a 4-OT-like enzyme could have been recruited to serve as a CaaD because it fortuitously had a low but useful level of dehalogenase activity (7,20). Indeed, CaaD exhibits vestigial tautomerase activity and, like 4-OT and YwhB, converts phenylenolpyruvate (8) to phenylpyruvate (9) (Scheme 3) (21).

cis-CaaD, a trimer composed of three identical monomers (149 amino acid residues each), is the title enzyme of a separate family within the tautomerase superfamily (3). This dehalogenase also exhibits vestigial tautomerase activity, albeit at a low-level, converting 8 to 9 (21). Its closest family member described to date, Cg10062 from Corynebacterium glutamicum, is considerably homologous in sequence (34% identical and 53% similar) to cis-CaaD, and a pairwise alignment suggests that all key catalytic residues of cis-CaaD are conserved in Cg10062 (22). As expected, Cg10062 possesses dehalogenase activity, converting both 2 and 3 with a strong preference for 3 (22).
However, the enzyme exhibits no significant tautomerase activity using 8.² Hence, a functional link of shared promiscuous activities between *cis*-CaaD and a tautomerase from the same family has not yet been identified.

Scheme 2.

Scheme 3.

Scheme 4.
As promiscuous activities could provide important clues regarding the progenitor of a newly diverged family member (7,23,24), we searched the sequence databases for a homologue of cis-CaaD that was tentatively annotated as a putative tautomerase and tested both its phenylenolpyruvate tautomerase (PPT) and cis-CaaD activity. Herein, we present the characterization of a cis-CaaD homologue (designated MsCCH2) from Mycobacterium smegmatis strain MC2 155 that functions as an efficient tautomerase, having a \(k_{cat}/K_m\) value of 9.8 x 10^3 M^{-1} s^{-1} for 8, and has low-level promiscuous cis-CaaD and CaaD activity. Hydratase assays and labeling studies using 2-oxo-3-pentynoate (12, Scheme 4) indicate that Pro-1 of MsCCH2 has a \(pK_a\) value that lies in between that of a typical tautomerase (\(pK_a\) of Pro-1 ~6) and that of cis-CaaD (\(pK_a\) of Pro-1 ~9). These results suggest that MsCCH2 could be characteristic of an evolutionary intermediate along the past route for the divergence of cis-CaaD from an unknown homotrimeric tautomerase.

Results

Identification of cis-CaaD Homologues. A sequence similarity search in the NCBI microbial database was performed with the BLASTP program using the cis-CaaD amino acid sequence as the query. This search yielded several bacterial proteins that shared significant sequence identity with cis-CaaD. The top hits included two sequences from M. smegmatis strain MC2 155, designated MsCCH1 and MsCCH2, as well as the previously characterized Cg10062 protein from Corynebacterium glutamicum (Figure 1) (22). The mycobacterial protein MsCCH2 is annotated as a putative tautomerase and was selected for further study.

Results

Identification of cis-CaaD Homologues. A sequence similarity search in the NCBI microbial database was performed with the BLASTP program using the cis-CaaD amino acid sequence as the query. This search yielded several bacterial proteins that shared significant sequence identity with cis-CaaD. The top hits included two sequences from M. smegmatis strain MC2 155, designated MsCCH1 and MsCCH2, as well as the previously characterized Cg10062 protein from Corynebacterium glutamicum (Figure 1) (22). The mycobacterial protein MsCCH2 is annotated as a putative tautomerase and was selected for further study.

Figure 1. Amino acid sequence alignment of cis-CaaD from coryneform bacterium strain FG41 (3) and three selected cis-CaaD homologues. The NCBI reference sequences for the selected proteins are as follows: Cg10062 from Corynebacterium glutamicum: NP_599314.1; MsCCH1 from Mycobacterium smegmatis strain MC2 155: YP_887666.1; MsCCH2 from Mycobacterium smegmatis strain MC2 155: YP_885986.1. Identical residues are shaded in black. The six active site residues of cis-CaaD that have been implicated as critical residues for its activity (3,12) are indicated by an asterisk.
The mature MsCCH2 protein is predicted to be 138 amino acids in length, assuming excision of the initiating methionine residue. The genomic context of the gene encoding MsCCH2 does not provide any clues about the biological function of this protein in *M. smegmatis*. The sequence of MsCCH2 is 28% identical and 39% similar to that of *cis*-CaaD. Despite the low sequence identity, four of the six active site residues that have been implicated as critical residues for *cis*-CaaD activity (Pro-1, His-28, Arg-70 and Glu-114) are conserved in MsCCH2 (Pro-1, His-28, Arg-68 and Glu-112) (Figure 1) (3,12). The other two residues, Arg-73 and Tyr-103, are not conserved and replaced by His-71 and Ile-101 in MsCCH2. To obtain insight into the functional properties of this *cis*-CaaD homologue, MsCCH2 was overproduced, purified and subjected to kinetic and mechanistic characterization (as described below).

**Expression and Purification of MsCCH2.** The gene coding for MsCCH2 was amplified from genomic DNA of *M. smegmatis* strain MC2 155 and cloned into the expression vector pET20b(+), resulting in the construct pMsCCH2. The MsCCH2 encoding gene in pMsCCH2 is under transcriptional control of the T7 promoter, and the enzyme was produced constitutively in *E. coli* BL21(DE3) as a C-terminal hexahistidine fusion protein. The recombinant enzyme was purified by a one-step Ni-Sepharose chromatography protocol, which typically provides ~15 mg of homogeneous enzyme per liter of culture. The purified MsCCH2 was analyzed by electrospray ionization mass spectrometry (ESI-MS) and gel filtration chromatography. Analysis of MsCCH2 by ESI-MS showed one major peak corresponding to a mass of 16093 ± 1 Da. A comparison of this value to the calculated subunit mass (16223 Da) indicates that the initiating methionine is removed during posttranslational processing, which results in a protein with an N-terminal proline. The native molecular mass of MsCCH2 was estimated by gel filtration chromatography to be ~48 kDa, which suggests that the native enzyme is a homotrimeric protein.

**Tautomerase Activity of MsCCH2.** It has previously been determined that *cis*-CaaD exhibits promiscuous low-level tautomerase activity, converting phenylenolpyruvate (8) to phenylpyruvate (9) (Scheme 3) (21), which may be a vestige of the function of its progenitor. This observation prompted us to examine whether MsCCH2 catalyzes the tautomerization of 8. The results show that MsCCH2 converts 8 to 9 and that the activity (in terms of $k_{cat}$) is 170-fold higher than that observed for *cis*-CaaD (Table 1). A comparison of the kinetic parameters further shows that the $K_m$ value for MsCCH2 is significantly higher (32-fold) than that measured for *cis*-CaaD. Together, this results in a 5.4-fold higher $k_{cat}/K_m$. Additionally, MsCCH2 also catalyzes the conversion of p-hydroxyphenylenolpyruvate (10) to p-hydroxyphenylpyruvate (11) (Scheme 3) with a $k_{cat} = 3.2 \text{ s}^{-1}$ and a $K_m$ of 0.89 mM, resulting in a $k_{cat}/K_m$ of $3.5 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ (Table 1). The formation of 11 in the reaction of MsCCH2 with 10 was confirmed by $^1$H NMR spectroscopy (see Supporting Information).
**Table 1.** Kinetic Parameters for cis-CaaD, MsCCH2, and MsCCH2-P1A using Phenylenalpyruvate (8) and (p-Hydroxyphenyl)enalpyruvate (10).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat}/K_m$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MsCCH2$^a$</td>
<td>8</td>
<td>$34 \pm 4$</td>
<td>$3500 \pm 300$</td>
<td>$9.8 \times 10^7$</td>
</tr>
<tr>
<td>MsCCH2$^a$</td>
<td>10</td>
<td>$3.2 \pm 0.3$</td>
<td>$890 \pm 90$</td>
<td>$3.5 \times 10^3$</td>
</tr>
<tr>
<td>P1A-MsCCH2$^a$</td>
<td>8</td>
<td>$0.8 \pm 0.02$</td>
<td>$250 \pm 30$</td>
<td>$3.3 \times 10^3$</td>
</tr>
<tr>
<td>P1A-MsCCH2$^a$</td>
<td>10</td>
<td>$0.2 \pm 0.02$</td>
<td>$790 \pm 150$</td>
<td>$2.5 \times 10^2$</td>
</tr>
<tr>
<td>cis-CaaD$^a$</td>
<td>8</td>
<td>$0.2 \pm 0.03$</td>
<td>$110 \pm 30$</td>
<td>$1.8 \times 10^3$</td>
</tr>
</tbody>
</table>

$^a$ These steady-state kinetic parameters were measured in 10 mM Na$_2$HPO$_4$ buffer (pH 7.3) at 22 °C. $^b$ These kinetic data were obtained from Poelarends et al. (21). Errors are standard deviations.

**Dehalogenase Activity of MsCCH2.** Having established that MsCCH2 exhibits high-level tautomerase activity, we next investigated whether the enzyme exhibits promiscuous dehalogenase activity. Accordingly, MsCCH2 was incubated with cis-3-chloroacrylic acid (3) and trans-3-chloroacrylic acid (2) (Scheme 1), and the reactions were monitored by $^1$H NMR spectroscopy. After incubation of 3 with MsCCH2 in 100 mM phosphate buffer (pH 9.2) for 14 days at 22 °C, the intensity of the two signals corresponding to 3 (6.25 and 6.17 ppm) decreased and four new signals appeared. Two signals (9.53 and 2.10 ppm) correspond to acetaldehyde, whereas the other two signals (5.11 and 1.18 ppm) correspond to its hydrate. Integration of the signals indicates that ~10% of 3 has been converted to acetaldehyde (and its hydrate). In contrast to cis-CaaD, MsCCH2 also processes 2: ~13% of 2 has been converted to acetaldehyde (and its hydrate) after 14 days at 22 °C.

The MsCCH2-catalyzed conversions of 2 and 3 were also followed in 100 mM phosphate buffer at pH 6.5. Integration of the signals indicates that ~6% of 2 and ~3% of 3 has been converted to acetaldehyde (and its hydrate) after 2 weeks at 22 °C. No product formation was detected for control reactions without enzyme (2 or 3 incubated in 100 mM phosphate buffer, either pH 6.5 or pH 9.2, for 14 days at 22 °C), ruling out a nonenzymatic dehalogenation.

A previously described colorimetric assay, which monitors halide release, was used to measure kinetic parameters (Table 2) (1,29). MsCCH2 catalyzes the dehalogenation of 2 with a $k_{cat} = 4 \times 10^{-4}$ s$^{-1}$ and a $K_m$ of 96 mM, resulting in a $k_{cat}/K_m$ of $\sim 4 \times 10^{-3}$ M$^{-1}$ s$^{-1}$. Saturation with 3 was not achieved for MsCCH2 and therefore only the $k_{cat}/K_m$ value was determined. A comparison of this value to that measured for the dehalogenation of 2 shows that MsCCH2 is ~2.5-fold more efficient in catalyzing the dehalogenation of 2 than 3. In comparison to the cis-CaaD-catalyzed dehalogenation of 3, MsCCH2 shows a $7.5 \times 10^6$-fold lower $k_{cat}/K_m$. Hence, MsCCH2 exhibits low-level dehalogenase activity, accepting both isomers of 3-chloroacrylate as substrates with a slight preference for the trans-isomer.
Reaction of MsCCH2 with 2-Oxo-3-Pentynoate. Previous work has shown that the tautomerase 4-OT is irreversibly inactivated by 2-oxo-3-pentynoate (12) due to the covalent modification of Pro-1 (30), whereas the dehalogenases CaaD and cis-CaaD convert 12 to acetopyruvate (13) (Scheme 4) (2,3). These dissimilar reactions reflect differences in the ionization state and reactivity of Pro-1 in the tautomerase versus dehalogenase active sites. In order to probe the active site of MsCCH2, the reaction of this enzyme with 12 was examined and compared to the well-characterized 4-OT and CaaD reactions with 12.

### Table 2. Kinetic Parameters for cis-CaaD, Cg10062, and MsCCH2 Using cis-3-chloroacrylate (3) and trans-3-chloroacrylate (2).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>( k_{cat} ) (s(^{-1}))</th>
<th>( K_m ) (mM)</th>
<th>( k_{cat}/K_m ) (M(^{-1})s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>MsCCH2(^{a})</td>
<td>3</td>
<td>ND(^{b})</td>
<td>ND</td>
<td>1.6 \times 10^{-3}</td>
</tr>
<tr>
<td>MsCCH2(^{c})</td>
<td>2</td>
<td>(4.0 \pm 0.1) \times 10^{-4}</td>
<td>96 \pm 6</td>
<td>4.0 \times 10^{-3}</td>
</tr>
<tr>
<td>Cg10062(^{c})</td>
<td>3</td>
<td>1.6 \pm 0.3</td>
<td>156 \pm 42</td>
<td>10</td>
</tr>
<tr>
<td>Cg10062(^{c})</td>
<td>2</td>
<td>(2.0 \pm 1) \times 10^{-3}</td>
<td>54 \pm 40</td>
<td>4.0 \times 10^{-3}</td>
</tr>
<tr>
<td>cis-CaaD(^{c})</td>
<td>3</td>
<td>4.6 \pm 0.3</td>
<td>0.152 \pm 0.02</td>
<td>3.0 \times 10^{4}</td>
</tr>
</tbody>
</table>

\(^{a}\) The steady-state kinetic parameters were measured in 50 mM Tris-SO\(_4\) buffer (pH 8.0) at 22 °C. \(^{b}\) Not determined. \(^{c}\) These kinetic data were obtained from Poelarends et al. (22). Errors are standard deviations.

As expected, incubation of 12 with CaaD (at pH 7.3) resulted in a decrease in the absorbance at 234 nm, corresponding to 12, accompanied by the appearance of one new absorbance peak at 294 nm, which corresponds to 13 (Figure 2A) (2). Incubation of 12 with 4-OT (at pH 7.3) resulted in the rapid inactivation of 4-OT and the appearance of one new absorbance peak at 340 nm, which likely corresponds to the enamine species 14 (Scheme 4) (Figure 2B) (30). Surprisingly, the conversion of 12 by MsCCH2 (at pH 7.3) resulted in the appearance of two new absorbance peaks, one at 294 nm and the other at 324 nm (Figure 2C). From comparison to the product absorbance peaks formed in the reactions of CaaD and 4-OT with 12, it can be inferred that the peak observed at 294 nm corresponds to acetopyruvate (13), whereas the peak observed at 324 nm likely corresponds to an enamine species (14), which could result from the reaction between 12 and Pro-1 of MsCCH2. Importantly, the formation of 13 and covalently modified MsCCH2 (presumably 14) in the reaction of 12 with MsCCH2 was confirmed by ESI-MS and MS/MS analyses (see Supporting Information).

Identification of the Modified Active Site Residue by Mass Spectrometry. In order to identify the site of modification, MsCCH2 was treated with 12 at pH 7.3, reduced with NaBH\(_4\), and analyzed by ESI-MS. The reconstruct of the ESI mass spectrum revealed two major peaks, one corresponding to the mass of unmodified MsCCH2 (16094 Da) and the other to the mass expected for the enzyme modified by a single molecule of 12 and reduced by NaBH\(_4\) (16209 Da). This MsCCH2 sample was then digested with
endoproteinase Glu-C (31), and the resulting peptide mixture was analyzed by MALDI-TOF-MS. A comparison of the peaks of the modified MsCCH2 component to those of the unmodified MsCCH2 component in the peptide mixture revealed a single modification by a species having a mass of 116 Da on the fragments Pro-1 to Glu-15 and Pro-1 to Glu-23 (Table 3). Analysis of the remaining peaks showed no modification of other fragments. These data indicate that a single site on the enzyme has been modified and that the site of modification is localized within the amino-terminal fragment Pro-1 to Glu-15.

Figure 2. UV spectra monitoring the dissimilar reactions of CaaD, 4-OT and MsCCH2 with 2-oxo-3-pentynoate (12, 400 µM) in 10 mM phosphate buffer at pH 7.3. (A) UV spectra monitoring the CaaD-catalyzed conversion of 12 ($\lambda_{\text{max}}$ = 234 nm) to acetopyruvate (13, $\lambda_{\text{max}}$ = 294 nm) (2). Spectra were recorded every 2 minutes. (B) UV spectrum of 4-OT inactivated by 12. The absorbance at 234 nm corresponds to unreacted 12, while the absorbance peak at 340 nm likely corresponds to enamine 14 (Scheme 4), which results from the reaction between 12 and Pro-1 of 4-OT (30). (C) UV spectra monitoring the MsCCH2-catalyzed conversion of 12 ($\lambda_{\text{max}}$ = 234 nm). The product peak at 294 nm corresponds to 13, whereas the peak observed at 324 nm likely corresponds to enamine 14, which results from the reaction between 12 and Pro-1 of MsCCH2. Spectra were recorded every 40 minutes.
To determine the actual site of the single modification, the unmodified and modified Pro-1 to Glu-15 and Pro-1 to Glu-23 peptides were subjected to MALDI-TOF/TOF-MS analysis. Because the two Pro-1 to Glu-23 peptides gave a more complete fragmentation pattern than the two Pro-1 to Glu-15 peptides, the masses of the fragment ions observed in the TOF/TOF spectra of the modified and unmodified Pro-1 to Glu-23 peptides were further analyzed. The spectrum of the precursor ion corresponding to the unmodified Pro-1 to Glu-23 peptide displayed characteristic b-ions resulting from the peptide fragments PVYTVTMSRGTLN and PV. MALDI-TOF/TOF-MS analysis of the precursor ion corresponding to the modified Pro-1 to Glu-23 peptide revealed an increase in mass of 116 Da for these b-ions (Table 3). Because one of these fragment ions is generated by the dipeptide Pro-1 to Val-2, we conclude that the active site Pro-1 residue is the sole site of modification by 12.

Table 3. Identification of Pro-1 as the sole site of labeling by 12 using Glu-C digestion and MALDI-TOF-MS and MALDI-TOF/TOF-MS analyses.

<table>
<thead>
<tr>
<th>Peptide fragment</th>
<th>Calculated mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVYTVTMSRGTLNGE</td>
<td>1624.8</td>
<td>1625.5</td>
</tr>
<tr>
<td>C&lt;sub&gt;6&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;O&lt;sub&gt;3&lt;/sub&gt;-PVYTVTMSRGTLNGE</td>
<td>1740.8</td>
<td>1741.6</td>
</tr>
<tr>
<td>PVYTVTMSRGTLNGETKAALAAE</td>
<td>2380.2</td>
<td>2381.8</td>
</tr>
<tr>
<td>PVYTVTMSRGTLN</td>
<td>1420.7</td>
<td>1421.9</td>
</tr>
<tr>
<td>PV</td>
<td>197.1</td>
<td>197.1</td>
</tr>
<tr>
<td>C&lt;sub&gt;6&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;O&lt;sub&gt;3&lt;/sub&gt;-PVYTVTMSRGTLNGETKAALAAE</td>
<td>2496.2</td>
<td>2497.8</td>
</tr>
<tr>
<td>C&lt;sub&gt;6&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;O&lt;sub&gt;3&lt;/sub&gt;-PVYTVTMSRGTLN</td>
<td>1536.7</td>
<td>1537.9</td>
</tr>
<tr>
<td>C&lt;sub&gt;6&lt;/sub&gt;H&lt;sub&gt;8&lt;/sub&gt;O&lt;sub&gt;3&lt;/sub&gt;-PV</td>
<td>313.1</td>
<td>313.2</td>
</tr>
</tbody>
</table>

<sup>a</sup> These values are calculated using the average molecular mass.  <sup>b</sup> This value corresponds to the total mass of the parent ion.  <sup>c</sup> The reported mass corresponds to the b-ion.

Effect of pH on Product Formation in Incubations of MsCCH2 with 2-Oxo-3-Pentynoate. To probe the ionization state of Pro-1, the reaction of MsCCH2 with 12 was examined at three different pH values (Figures 3 and 4). Incubation of MsCCH2 with 12 at pH 6.5 resulted in the formation of 13, but no absorbance peak at 324 nm (indicative of covalently modified MsCCH2; that is, the presumed enamine species 14) was observed (Figure 3A). Indeed, analysis of this incubation mixture by ESI-MS showed one major peak corresponding to the mass (16093) of unmodified MsCCH2 (Figure 4A). These observations suggest that MsCCH2 exhibits hydratase activity at pH 6.5, converting 12 to 13, which implies that Pro-1 is cationic at pH 6.5 and not functioning as nucleophile. However, incubation of 12 with MsCCH2 at pH 8.5 did not result in formation of 13, but only resulted in the formation of covalently modified MsCCH2, as indicated by the absorbance peak at 324 nm (Figure 3C). As expected, analysis of this incubation mixture by ESI-MS showed one major peak corresponding to the mass (16206) of modified
MsCCH2 (Figure 4C). This suggests that at pH 8.5, Pro-1 has the correct protonation state to be able to act as a nucleophile and attacks 12 to form a covalently modified enzyme (14). Consistent with the preceding experiment at pH 7.3, incubation of 12 with MsCCH2 at pH 7.5 resulted in the formation of both 13 and covalently modified enzyme (Figures 3B and 4B). Assuming a pKₐ of Pro-1 of about 7.5, ~50% of the enzyme would be in the correct protonation state to react with 12 to give 13, and ~50% of the enzyme would be in the correct protonation state to attack 12 to give 14.

**Figure 3.** UV spectra monitoring the reaction of MsCCH2 with 2-oxo-3-pentynoate (12, 400 µM) in 10 mM phosphate buffer at three different pH values. (A) Spectra monitoring the reaction at pH 6.5. Spectra were recorded every 40 minutes. (B) Spectra monitoring the reaction at pH 7.5. Spectra were recorded every 40 minutes. (C) Spectra monitoring the reaction at pH 8.5. Spectra were recorded every 5 minutes. The decrease in absorbance at 234 nm corresponds to the loss of 12, whereas the increases in absorbance at 294 nm and 324 nm correspond to the formation of 13 and 14, respectively.
Figure 4. ESI-MS analysis of the MsCCH2 protein in the three incubation mixtures described in Figure 3. (A) ESI-MS analysis of MsCCH2 after incubation with 12 at pH 6.5; see Figure 3A. (B) ESI-MS analysis of MsCCH2 after incubation with 12 at pH 7.5; see Figure 3B. (C) ESI-MS analysis of MsCCH2 after incubation with 12 at pH 8.5; see Figure 3C. The major peaks in the spectra correspond to either unmodified MsCCH2 (16093 Da) or MsCCH2 modified by a single molecule of 12 (16206 Da). The additional minor peaks at the right of each major peak result from (one or two) sodium adducts.

Kinetic Properties of the P1A Mutant of MsCCH2. The results above show that incubation of MsCCH2 with 12 (at pH 8.5) leads to the modification of Pro-1 and the concomitant loss of hydratase activity. To further assess the importance of Pro-1 for the activities of MsCCH2, the P1A mutant was constructed, overproduced in E. coli BL21(DE3) as a C-terminal hexahistidine fusion protein, and purified to homogeneity using the metal affinity chromatography protocol developed for wild-type MsCCH2. ESI-MS analysis revealed a mass of 16067 Da (calc. 16066 Da for the enzyme without the initiating methionine) for the P1A mutant, indicating that the initiating methionine is removed during posttranslational processing. The kinetic properties of the P1A mutant
were analyzed using 2, 3, 8, 10, and 12 as substrates, and compared to those measured for the wild-type enzyme.

For 8, the P1A mutant showed a 41-fold decrease in $k_{\text{cat}}$ and a 14-fold decrease in $K_m$, resulting in a ~3-fold decrease in $k_{\text{cat}}/K_m$ (Table 1). For 10, the P1A mutant showed a 16-fold decrease in $k_{\text{cat}}$, whereas the $K_m$ was not significantly affected. Hence, this resulted in a 16-fold decrease in $k_{\text{cat}}/K_m$. The P1A mutant had no detectable dehalogenase activity towards 2 and 3 using the colorimetric assay, and no detectable activity towards 12 (at pH 7.3) using the UV spectroscopic assay (i.e., the reaction mixture showed no product absorbance peaks). The CaaD and cis-CaaD activities of the P1A mutant were also assessed by $^1$H NMR spectroscopy after a lengthy incubation period (14 days at 22 °C) in 100 mM phosphate buffer (pH 9.2). However, the reaction mixtures showed no product. These findings are not surprising in view of the low-level dehalogenase activities of the wild-type enzyme. Taken together, the results suggest that Pro-1 is critical for the tautomerase, dehalogenase, and hydratase activities of MsCCH2.

Discussion

The mycobacterial enzyme (MsCCH2) characterized in this study proficiently tautomerizes phenylenolpyruvate (8) to phenylpyruvate (9) and $p$-hydroxyphenylenolpyruvate (10) to $p$-hydroxyphenylpyruvate (11) (Scheme 3). MsCCH2 catalyzes the conversion of 8 to 9 with a $k_{\text{cat}}/K_m$ of $9.8 \times 10^3$ M$^{-1}$ s$^{-1}$, which is somewhat lower (~40-fold) than the values measured for the same conversion catalyzed by the tautomerasers 4-OT ($k_{\text{cat}}/K_m = 3.7 \times 10^5$ M$^{-1}$ s$^{-1}$) (32) and YwhB ($k_{\text{cat}}/K_m = 4.2 \times 10^5$ M$^{-1}$ s$^{-1}$) (19). The lower catalytic efficiency of the MsCCH2-catalyzed reaction results mainly from a higher $K_m$ value, which suggests suboptimal binding of 8 in the active site of MsCCH2. A comparison of the kinetic parameters for the MsCCH2 ($k_{\text{cat}} = 3.2 \pm 0.3$ s$^{-1}$, $K_m = 890 \pm 90$ µM) and YwhB ($k_{\text{cat}} = 4.1 \pm 0.4$ s$^{-1}$, $K_m = 160 \pm 22$ µM) (19) catalyzed conversion of 10 to 11 shows similar catalytic efficiencies for both enzymes. It is important to emphasize that we have also tested compound 6, the native substrate of 4-OT (Scheme 2) (18,19), as potential substrate for MsCCH2, but no detectable 1,5-keto-enol tautomerase activity was observed. These results indicate that MsCCH2 functions as an effective 1,3-keto-enol tautomerase, converting 8 to 9 and 10 to 11, given that the physiological substrate for this tautomerase has yet to be identified.

Similarities in sequence and quaternary structure firmly link MsCCH2 to the cis-CaaD family, which is one of the five known families in the tautomerase superfamily (3,12). With the previously studied family members being cis-CaaD (3) and Cg10062 (22), MsCCH2 is the first known family member that exhibits a pronounced tautomerase activity. All known tautomerase superfamily members are characterized by a catalytic amino-terminal proline (6,7). The importance of the conserved Pro-1 in MsCCH2 for its tautomerase activity was demonstrated by mutating this residue to an alanine. For both substrates (8 and 10), the major effect of this mutation is observed in $k_{\text{cat}}$ (Table 1). Hence, the function of Pro-1 is predominantly catalytic. In the well-studied tautomerasers 4-OT and YwhB, the catalytic Pro-1 functions as the general base and is responsible for proton transfer (19,33-35). It is therefore reasonable to suggest that
Pro-1 in MsCCH2 also functions as the general base and may abstract the C-2 enol proton (of \(8\) or \(10\)) for delivery to the C-3 position (yielding \(9\) or \(11\)).

MsCCH2 also exhibits dehalogenase activity using trans-3-chloroacrylate (\(2\)) and cis-3-chloroacrylate (\(3\)) as substrates (Scheme 1), but this activity is much lower (10^6-fold) compared to its tautomerase activity (as assessed by \(k_{\text{cat}}/K_{\text{m}}\) values). Hence, the tautomerase activity is the primary activity of MsCCH2. A comparison of the kinetic parameters for the conversion of \(3\) to \(4\) shows that the catalytic efficiency (\(k_{\text{cat}}/K_{\text{m}}\)) of MsCCH2 is 10^4- and 10^7-fold lower than that of Cg10062 and cis-CaaD (Table 2) (22). Unlike cis-CaaD, MsCCH2 also catalyzes the conversion of \(2\) to \(4\) but with a catalytic efficiency that is 10-fold lower than that measured for Cg10062 (Table 2). Replacement of Pro-1 by an alanine completely abolished the dehalogenase activity of MsCCH2, indicating that dehalogenation, like tautomerization, is an active site process that involves Pro-1.

Although MsCCH2 has a low-level dehalogenase activity, the rates of dehalogenation are still significant in comparison with the reported nonenzymatic rate of \(\sim 2.2 \times 10^{-12} \text{ s}^{-1}\) at 25 °C and pH 7 (36). Using this value for the spontaneous, uncatalyzed dehalogenation of \(2\) and the \(k_{\text{cat}}\) value for the MsCCH2-catalyzed dehalogenation of \(2\) (Table 2), it can be estimated that MsCCH2 affords a \(\sim 2 \times 10^8\)-fold rate enhancement. Although the individual \(k_{\text{cat}}\) value for the MsCCH2-catalyzed dehalogenation of \(3\) could not be determined, a similar 10^8-fold rate enhancement can be assumed. For comparison, cis-CaaD, which is thought to have evolved for the purpose of degrading \(3\) (3,6), affords a \(\sim 2 \times 10^{12}\)-fold rate enhancement for the conversion of \(3\) to \(4\).

On the basis of experimental studies (3,11-13), a mechanism for the cis-CaaD-catalyzed conversion of \(3\) to \(4\) was proposed (Scheme 5). The first step in catalysis is the nucleophilic attack of an activated water molecule on C-3 of \(3\) to form an enediolate intermediate. The enzyme presumably uses the side chains of two residues, Glu-114 and Tyr-103, to activate the nucleophilic water molecule. The side chains of His-28, Arg-70, and Arg-73 interact with the C-1 carboxylate group of \(3\). These interactions position and polarize the substrate to facilitate the attack of water and stabilize the resulting enediolate species. Tautomerization of the enediolate intermediate may be assisted by Pro-1, which is thought to place a proton at the C-2 atom, and generates an unstable chlorohydrin intermediate (Scheme 5a). Collapse of the proposed chlorohydrin to afford \(4\) and HCl could be an enzymatic or a nonenzymatic process. Alternatively, the enediolate species can undergo an elimination reaction to release the chlorine as chloride ion, followed by tautomerization of the enol intermediate and protonation at C-2 by Pro-1, yielding \(4\) (Scheme 5b).

In the context of this proposed cis-CaaD mechanism, the much lower dehalogenase activity of MsCCH2 (compared to cis-CaaD) could potentially be explained by active site differences. Sequence comparisons suggested that four of the six active site residues that have been implicated as critical residues for cis-CaaD activity (Pro-1, His-28, Arg-70 and Glu-114) are conserved in MsCCH2 (Pro-1, His-28, Arg-68 and Glu-112) (Figure 1) (3,12). However, two active site residues, Arg-73 and Tyr-103, are replaced by His-71 and
Ile-101 in MsCCH2. The replacement of Arg-73 in cis-CaaD by His-71 in MsCCH2 may result in less efficient binding and polarization of 3, thereby affecting water addition. The replacement of Tyr-103 in cis-CaaD by Ile-101 in MsCCH2 may even be more significant because it eliminates one of the presumed water-activating residues. These active site differences may reflect the different reaction specificities of MsCCH2 and cis-CaaD, and suggest that MsCCH2 could be only a few mutations away from being a highly specific and efficient cis-CaaD. Hence, mutation of His-71 to an arginine accompanied by the replacement of Ile-101 with a tyrosine might increase the dehalogenase activity and decrease the tautomerase activity of MsCCH2. The consequences of these mutations are currently being examined.

Another potential explanation for the lower dehalogenase activity and increased tautomerase activity of MsCCH2 could be a lowered pK_a of its Pro-1 residue when compared to the pK_a of Pro-1 (~9.3) in cis-CaaD (11). Support for this view comes from studies with 2-oxo-3-pentynoate (12, Scheme 4). Previous work has shown that the reactions of 4-OT and CaaD, the best characterized members of the 4-OT family in the tautomerase superfamily, with 12 reflect both the ionization state of Pro-1 (neutral versus cationic) and the environment of the active site (2,30). Whereas 4-OT is irreversibly inactivated by 12 due to the covalent modification of Pro-1, CaaD converts 12 to acetopyruvate (13) (Scheme 4). These dissimilar reactions reflect differences between the two active sites. Pro-1 of 4-OT has a pK_a of ~6.4, enabling it to function as a general base catalyst in its physiological tautomerase activity (conversion of 6 to 7, Scheme 2) (33,34). Hence, at neutral pH, Pro-1 functions as a nucleophile and attacks C-4 of 12 in a Michael-type reaction, presumably yielding 14 (Scheme 4) (30). In contrast, the pK_a of Pro-1 in CaaD is ~9.2, enabling it to function as a general acid catalyst in its physiological dehalogenase activity (conversion of 2 to 4, Scheme 1) (10). Thus, a Michael-type reaction between the amino group of proline and 12 is not favored, and in
the active site of CaaD, which has evolved to carry out water addition, 12 is processed to 13 (2). Like CaaD, cis-CaaD and Cg10062 process 12 to 13, instead of being inactivated by 12, which implies that their Pro-1 residues have a pK\textsubscript{a} comparable to that determined for Pro-1 of CaaD (3,22).

In view of these observations, the reaction of MsCCH2 with 12 was examined at different pH values. At pH 8.5, MsCCH2 is inactivated by 12 due to the covalent modification of Pro-1. This suggests that Pro-1 is neutral at pH 8.5 and functions as a nucleophile, attacking C-4 of 12 in a Michael-type reaction. The identification of 15 by mass spectrometry experiments, coupled with the observed \(\lambda_{\text{max}}\) of 324 nm for the reaction product, strongly suggests that the reaction results in the formation of 14. Indeed, reduction with NaBH\textsubscript{4} results in the complete loss of the absorbance at 324 nm, consistent with the rearrangement of 14 into the corresponding imine, which is reduced by NaBH\textsubscript{4} to give 15. At pH 6.5, however, MsCCH2 exhibits hydratase activity and converts 12 to 13. This implies that Pro-1 is cationic at pH 6.5 and not functioning as a nucleophile. At pH 7.3-7.5, MsCCH2 converts 12 to both 13 and 14, indicating that the hydratase and inactivation reactions occur simultaneously. These observations, coupled with the results of control reactions using 12 and 4-OT or CaaD at pH 7.3, suggest that Pro-1 of MsCCH2 has a pK\textsubscript{a} value that lies in between that of a typical tautomerase like 4-OT (pK\textsubscript{a} of Pro-1 ~6.4) and that of the dehalogenases CaaD and cis-CaaD (pK\textsubscript{a} of Pro-1 ~9.2). This makes MsCCH2 the first known tautomerase superfamily member with such an intermediate pK\textsubscript{a} value for its catalytic Pro-1 residue. This intermediate pK\textsubscript{a} value, estimated to be ~7.5, may enable Pro-1 to serve (at cellular pH) as a general base catalyst in its primary tautomerase activity and as a general acid catalyst in its promiscuous dehalogenase activity.

Given that Pro-1 in CaaD and cis-CaaD has a pK\textsubscript{a} of ~9.2 (10,11), enabling it to function as a general acid catalyst in the physiological dehalogenase activity, an interesting route for the laboratory evolution of MsCCH2 into a more efficient dehalogenase can be proposed. The introduction of one or more hydrophilic residues could make the active site of MsCCH2 less hydrophobic, thereby making the active site more amenable for water addition as well as raising the pK\textsubscript{a} of Pro-1. This would increase the concentration of enzyme with Pro-1 in the correct protonation state to function as a general acid catalyst. The appropriate experiments to investigate this intriguing possibility are underway.

In conclusion, the identification of MsCCH2 as tautomerase with promiscuous dehalogenase and hydratase activities now establishes a functional link between the recently diverged cis-CaaD and a tautomerase of the same family. Hence, the major phenylpyruvate tautomerase activity of MsCCH2 comprises the promiscuous activity of cis-CaaD, and the primary dehalogenase activity of cis-CaaD comprises the promiscuous activity of MsCCH2. In addition, both enzymes possess significant hydratase activity. A reasonable route for the divergence of cis-CaaD from an unknown tautomerase is presented in Figure 5. The ancestral tautomerase may have been a specialist tautomerase with a low pK\textsubscript{a} value for its catalytic Pro-1 residue. Mutations in the
Figure 5. Proposed evolutionary pathway for the divergence of an ancestral tautomerase, which may have been a specialist, to cis-CaaD, a specialist dehalogenase that is selective for the cis-isomer of 3-chloroacrylate (3). The family members MsCCH2 (this study) and Cg10062 (22) could be characteristic of evolutionary intermediates 1 and 2, respectively. In the course of divergence, the pKₐ of the essential Pro-1 residue likely increased from ~6 (in the ancestral tautomerase) to ~7.5 (in the generalist intermediate 1) to ~9 (in the specialist dehalogenase), allowing this residue to efficiently serve as a general acid in the newly evolved dehalogenase activity. For clarity, and because the three-dimensional structures of MsCCH2 and Cg10062 are not reported yet, the β-α-β building block carrying the N-terminal proline is shown.

specialist may have raised the pKₐ of its Pro-1 residue to about 7.5. This may have introduced a low-level promiscuous dehalogenase activity in the enzyme, without significantly decreasing the original tautomerase activity. This thus yielded, in effect, a generalist intermediate (intermediate 1 in Figure 5) exhibiting both tautomerase and dehalogenase activity. MsCCH2 could be characteristic of this generalist intermediate, and may therefore resemble the progenitor of cis-CaaD, using its promiscuous dehalogenase activity as an essential starting point. Further mutation and selection for an increase of the promiscuous dehalogenase activity may have resulted in intermediate 2, and finally in the specialist dehalogenase, cis-CaaD. Cg10062, which has pronounced dehalogenase activity, but lacks isomer specificity (22), could be characteristic of intermediate 2. In the course of divergence, the pKₐ of the essential Pro-1 residue likely increased from ~6 (in the ancestral tautomerase) to ~7.5 (in the generalist intermediate 1) to ~9 (in intermediate 2 and the specialist dehalogenase), allowing this residue to efficiently serve as a general acid in the newly evolved dehalogenase activity. Although this route is highly speculative, laboratory evolution of the promiscuous dehalogenase activity of MsCCH2 could provide important insight into pathways that could lead to the rapid divergent evolution of an efficient cis-CaaD enzyme.
**Acknowledgements**

We thank Professor William R. Jacobs, Jr. (Howard Hughes Medical Institute, Chevy Chase, MD) for the kind gift of genomic DNA of *Mycobacterium smegmatis* strain MC2 155 and Professor Christian P. Whitman (University of Texas at Austin, TX) for the kind gift of 2-oxo-3-pentynoate. We thank Annie van Dam and Margot Jeronimus-Stratingh (University of Groningen) for their expert assistance in acquiring the MS spectra.

**Materials and methods**

**Materials.** The sources of the chemicals, biochemicals, components of buffers and media, PCR purification, gel extraction, and Miniprep kits, Ni-NTA sepharose, pre-packed PD-10 Sephadex G-25 columns, and the oligonucleotides, enzymes and reagents used in the molecular biology procedures are reported elsewhere (25). Phenylpyruvate, (p-hydroxyphenyl)-pyruvate, trans-3-chloroacrylic acid, and cis-3-chloroacrylic acid were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). 2-Oxo-3-pentynoate was a kind gift of Professor C.P. Whitman (University of Texas at Austin, TX).

**General Methods.** BLASTP searches of the National Center for Biotechnology Information (NCBI) databases were performed using the cis-CaaD amino acid sequence (GenBank: AAR00932.1) as the query sequence. Amino acid sequences were aligned using a version of the CLUSTALW multiple-sequence alignment routines available in the computational tools at the EMBL-EBI Website. Techniques for restriction enzyme digestions, ligation, transformation, and other standard molecular biology manipulations were based on methods described elsewhere (26) or as suggested by the manufacturer. The PCR was carried out in a DNA thermal cycler (model GS-1) obtained from Biolegio (Nijmegen, The Netherlands). DNA sequencing was performed by ServiceXS (Leiden, The Netherlands) or Macrogen (Seoul, Korea). Protein was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on gels containing 10% polyacrylamide. The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the method of Waddell (27). Kinetic data were obtained on a V-650 or V-660 spectrophotometer from Jasco (Ijsselstein, The Netherlands). The kinetic data were fitted by nonlinear regression data analysis using the Grafit program (Erithacus, Software Ltd., Horley, U.K.) obtained from Sigma Chemical Co. 1H NMR spectra were recorded on a Varian Inova 500 (500 MHz) spectrometer using a pulse sequence for selective presaturation of the water signal. Chemical shifts for protons are reported in parts per million scale (Δ scale) downfield from tetramethylsilane and are referenced to protium (H2O: Δ = 4.67). The native molecular mass of purified MsCCH2 was determined by Superdex 75 gel filtration using a fast protein liquid chromatography system according to the instructions provided with the HiLoad 16/60 Superdex 75 prep grade column (Pharmacia). The standard molecular mass markers were albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa) and ribonuclease A (13.7 kDa). ESI-MS and MS/MS spectra were recorded using an LCQ electrospray mass spectrometer (Applied Biosystems, Foster City, CA). MALDI-TOF and TOF/TOF mass spectra were recorded using a 4700 Proteomics Analyzer (Applied
Biosystems). The enzymes CaaD and 4-OT were purified according to a previously published procedure (28).

Construction of Expression Vectors for MsCCH2 and the MsCCH2-P1A Mutant, and Expression and Purification of the two His-Tagged Proteins. The experimental procedures used for the construction of expression vectors for MsCCH2 wild-type and the MsCCH2-P1A mutant, and the overproduction and purification of these two proteins are provided in the Supporting Information.

Enzyme Assays. The tautomerization activities of MsCCH2 and the P1A mutant were measured by monitoring the ketonization of 8 to 9 and 10 to 11 in 10 mM Na2HPO4 buffer (pH 7.3) at 22 °C. Stock solutions of 8 and 10 were generated by dissolving the appropriate amount of phenylpyruvic acid or (p-hydroxyphenyl)pyruvic acid in absolute ethanol. The crystalline free acid of phenylpyruvic acid or (p-hydroxyphenyl)pyruvic acid is exclusively the enol form. The ketonization of 8 to 9 was monitored by following the decrease in absorbance at either 310 nm (ε = 1400 M⁻¹ cm⁻¹) or 315 nm (ε = 300 M⁻¹ cm⁻¹) using substrate concentrations ranging from 0.05 to 8 mM. The ketonization of 10 to 11 was monitored by following the decrease in absorbance at 325 nm (ε = 1800 M⁻¹ cm⁻¹) using substrate concentrations ranging from 0.1 to 1.8 mM. An appropriate quantity of enzyme (from a 5 mg/mL stock solution) was diluted into the sodium phosphate buffer (1 mL in a cuvette), and the assay was initiated by the addition of a small aliquot (5-20 µL) of substrate (either 8 or 10) from a stock solution. At all substrate concentrations, the nonenzymatic rate was subtracted from the enzymatic rate of ketonization.

The dehalogenation activities of MsCCH2 and the P1A mutant were measured by following the dechlorination of 2 and 3 at 22 °C in 50 mM Tris-SO4 buffer (pH 8.0) using a colorimetric assay (1,29). An appropriate amount of enzyme (1 mg) was incubated with the desired concentration of 2 or 3 in 3 mL of the Tris-SO4 buffer. The concentrations of substrate in the assay ranged from 20 to 350 mM. Stock solutions (400 mM) of 2 and 3 were made up in 50 mM Tris-SO4 buffer. The pH of the stock solution was adjusted to 8.0. Chloride concentrations were measured colorimetrically at different time intervals.

1H NMR Spectroscopic Analysis of the Reaction of MsCCH2 with 2, 3 or 10. 1H NMR spectra monitoring the MsCCH2-catalyzed conversion of trans-3-chloroacrylate (2), cis-3-chloroacrylate (3), or (p-hydroxyphenyl)enolpyruvate (10) were recorded according to protocols reported elsewhere (2,3,25). The modifications to these protocols are provided in the Supporting Information.

UV Spectroscopic Analysis of the Reactions of CaaD, 4-OT, and MsCCH2 with 2-Oxo-3-Pentynoate (12). UV spectra monitoring the CaaD, 4-OT, and MsCCH2 reactions with 12 were recorded as follows. An appropriate amount of enzyme (45 µg of CaaD, 0.5 mg of 4-OT or 0.5 mg of MsCCH2) was incubated with 12 (400 µM) in 1 mL of 10 mM Na2HPO4 buffer at 22 °C, with the final pH of the incubation mixture ranging from 6.5 to 8.5. A fresh stock solution of 12 (40 mM) was made up in 10 mM Na2HPO4 buffer, and the pH was adjusted to 7.3. The hydration of 12 (λmax = 234 nm) was monitored by following the formation of 13 at 294 nm (2). The covalent modification of 4-OT or MsCCH2 by 12 was
monitored by following the formation of 14 (for 4-OT, $\lambda_{\max} = 340$ nm; for MsCCH2, $\lambda_{\max} = 324$ nm).

To determine the mass of (unmodified and modified) MsCCH2 after its incubation with 12 in 10 mM Na$_2$HPO$_4$ buffer at pH 6.5, 7.3, 7.5, or 8.5, a sample was withdrawn from the incubation mixture, diluted 10-fold in 5 mM NH$_4$HCO$_3$ buffer (pH 7.3), and directly analyzed by ESI-MS.

**Mass Spectral Analysis of the Product of the MsCCH2-Catalyzed Hydration of 2-Oxo-3-Pentynoate (12).** The experimental procedure used for the identification of 13 as the product of the MsCCH2-catalyzed hydration of 12 is provided in the Supporting Information.

**Mass Spectral Analysis of Modified MsCCH2 and Peptide Mapping.** The MsCCH2 sample was made up as follows. The enzyme (1 mg) was incubated with 12 (400 $\mu$M) in 1 mL of 10 mM NaH$_2$PO$_4$ buffer (pH 7.3) for 3 h at 22 °C. The formation of the covalent adduct (14) was monitored spectrophotometrically at 324 nm. An aliquot of a 500 mM stock solution of NaBH$_4$ in water was added to the sample to give a final NaBH$_4$ concentration of 25 mM, after which the mixture was incubated overnight at 22 °C. This protocol was repeated (3 times) until the absorbance at 324 nm was completely lost. Subsequently, the buffer was exchanged against 5 mM NH$_4$HCO$_3$ buffer (pH 7.3) using a pre-packed PD-10 Sephadex G-25 gel filtration column. An aliquot of this protein sample was directly analyzed by ESI-MS.

For the peptide mapping studies, a quantity (50 $\mu$g) of the MsCCH2 sample was vacuum-dried. The protein pellet was dissolved in 10 $\mu$L of 10 M guanidine-HCl and incubated for 2 h at 37 °C. Subsequently, the sample was diluted 10-fold by the addition of 90 $\mu$L of 100 mM NH$_4$HCO$_3$ buffer (pH 8.0) and incubated for 48 h at 37 °C with protease Glu-C (0.5 $\mu$L from a 10 mg/mL stock solution in water). The sample was analyzed by MALDI-TOF MS without further purification. Selected ions of modified and unmodified peptide fragments were subjected to TOF/TOF MS analysis.

**Footnotes**

1Abbreviations: CaaD, trans-3-chloroacrylic acid dehalogenase; cis-CaaD, cis-3-chloroacrylic acid dehalogenase; EBI, European Bioinformatics Institute; EMBL, European Molecular Biology Laboratory; ESI-MS, electrospray ionization mass spectrometry; MALDI-TOF, matrix assisted laser desorption/ionization time-of-flight; MsCCH, Mycobacterium smegmatis cis-CaaD homologue; NCBI, National Center for Biotechnology Information; NMR, nuclear magnetic resonance; 4-OT, 4-oxalocrotonate tautomerase; PCR, polymerase chain reaction; PPT, phenyl(enol)pyruvate tautomerase; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

2G. J. Poelarends and C. P. Whitman, unpublished results.

3Without this reduction step, the bound species is lost under the conditions of the MALDI-TOF MS experiments. Reduction with NaBH$_4$ results in the complete loss of the
absorbance at 324 nm, consistent with the rearrangement of the presumed enamine species (14, Scheme 4) into the corresponding imine, which is reduced by NaBH₄.

The pKₐ of the catalytic Pro-1 in CaaD has been determined to be 9.2 by direct titration using ¹⁵N NMR spectroscopy (10). A direct titration of Pro-1 in cis-CaaD has not been conducted, but a pH-rate profile of the cis-CaaD reaction implicated an acid catalyst with a pKₐ of ~9.3 (11). It is assumed that this pKₐ value corresponds to that of Pro-1 on the basis of the expected mechanistic parallels with CaaD.
References


Supplementary information

The experimental procedures used for the construction, expression, overproduction, and purification of His-tagged MsCCH2 wild-type and MsCCH2-P1A mutant, and the 1H NMR spectroscopic analysis of the reaction of MsCCH2 with trans-3-chloroacrylate (2) and cis-3-chloroacrylate (3) are provided below. In addition, the experimental procedures used for and results of the 1H NMR spectroscopic analysis of the reaction of MsCCH2 with (p-hydroxyphenyl)enolpyruvate (10) and the mass spectral analysis of the products of the MsCCH2-catalyzed reaction with 2-oxo-3-pentynoate (12) are provided below.

Supplementary methods

Construction of the MsCCH2 Expression Vector. The gene coding for MsCCH2 was amplified using two synthetic primers, genomic DNA of M. smegmatis strain MC2 155, and the PCR reagents supplied in the Phusion DNA Polymerase system. The genomic DNA of strain MC2 155 (GenBank sequence: CP000480) was kindly provided by Professor William R. Jacobs, Jr. (Howard Hughes Medical Institute, Chevy Chase, MD). The forward primer 5'-A TAG CAG GTA CAT ATG CCC GTC TAC ACC GTC AC-3' contains an NdeI restriction site (in bold) followed by 17 bases corresponding to the coding sequence of the gene. To clone the gene for MsCCH2 in frame with the sequence that codes for the polyhistidine region of the expression vector pET-20b(+), the reverse primer 5'-G TGA TGT TAT AAG CTT CCG AGA ACC TGA GCT CT-3' (primer R), which contains a HindIII restriction site (in bold), was used. The amplification mixture was made up in Phusion High-Fidelity buffer and contained 250 μM of each dNTP, 150 ng of each primer, 250 ng of total genomic DNA, and 1 U of Phusion DNA polymerase. The cycling parameters were 95 °C for 5 min followed by 35 cycles of 98 °C for 30 s, 45 °C for 30 s and 72 °C for 20 s, with a final elongation step of 72 °C for 10 min. The resulting PCR product and the pET-20b(+) vector (Novagen) were digested with NdeI and HindIII restriction enzymes, after which the vector was dephosphorylated with alkaline phosphatase. Following purification, the PCR product and vector were ligated using T4 DNA ligase. An aliquot of the ligation mixture was transformed into competent Escherichia coli BL21 (DE3) cells. Transformants were selected at 37 °C on LB/ampicillin plates. Plasmid DNA was isolated from several colonies and analyzed by restriction analysis for the presence of the insert. The cloned gene was sequenced to verify that no mutations had been introduced during the amplification of the gene. The newly constructed expression vector for MsCCH2 was named pMsCCH2.

Construction of the MsCCH2-P1A Mutant. The MsCCH2-P1A mutant was generated by PCR using plasmid pMsCCH2 as the template and primers F (5'-A TAG CAG GTC CAT ATG GCC GTC TAC ACC GTC AC-3') and R. Primer F contains an NdeI restriction site (in bold) and the mutated codon (underlined), followed by 14 bases corresponding to the coding sequence of the gene. The PCR product was generated and cloned into the pET-20b(+) vector as described above for the wild-type MsCCH2 encoding gene. The cloned mutant gene was completely sequenced to verify that only the intended mutation had been
introduced. The expression vector for the MsCCH2-P1A mutant was named pMsCCH2-P1A.

**Expression and Purification of His-Tagged MsCCH2 Proteins.** The MsCCH2 proteins, either wild-type or P1A mutant, were produced constitutively in *E. coli* BL21(DE3) using the T7 expression system. Fresh BL21(DE3) transformants containing the appropriate expression plasmid, either pMsCCH2 or pMsCCH2-P1A, were collected from a LB/ampicillin plate and used to inoculate LB/ampicillin medium (5 mL). After growth for 5 h at 37 °C, this culture was used to inoculate fresh LB/ampicillin medium (1 L) to a starting OD$_{600}$ of about 0.01. After overnight growth at 30 °C with vigorous shaking, cells were harvested by centrifugation (10 min at 2300g) and resuspended in 10 mM NaH$_2$PO$_4$ buffer, pH 7.3 (buffer A), to a total volume of about 10 mL. Protease inhibitors (Complete Mini, Roche, Mannheim, Germany) were added and cells were disrupted by sonication for 5 × 1 min (with 5 min rest in between each cycle) at 50% duty cycle/50% output in a Branson Sonifier 450 (Branson Ultrasonics Corporation, Danbury, CT). Unbroken cells and debris were removed by centrifugation (45 min at 18000g). The supernatant was filtered through a 0.2 μm pore diameter filter, made 10 mM in imidazole (from a 1 M stock solution in water, pH adjusted to 7.3), and applied to a gravity flow column containing 0.5 mL Ni-Sepharose 6 Fast Flow resin. After overnight incubation at 4 °C, while mixing at 60 rpm on a rotor, the nonbound proteins were removed from the column by gravity flow. The column was washed with 40 mM imidazole in buffer A, after which retained proteins were eluted with 1 mL of 250 mM imidazole in buffer A. Subsequently, the protein sample was applied to a PD-10 gel filtration column, which was previously equilibrated in buffer A, and proteins were eluted with buffer A. Fractions (1 mL) were analyzed by SDS-PAGE, and those containing highly purified MsCCH2 were combined and concentrated to a protein concentration of 5-20 mg/mL using a Vivaspin centrifugal concentrator equipped with a 5000 Da molecular weight cut-off filtration membrane (Sartorius Stedim Biotech S.A., France).

**$^1$H NMR Spectroscopic Analysis of the Reaction of MsCCH2 with p-Hydroxyphenylenolpyruvate (10).** A series of $^1$H NMR spectra monitoring the MsCCH2-catalyzed conversion of 10 to p-hydroxyphenylpyruvate (11) was recorded using a procedure described elsewhere (1) with the following modifications. An aliquot (550 μL) of a solution of 10 (50 μmol in 2 mL of 50 mM acetate buffer, pH 6.2) was added to an NMR tube containing D$_2$O (50 μL) and MsCCH2 (50 μL from a 5 mg/mL stock solution in 10 mM Na$_2$HPO$_4$ buffer, pH 7.3). The control sample for analyzing the nonenzymatic rate was prepared in the same way, but now 50 μL of 10 mM Na$_2$HPO$_4$ buffer (pH 7.3) was added to the NMR tube instead of 50 μL of MsCCH2. $^1$H NMR spectra were recorded directly after mixing and then after 5, 10 and 15 min. The $^1$H NMR signals for 10 and 11 are reported elsewhere (1).

**$^1$H NMR Spectroscopic Analysis of the Reaction of MsCCH2 with 2 and 3.** A series of $^1$H NMR spectra monitoring the MsCCH2-catalyzed dehalogenation of trans-3-chloroacrylate (2) or cis-3-chloroacrylate (3) was recorded according to protocols reported elsewhere (2,3) with the following modifications. An aliquot of MsCCH2 (50 μL from a 20 mg/mL stock solution in 10 mM Na$_2$HPO$_4$ buffer, pH 7.3) was added to an
NMR tube containing D$_2$O (50 µL) and 500 µL of a 60 mM solution of 2 or 3 in 100 mM Na$_2$HPO$_4$ buffer (pH adjusted to either 6.5 or 9.2). The first $^1$H NMR spectrum was recorded directly after mixing, and then after 1 and 14 days. After 14 days of incubation at either pH 6.5 or pH 9.2, the spectrum for the mixture containing MsCCH2 and 2 exhibited signals corresponding to 2, acetaldehyde, and the hydrate of acetaldehyde. The spectrum for the mixture containing MsCCH2 and 3 exhibited signals corresponding to 3, acetaldehyde, and the hydrate of acetaldehyde (both at pH 6.5 and 9.2). The $^1$H NMR signals for 2, 3, acetaldehyde, and the hydrate of acetaldehyde are reported elsewhere (2,3,4).

**Mass Spectral Analysis of the Product of the MsCCH2-Catalyzed Hydration of 2-Oxo-3-Pentynoate (12).** To identify acetopyruvate (13) as the product of the MsCCH2-catalyzed hydration of 12 the following reactions were performed. An aliquot of MsCCH2 (from a 5 mg/mL stock solution in 5 mM NH$_4$HCO$_3$ buffer, pH 7.3) was diluted into 1 mL of 5 mM NH$_4$HCO$_3$ buffer (pH 6.8), yielding a final enzyme concentration of 62 µM. The reaction was initiated by the addition of 10 µL of 12 (from a 40 mM stock solution in 10 mM Na$_2$HPO$_4$ buffer, pH adjusted to 7.3), yielding a final substrate concentration of 400 µM. In a separate control reaction, CaaD (27 µM) was incubated with 12 (400 µM) under identical conditions to produce 13, which was used as reference compound. The two reactions were monitored spectrophotometrically at 294 nm ($\lambda_{\text{max}}$ of 13) until completion, after which the enzyme was removed from each reaction mixture by centrifugation on vivaSpin centrifugal concentrator equipped with a 2000 Da molecular weight cut-off filtration membrane. The flow-through of each reaction mixture was placed on ice and directly analyzed by ESI-MS and MS/MS.

**Supplementary results**

**MsCCH2 Catalyzes the Conversion of 10 to 11.** The MsCCH2-catalyzed tautomerization of 10 was also monitored by $^1$H NMR spectroscopy to verify that the reaction results in the formation of 11. The spectrum recorded immediately after dissolving $p$-hydroxyphenylpyruvate in acetate buffer revealed two doublets in the aromatic region (7.53 and 6.78 ppm) and a singlet representing an olefinic proton at 6.20 ppm. These signals are consistent with the enol form of the substrate. The MsCCH2-catalyzed conversion of the enol form yields the keto form, as indicated by changes in the chemical shifts of the aromatic protons (to 7.00 and 6.75 ppm) and the appearance of a singlet at 3.87 ppm, which corresponds to the methylene protons. After a 10 min-incubation period, the ratio of keto:enol is ~5 (Figure S1A). For comparison, a 10 min-incubation period in the absence of enzyme yields a keto:enol ratio of ~0.12 (Figure S1B). These observations confirm that 11 is the product of the MsCCH2-catalyzed conversion of 10.

**MsCCH2 Catalyzes the Conversion of 12 to 13 and 14.** While 13 has a characteristic $\lambda_{\text{max}}$ of 294 nm (2), its formation in the reaction of 12 with MsCCH2 was confirmed by ESI-MS and MS/MS analyses. Accordingly, MsCCH2 was incubated with 12 in 5 mM NH$_4$HCO$_3$ buffer (pH 6.8). In a separate control reaction, CaaD was incubated with 12 under identical conditions to produce 13, which was used as reference compound. The two
reactions were monitored spectrophotometrically at 294 nm until completion, after which the enzymes were removed from the reaction mixtures by centrifugation on vivaspin concentrators. The flow-through of each reaction mixture was analyzed by ESI-MS. The spectrum revealed a major product for the MsCCH2-catalyzed reaction with a mass (129.2 Da) nearly identical to that found for the product of the CaaD-catalyzed reaction (129.1 Da). MS/MS analysis of the precursor ion of the 129.2 Da product of the MsCCH2-catalyzed reaction revealed fragment ions with masses (44.9, 57.0, and 84.9 Da) identical to those found with the 129.1 Da product of the CaaD-catalyzed reaction. Based on these results, we conclude that 13 is indeed the product of the MsCCH2-catalyzed hydration of 12.

We next investigated whether the absorbance peak at 324 nm corresponds to the covalent modification of MsCCH2 by 12, presumably yielding enamine species 14. Accordingly, MsCCH2 was treated with 12 at pH 7.3, after which the enzyme was separated from the free substrate (12) and product (13) by centrifugation on a vivaspin column. Analysis of the purified enzyme by UV spectroscopy demonstrated that the absorbance at 324 nm is indeed associated with a covalent adduct on the enzyme rather than an unbound reaction product.

MsCCH2 was also treated with 12 at pH 7.3 and directly analyzed by ESI-MS. The reconstruct of the ESI mass spectrum revealed two major peaks, one corresponding to the mass of unmodified MsCCH2 (16092 Da) and the other to the mass expected for modified MsCCH2 (16204 Da). Hence, the difference between the masses is 112 Da, consistent with covalent modification of the enzyme by a single molecule of 12 (5). A comparison of the signal intensities in the ESI mass spectrum suggests that about 30% of the enzyme is covalently modified by 12.
Supplementary figures

Figure S1A

Figure S1B

Figure S1. $^1$H NMR spectra monitoring the conversion of $p$-hydroxyphenylpyruvate (10) to $p$-hydroxyphenylpyruvate (11). (A) Spectrum of the MsCCH2-catalyzed conversion of 10 to 11. (B) Spectrum of the non-enzymatic conversion of 10 to 11. Both spectra were recorded 10 min after the addition of 10 to the reaction mixture. The $^1$H NMR signals for compounds 10 and 11 are reported elsewhere (7).
Supplementary references


