# Catalytic promiscuity of 4-oxalocrotonate tautomerase

Zandvoort, Ellen

**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:*
2012

*Citation for published version (APA):*
Zandvoort, E. (2012). *Catalytic promiscuity of 4-oxalocrotonate tautomerase: discovery and characterization of C-C bond-forming activities*. s.n.

# Catalytic Promiscuity of 4-Oxalocrotonate Tautomerase

Discovery and Characterization of C-C Bond-Forming Activities

**Ellen Zandvoort**

**RIJKSUNIVERSITEIT GRONINGEN**


# Catalytic Promiscuity of 4-Oxalocrotonate Tautomerase

Discovery and Characterization of C-C Bond-Forming Activities


**Proefschrift**



ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
vrijdag 7 december 2012
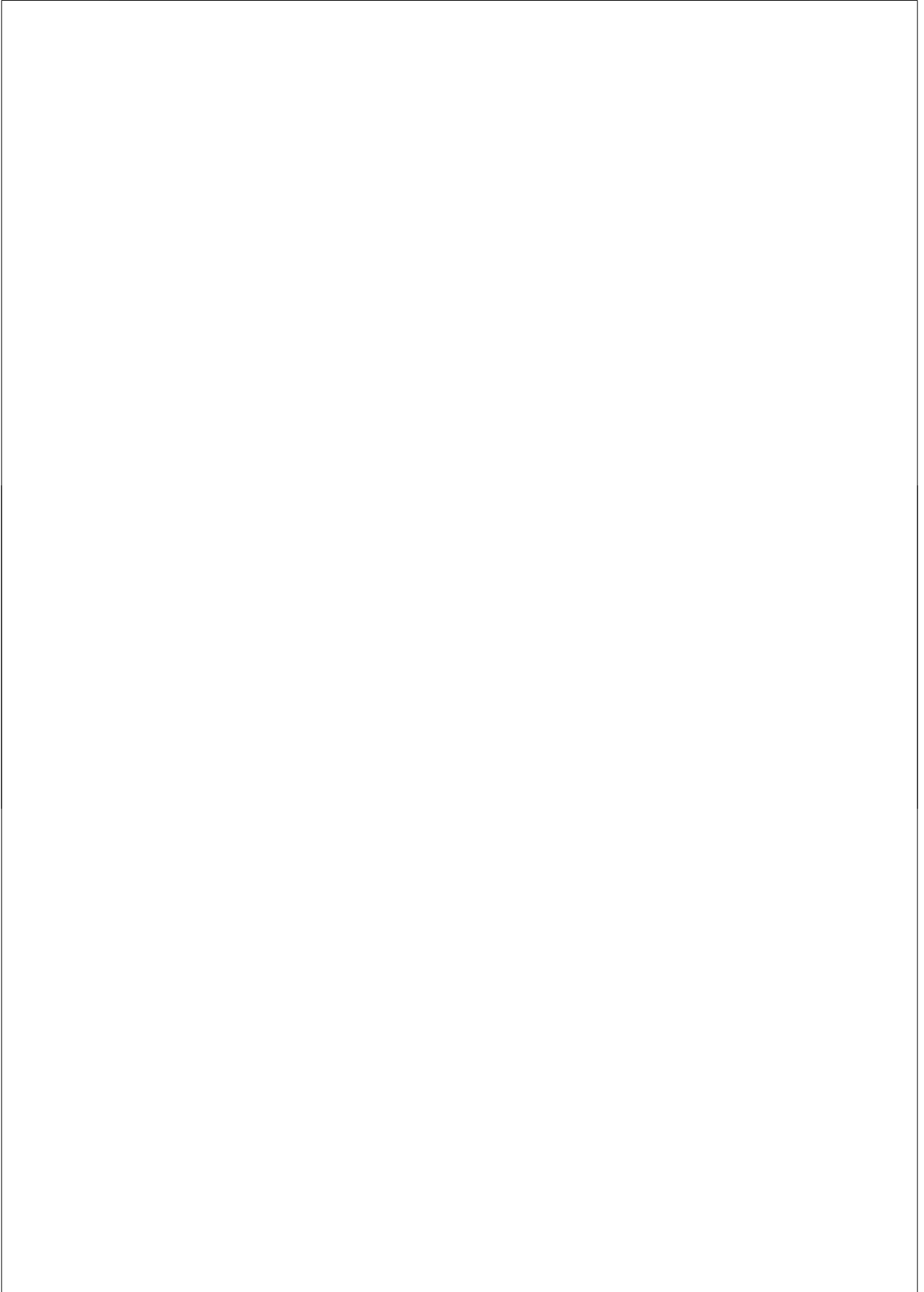om 14.30 uur



door



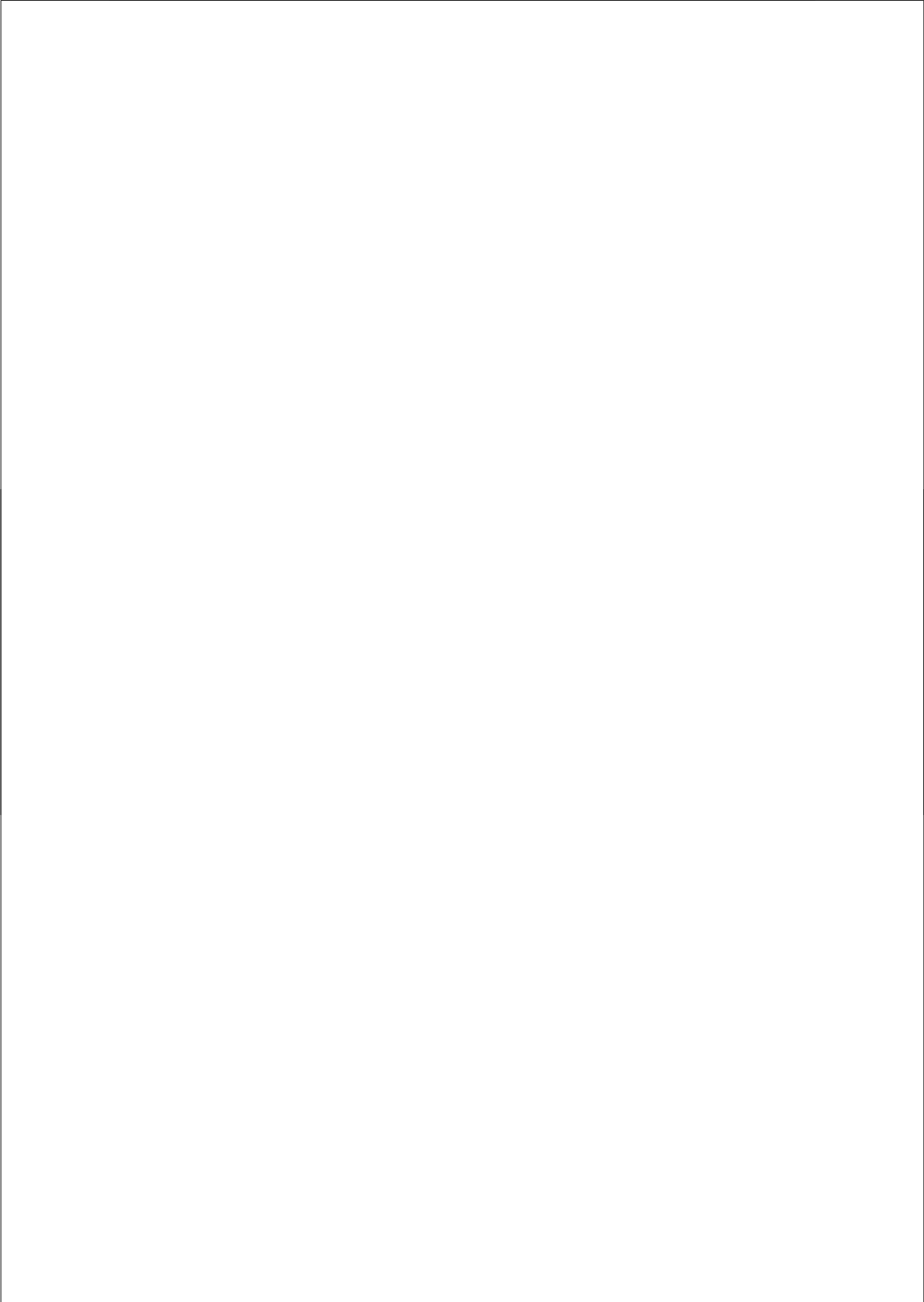**Ellen Zandvoort**


geboren op 25 oktober 1983
te Drachten

# Contents

# Aim and Scope of this Thesis

The formation of a carbon-carbon bond is one of the most important reactions for the synthesis of new compounds. During the last two decades, L-proline and related cyclic secondary amines were found to be useful organocatalysts for various C-C bond-forming reactions, including aldol condensations, Michael additions and alkylations, through enamine catalysis. However, the major drawback of most of these organocatalytic synthesis reactions is the need for large amounts of catalyst and organic solvents.

4-oxalocrotonate tautomerase (4-OT) is a member of the tautomerase superfamily, a group of enzymes that has the unique feature of using an amino-terminal proline in catalysis. 4-OT is part of a catabolic pathway for aromatic hydrocarbons in *Pseudomonas putida* mt-2, where it catalyzes the conversion of 2-hydroxy-2,4-hexadienedioate to 2-oxo-3-hexenedioate. In this natural activity of 4-OT, Pro-1 functions as a general base. Pro-1 can act as a general base because the prolyl nitrogen has a p$K_a$ of ~6.4 so that it exists largely as the uncharged species at neutral pH.

**Inspired by the success of proline in organocatalysis, the work described in this thesis aimed to explore the proline-based tautomerase 4-OT for catalyzing unnatural carbon-carbon bond-forming reactions. The long term goal of this research line is to develop a new class of biocatalysts, based on the 4-OT scaffold, that use the powerful proline-based enamine mechanism of organocatalysts, but take advantage of the water solubility and stereoselectivity of enzymes.**

Many enzymes are catalytically promiscuous and catalyze one or more alternative reactions in addition to the one they evolved for. The catalytic promiscuity of enzymes has become a source of inspiration for the design of new biocatalysts. Within the tautomerase superfamily, many examples of catalytic promiscuity exist, as is set out in **Chapter 1**.

So far, most promiscuous activities have been discovered either by chance or by looking for a specific reaction based on an enzyme's close relatives. The challenge, however, is to use mechanistic reasoning to discover new promiscuous reactions that may serve as a springboard to develop new biocatalysts. **Chapter 2** describes a systematic screening approach which is based on the presence of a nucleophilic proline residue within the active site of 4-OT, which provides a potential catalytic framework that can give rise to promiscuous carbonyl transformation activities. The use of this approach led to the discovery of a promiscuous aldolase activity in 4-OT, in which acetaldehyde is coupled to benzaldehyde to yield cinnamaldehyde. In addition, a 4-OT variant with a single amino acid

substitution (L8R) was constructed, which has improved aldolase activity (16-fold in terms of $k_{cat}/K_m$).

Although an improved variant (L8R) has been found, its aldolase activity still remains a low-level activity. By protein engineering, a different 4-OT mutant (F50A) was obtained with 600-fold higher activity (in terms of $k_{cat}/K_m$) compared to 4-OT WT (**Chapter 3**). This mutant enzyme was used to study the mechanism of the aldolase activity in more detail.

In **Chapter 4** a different promiscuous carbon-carbon bond-forming activity of 4-OT is discussed: the Michael-type addition of acetaldehyde to β-nitrostyrenes. 4-OT catalyzes these reactions with high stereoselectivity, yielding the corresponding γ-nitroaldehydes, which are important precursors for various pharmaceuticals.

In the course of screening for promiscuous Michael-type addition activity in 4-OT, it has been found that 4-OT catalyzes the isomerization of *cis*-nitrostyrene to *trans*-nitrostyrene. Although this is not a carbon-carbon bond forming reaction, it is a promiscuous activity of 4-OT that is of high mechanistic interest. Therefore, the discovery and characterization of this activity is described in **Chapter 5**.

Finally, in **Chapter 6** the work described in this thesis is summarized, and suggestions for future research are presented.

# 1

# Introduction to Enzyme Promiscuity in the Tautomerase Superfamily

*B. J. Baas, E. Zandvoort, and G. J. Poelarends*

## Introduction

The concept of enzymes being highly specific catalysts towards a single chemical transformation, for which nature has invented them, is nowadays seen as a somewhat restricted view.[1] During the past few decades enzymes and the families and superfamilies they are part of have been intensively studied and characterized.[2] The results of these studies show that enzymes often do not accept just a single substrate, but instead accept a range of substrates. In addition, in a growing number of enzymes a feature termed catalytic promiscuity has been identified, which can be defined as the capability of a single active site to catalyze chemically, and often mechanistically, distinct transformations.[3] The study of catalytic promiscuity of enzymes provides various insights, which can be divided into three topics.

First, the investigation of catalytic promiscuity of enzymes extends our comprehension of the evolution of enzymes in nature.[4] Enzymes that cluster together in families and superfamilies tend to share activities, the native activity of one member being a catalytically promiscuous one for the other.[5] In such a case, catalytic promiscuity could be viewed as a vestigial feature of that enzyme's evolutionary progenitor (or ancestor). Therefore, catalytic promiscuity is thought to play a role in the divergent evolution of enzyme function. Already in the 1970s, it was postulated that if a catalytically promiscuous activity could provide a selective advantage for the host organism, such a promiscuous activity could be the starting point for the evolution of a new enzyme.[6] A low-level promiscuous activity can be amplified by accumulation of mutations, preceded or followed by gene duplication. The promiscuous activity of the parent enzyme then becomes the primary activity of the newly evolved enzyme. This concept is nowadays widely accepted and worked out in more detail by a number of key studies on enzyme promiscuity.[1a, 4, 7]

Second, the characterization of catalytically promiscuous activities of enzymes can aid in the identification of active site residues as catalytically important.[8] Moreover, insight can be gained into the mechanistic roles of active site residues. Often it is found that the same set of catalytic residues is involved in both the enzyme's native and promiscuous activity, but that these residues fulfill a different mechanistic role.[9]

Third, the presence of catalytic promiscuity shows the chemical versatility of an enzyme's active site.[1a] In fact, catalytic promiscuity might even be predicted based on knowledge of the chemical and mechanistic properties of active site residues.[3,10] Here lies a formidable challenge: the use of mechanistic reasoning to discover new promiscuous activities in existing enzymes, which could be exploited to generate novel biocatalysts.

These three topics will be addressed in this chapter based on recent results obtained from studies on the enzymes of the tautomerase superfamily. The enzymes belonging to this superfamily share two defining features: all members exhibit a β-α-β-fold as the basic

structural motif, and possess an amino-terminal proline as a key catalytic residue (Figure 1.1).[5b,11] The known members of the tautomerase superfamily can be divided into five families, the first studied member being the title enzyme of each family: the 4-oxalocrotonate tautomerase (4-OT), *cis*-3-chloroacrylic acid dehalogenase (*cis*-CaaD), malonate semialdehyde decarboxylase (MSAD), 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI), and the macrophage migration inhibitory factor (MIF) family.[12] Catalytic promiscuity is a feature found in all five families. In addition, the chemical nature of these reactions is diverse, involving the formation or breakage of carbon-hydrogen, carbon-carbon, carbon-oxygen and carbon-halogen bonds.[5b] The tautomerase superfamily is therefore an excellent group of enzymes to study the different facettes of catalytic promiscuity. In the remaining part of this chapter, recent advances in the study of catalytic promiscuity in the tautomerase superfamily and the implications derived thereof are reviewed.
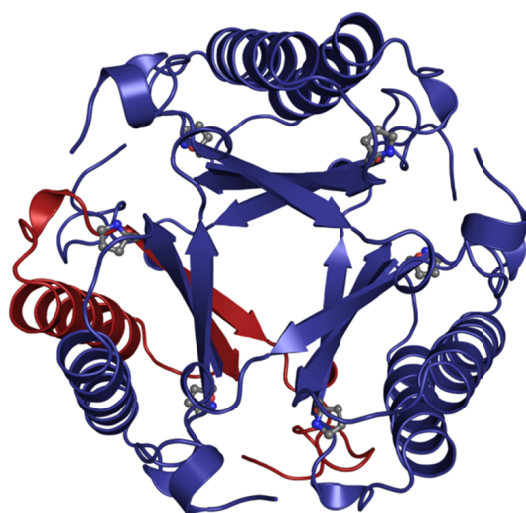


**Figure 1.1.** A ribbon diagram showing the β-α-β building block (highlighted in red) and the amino-terminal proline residues (in ball/stick representation), that are characteristic features of members of the tautomerase superfamily, here represented by 4-oxalocrotonate tautomerase (4-OT).

**Catalytic promiscuity and divergent evolution in the tautomerase superfamily.**
A fascinating example of catalytic promiscuity is the hydrolytic dehalogenation of 3*E*-haloacrylates by various members of the tautomerase superfamily. Two enzymes from this superfamily, *trans*-3-chloroacrylic acid dehalogenase (CaaD) and *cis*-CaaD, utilize the *trans*- and *cis*-isomer of 3-chloroacrylate (*trans*-3-CAA and *cis*-3-CAA), respectively, as their primary substrate (Table 1.1).[13,14] Both CaaD and *cis*-CaaD are enzymes found in soil-

dwelling bacteria, where they are part of a catabolic pathway evolved for the degradation of 1,3-dichloropropene (Scheme 1.1), a xenobiotic nematocide introduced into the environment in the 1950s.[15] The enzymes catalyze the hydrolytic dehalogenation of *trans*-3-CAA, or *cis*-3-CAA to yield malonate semialdehyde, one of the intermediary steps in the degradation of this nematocide. The end product of the pathway is acetaldehyde, which is utilized by the host organism as a source of carbon and energy.[15b,16] The crystal structures of both enzymes show a β-α-β-fold, and mutagenesis studies have implicated Pro-1 as essential for catalytic activity.[17,18] CaaD and *cis*-CaaD are highly proficient enzymes, which both afford approximately a $2 \times 10^{12}$-fold rate enhancement when compared to the rate of spontaneous uncatalyzed hydrolysis of *trans*-CAA.[19] It is quite surprising that two enzymes with high catalytic proficiency towards a man-made compound seemingly have evolved within just a few decades. Hence, CaaD and *cis*-CaaD appear to be the product of rapid divergent evolution and catalytic promiscuity might have played an important role in this process. Intriguingly, both CaaD and *cis*-CaaD promiscuously catalyze the enol-keto tautomerization of phenylenolpyruvate to phenylpyruvate (PPT activity) (Table 1.1, Scheme 1.2).[13] The presence of this tautomerase activity establishes a functional link with the known tautomerases in the superfamily. This promiscuous PPT activity could be a vestige of an evolutionary tautomerase ancestor of these recently evolved dehalogenases. Thus, the characterization of related tautomerase family members of CaaD and *cis*-CaaD might shed light on the evolutionary origin of these dehalogenating enzymes and on the evolutionary route by which they have emerged.
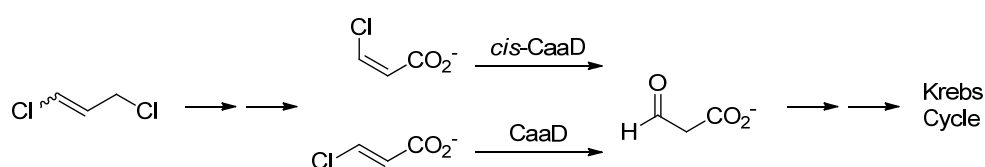


**Scheme 1.1.** The dehalogenation of *trans*- and *cis*-3-chloroacrylate (*trans*- and *cis*-3-CAA) by CaaD and *cis*-CaaD, respectively, as part of a degradation pathway for the nematocide 1,3-dichloropropene.
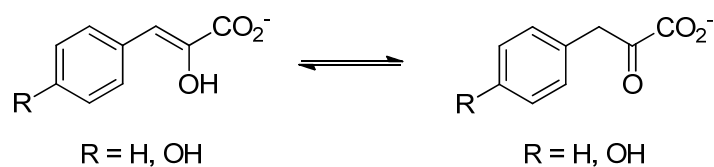


**Scheme 1.2.** The tautomerization of (4-hydroxy)phenylenolpyruvate to (4-hydroxy)phenylpyruvate.

**Table 1.1.** Kinetic parameters of the activities of tautomerase superfamily members.

| Enzyme | Substrate | pH | $k_{cat}$ (s$^{-1}$) | $K_m$ (μM) | $k_{cat}/K_m$ (M$^{-1}$ s$^{-1}$) |
|---|---|---|---|---|---|
| CaaD[13] | *trans*-3-CAA | 9.0 | 2.4±0.1 | 34±2 | $7.1 \times 10^4$ |
| CaaD[13] | phenylpyruvate | 9.0 | 1.4±0.3 | 61±23 | $2.3 \times 10^4$ |
| CaaD[33] | 2-OP | 9.0 | 0.7±0.02 | 110±4 | $6.4 \times 10^3$ |
| *cis*-CaaD[14] | *cis*-3-CAA | 8.0 | 4.7±0.2 | 210±10 | $(2.2±0.1) \times 10^4$ |
| *cis*-CaaD[13] | phenylpyruvate | 9.0 | 0.20±0.03 | 110±30 | $1.8 \times 10^3$ |
| *cis*-CaaD[14] | 2,3-butadien-oate | 8.0 | 6.0±0.1 | 690±30 | $(8.7±0.4) \times 10^3$ |
| *cis*-CaaD[12b] | 2-OP | 9.0 | $(7±0.5) \times 10^{-3}$ | 620±60 | $0.011 \times 10^3$ |
| 4-OT[25] | 2-hydroxy-muconate | 7.3 | $(3.5±0.5) \times 10^3$ | 180±30 | $1.9 \times 10^7$ |
| 4-OT[25] | *trans*-3-CAA | 8.2 | $(1.0±0.2) \times 10^{-3}$ | $(91±34) \times 10^3$ | $1.1 \times 10^{-2}$ |
| 4-OT L8R[25] | *trans*-3-CAA | 8.2 | $(8.8±0.3) \times 10^{-3}$ | $(16±1) \times 10^3$ | $5.5 \times 10^{-1}$ |
| YwhB[21] | *trans*-3-CAA | 7.8 | - | - | $(4.4±1.0) \times 10^{-2}$ |
| Cg10062[27] | 2-OP | 9.0 | 0.33±0.03 | $(6.20±0.75) \times 10^3$ | 50 |
| Cg10062[27] | *trans*-3-CAA | 9.0 | 0.06±0.01 | $(78±36) \times 10^3$ | $8.0 \times 10^{-1}$ |
| Cg10062[27] | *cis*-3-CAA | 9.0 | 3.5±1.1 | $(19±1) \times 10^3$ | 184 |
| MsCCH2[28] | phenylpyruvate | 7.3 | 34±4 | $(3.5±0.3) \times 10^3$ | $9.8 \times 10^3$ |
| MsCCH2[28] | 4-hydroxy-phenylpyruvate | 7.3 | 3.2±0.3 | 890±90 | $3.5 \times 10^3$ |
| MsCCH2[28] | *trans*-3-CAA | 8.0 | $(4.0±0.1) \times 10^{-4}$ | $(96±6) \times 10^3$ | $4.0 \times 10^{-3}$ |
| MsCCH2[28] | *cis*-3-CAA | 8.0 | - | - | $1.6 \times 10^{-3}$ |
| MIF[30] | phenylpyruvate | 7.5 | 114±13 | 500±120 | $2.3 \times 10^5$ |
| MIF[30] | *trans*-3-CAA | 6.5 | $3.0 \times 10^{-5}$ | $(5.8±1.2) \times 10^3$ | $5.0 \times 10^{-3}$ |
| MIF I64V V106L[30] | *trans*-3-CAA | 6.5 | $6.0 \times 10^{-3}$ | $(18.2±4) \times 10^3$ | $3.3 \times 10^{-1}$ |

CaaD is a member of the 4-OT family of enzymes, and both CaaD and 4-OT are well-studied members of the tautomerase superfamily.[20] The conversion of *trans*-3-CAA as a case of catalytic promiscuity was first identified in 4-OT.[21] 4-OT, from *Pseudomonas putida* mt-2, is part of a catabolic pathway for the degradation of aromatic hydrocarbons like toluene and benzene.[22] The enzyme catalyzes the 1,5-enol-keto tautomerization of 2-hydroxymuconate to 2-oxohex-3-enedioate (Table 1.1, Scheme 1.3) where Pro-1 acts as a general base facilitating the transfer of the hydroxyl proton to the C-5 position. In addition to its primary function as a tautomerase, 4-OT exhibits low-level CaaD activity (Table 1.1). Albeit seven orders of magnitude lower than the reported catalytic efficiency of CaaD, this example of catalytic promiscuity of 4-OT is still remarkable, since the hydrolytic dehalogenation of *trans*-3-CAA is accelerated $10^9$-fold when compared to the rate of spontaneous hydrolysis of this substrate.[19]
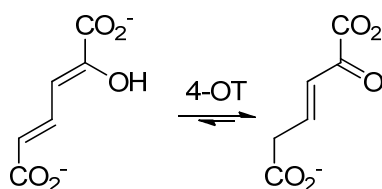


**Scheme 1.3.** The 4-OT-catalyzed tautomerization of 2-hydroxymuconate (i.e., 2-hydroxyhexa-2,4-dienedioate) to 2-oxohex-3-enedioate.

In-depth characterization of CaaD and 4-OT has resulted in detailed insight into their structure and catalytic mechanism. In addition to similarities such as a β-α-β structural fold and a catalytic Pro-1 residue, which is present in an active site pocket, clear differences were observed in terms of their overall catalytic machinery and the role of Pro-1. The $pK_a$ value of Pro-1 was determined for both enzymes and it reflects the different role this residue plays in the catalytic mechanism. For Pro-1 in CaaD, a $pK_a$ value of 9.3 was found, which enables Pro-1 to function as a general acid at neutral pH.[23] In contrast, Pro-1 of 4-OT exhibits a $pK_a$ of ~6.4, allowing it to function as a general base at cellular pH.[24] With regard to the other catalytic residues, two major differences were observed. In CaaD, αGlu-52 activates a water molecule for the addition to C-3 of *trans*-3-CAA. The resulting enediolate intermediate is stabilized by a pair of arginine residues, αArg-8 and αArg-11.[13] Both αGlu-52 and αArg-8 in CaaD are lacking at the corresponding positions in 4-OT. Introduction of these residues in 4-OT yielded interesting results.[25] Two mutants, 4-OT L8R and 4-OT L8R I52E, exhibit an increase in catalytic efficiency mainly due to a $k_{cat}$ or a $K_m$ effect, respectively. For the best mutant, 4-OT L8R, a 50-fold increase in $k_{cat}/K_m$ was found (Table 1.1). These results indicate that the CaaD activity of 4-OT, which belongs to the

same family as CaaD, can be increased relatively easily, requiring only a single amino acid substitution.

The ability to catalyze the hydrolytic dehalogenation of *trans*-3-CAA in the 4-OT family is not limited to 4-OT and CaaD. YwhB, a 4-OT homologue from *Bacillus subtilis*, was found to be a promiscuous dehalogenase as well.[21,26] The enzyme shares 36% pairwise sequence identity with 4-OT. Both Pro-1 and Arg-11 are conserved and are key catalytic residues for the observed activity, as was shown by mutagenesis studies. YwhB is an example of the catalytic ability of 4-OT-like enzymes to function as a promiscuous dehalogenase towards *trans*-3-CAA (Table 1.1). Although the direct progenitor of CaaD is unknown, it is thus likely that CaaD has evolved by divergent evolution from a tautomerase-like ancestor which displayed promiscuous dehalogenase activity.

Until recently, the evolutionary origins of *cis*-CaaD were little understood. *cis*-CaaD belongs to a different family of the tautomerase superfamily as CaaD, and it is therefore unlikely that both enzymes share the same ancestral lineage. CaaD and *cis*-CaaD are very alike in terms of their overall structure and catalytic machinery, however, there are a number of major differences.[17,18] The most obvious difference is the distinct oligomeric structure of the two enzymes. CaaD is a heterohexamer, which can be viewed as a trimer of dimers, where each dimer is built from an α- and a β-subunit. Each subunit is characterized by a β-α-β structural motif and an amino-terminal proline, with βPro-1 being the catalytically important Pro-1. *cis*-CaaD, however, is a homotrimer in which each monomer has a catalytically important Pro-1 residue; each trimer is made up of two consecutive β-α-β-fold motifs that apparently became fused. Whether this fusion event has taken place during the evolution of *cis*-CaaD, or of its progenitor, remains open to debate. Another difference was observed from the crystal structures of both enzymes, which revealed a different active site geometry. The active site of CaaD has an elongated shape, which can readily accommodate *trans*-3-CAA. In *cis*-CaaD however, a u-shaped active site pocket was found, analogous to the bent shape of the *cis*-isomer that it converts. This difference in active site geometry likely is the origin of the isomer specificity of the two enzymes.[5b]

Some insight into the evolutionary process that gave rise to *cis*-CaaD and into its progenitor enzyme was gained from the study of two members from the *cis*-CaaD family, Cg10062 and MsCCH2.[27,28] The gene encoding Cg10062 was identified in the genome of *Corynebacterium glutamicum* based on sequence homology with the gene encoding *cis*-CaaD. Despite only 34% sequence identity, six key catalytic residues of *cis*-CaaD are conserved in Cg10062. Similar to *cis*-CaaD, Cg10062 was found to be a hydratase, converting 2-oxo-3-pentynoate (2-OP) to acetopyruvate (Table 1.1, Scheme 1.4). In addition, Cg10062 was found to catalyze the hydrolytic dehalogenation of *cis*-3-CAA at a low level, the $k_{cat}/K_m$ being ~160-fold lower compared to *cis*-CaaD (Table 1.1). Of the six

conserved residues of the catalytic machinery of *cis*-CaaD, four (Pro-1, Arg-70, Arg-73 and Glu-114) were found to be crucial for the dehalogenase activity of Cg10062, since mutagenesis of these residues results in complete loss of activity. Surprisingly, the enzyme was found to accept *trans*-3-CAA as a substrate as well. Although the catalytic efficiency towards the *trans*-isomer is over 200-fold lower compared to the value determined for the *cis*-isomer, it shows that Cg10062 is not isomer-specific. The low-level *cis*-CaaD activity combined with a lack of isomer specificity shows that Cg10062 is not a fully functional dehalogenase. This finding shows that the presence of the core set of catalytic residues of *cis*-CaaD alone is not the sole prerequisite for an enzyme being an efficient and specific dehalogenase. Discriminating between isomers and achieving a high catalytic efficiency likely originates from optimized positioning of all residues that line the active site pocket. The absence of a clear genomic context of the *cg10062* gene prevented the researchers of identifying the native activity of Cg10062 in its host organism. Although the strain of *C. glutamicum* was isolated from a soil sample after the introduction of 1,3-dichloropropene into the environment, the absence of genes associated with a catabolic degradation pathway for this compound in the organism's genome indicated that Cg10062 is unlikely to function as a *cis*-3-CAA dehalogenase in this strain.[29]
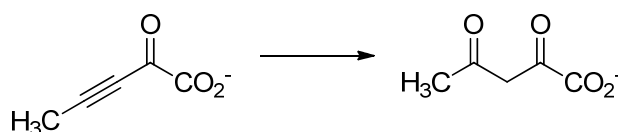


**Scheme 1.4.** The hydration of 2-oxo-3-pentynoate (2-OP) to acetopyruvate.

The study on Cg10062 shows that within the proteome of the prokaryotic kingdom, *cis*-CaaD-like enzymes do exist, which may represent an intermediary stage in the evolution towards a fully functional dehalogenase. However, without an activity that can be firmly linked to other members of the tautomerase superfamily, the promiscuous hydratase and dehalogenase activities of Cg10062 alone offer few clues about the ancestral enzyme from which *cis*-CaaD may have evolved. To gain further insight into this, another member of the *cis*-CaaD family, termed MsCCH2, was cloned and characterized.[28] The gene encoding MsCCH2 was identified already some years ago in the genome of *Mycobacterium smegmatis* MC2 155.[12b] A pairwise sequence alignment showed 28% identity and 39% similarity to *cis*-CaaD and revealed that four key catalytic residues of *cis*-CaaD (Pro-1, His-28, Arg-68 and Glu-112) are conserved in MsCCH2. The relatively low sequence identity to *cis*-CaaD and the incomplete catalytic machinery indicate that MsCCH2 is a more distant family member of *cis*-CaaD as Cg10062. The genomic context of the *mscch2* gene does not reveal any clear information about the native activity of MsCCH2 in its host organism. Characterization of

MsCCH2 using likely substrates for a tautomerase superfamily member, revealed that MsCCH2 is a robust tautomerase, catalyzing the conversion of phenylenolpyruvate to phenylpyruvate (Table 1.1, Scheme 1.2 (R = H)). Additionally, the enzyme also catalyzes the tautomerization of (4-hydroxyphenyl)-enolpyruvate to (4-hydroxyphenyl)pyruvate (Table 1.1, Scheme 1.2 (R = OH)), albeit with lower efficiency.

Two promiscuous activities were identified in MsCCH2.[28] The enzyme acts as a hydratase in the conversion of 2-OP to acetopyruvate and as a dehalogenase using both *cis*- and *trans*-3-CAA (Table 1.1). In addition to sequence homology, these enzymatic activities firmly place MsCCH2 in the *cis*-CaaD family. MsCCH2 offers intriguing clues about the evolutionary origins of *cis*-CaaD. MsCCH2 is a robust tautomerase and a promiscuous dehalogenase without specificity towards one of the two isomers of 3-CAA. *cis*-CaaD, however, is a specialized dehalogenase with promiscuous PPT activity, an activity that could be a vestige from its evolutionary progenitor. The tautomerase activity of MsCCH2 combined with low-level dehalogenase activities makes MsCCH2 a suitable starting template for the evolution towards a specialized dehalogenase. These properties could resemble those of the ancestral enzyme of *cis*-CaaD. Accumulation of beneficial mutations could then increase the catalytic efficiency, and yield an enzyme that is somewhere in between a specialist tautomerase and a specialist dehalogenase. An enzyme like Cg10062 might resemble such a generalist intermediate on the evolutionary route toward a fully functional and isomer-specific dehalogenase.

The recurring ability to catalyze a phenylenolpyruvate tautomerization is an interesting observation in the study of the role of catalytic promiscuity in the evolution of CaaD and *cis*-CaaD (Table 1.1). These observations prompted Wasiel *et al*. to investigate whether the mammalian cytokine MIF acts as a dehalogenase towards *cis*- or *trans*-3-CAA.[30] MIF is a member of the tautomerase superfamily and in addition to its cytokine function, it is also known as the enzyme phenylpyruvate tautomerase.[31] MIF is the most efficient phenylpyruvate tautomerase characterized to this date (Table 1.1).[30] This activity clearly establishes a functional link between mammalian MIF and the bacterial dehalogenases CaaD and *cis*-CaaD. Intriguingly, MIF was found to act as a promiscuous dehalogenase towards *trans*-3-CAA (Table 1.1). To probe the evolvability of MIF towards this activity, a number of mutant libraries were screened, targeting residues in the active site pocket of MIF. This strategy yielded a double mutant of MIF (I64V V106L), which exhibits a 200-fold improvement in $k_{cat}$ (Table 1.1). Although the $k_{cat}$ ( $6.0 \times 10^{-3}$ s$^{-1}$) is rather low, the enzyme affords a $10^{9}$-fold rate enhancement when compared to the rate of spontaneous hydrolysis of the substrate.[19] Such a rate enhancement lies in the order of magnitude usually observed for more convential enzymes acting on their natural substrates. These findings

clearly demonstrate that MIF is a promiscuous dehalogenase towards *trans*-3-CAA, and that this activity can be significantly enhanced by only a few mutations.

The results reviewed here strengthen the notion that there is an evolutionary and functional link between the tautomerases and the dehalogenases in the tautomerase superfamily. The presence of low-level promiscuous tautomerase activity in the dehalogenases, and, vice versa, a low-level promiscuous dehalogenase activity in the tautomerases within the superfamily supports the hypothesis that a tautomerase likely has been the progenitor of the enzymes CaaD and *cis*-CaaD. The recurrence, specifically, of the PPT activity may indicate that these ancestral enzymes also had this activity and that a tautomerase active site capable of catalyzing a enol-keto tautomerization of phenylenolpyruvate, is also equipped to catalyze the hydrolytic dehalogenation of 3-CAA. The mechanistic relationship between these two activities may lie in a shared reaction intermediate, which is formed in both reactions. As was already postulated in the study of the PPT activity of CaaD, *trans*-3-CAA and phenylpyruvate (i.e. 2-hydroxy-3-phenylacrylate) both possess an acrylate functionality.[13] The ability to bind and polarize this functional group might make an enzyme well-suited to function as a tautomerase and act as a template for evolution towards a dehalogenase that acts on 3-CAA. The results from in-depth studies on CaaD and *cis*-CaaD, combined with the recent insights gained from the study of closely related enzymes (Cg10062 and MsCCH2), and a more distantly related enzyme (MIF), highlight the possible role of catalytic promiscuity in the divergent evolution of enzymatic activities.

**Mechanistic insight from the analysis of promiscuous enzymatic activities.**

The study of catalytic promiscuity often leads to a more refined insight in the catalytic mechanism of the reactions catalyzed by enzymes.[11] For example, the occurrence of a specific intermediate in the mechanism of a promiscuous activity may suggest that a similar intermediate is involved in the mechanism of the native reaction catalyzed by that enzyme.[14] Alternatively, the mechanism of promiscuous activities might be totally different from that of the native activity and in such a case promiscuity demonstrates the versatile chemical nature of an enzyme's active site.[25] In many cases, the same active site residues are implicated in both the promiscuous and native reactions, but they play very different mechanistic roles.

One example of a shared intermediate in the native and promiscuous reaction is highlighted by the analysis of the promiscuous PPT activity of CaaD.[13] The precise mechanistic route by which CaaD catalyzes the dehalogenation of *trans*-3-CAA has not been clearly established, however, two likely mechanisms have been postulated (Scheme 1.5).[13,23] Both mechanisms involve the same initial step, which is the Michael-type addition

of water to the double bond of *trans*-3-CAA, yielding a presumed enediolate intermediate stabilized by two active site arginines. This proposed intermediate can then undergo two possible fates. The enediolate intermediate can undergo ketonization accompanied by protonation of the C-2 carbon atom by Pro-1, yielding an unstable halohydrin species. Expulsion of the chlorine atom, either chemically or enzymatically, yields the final product malonate semialdehyde (Scheme 1.5, A). In an alternative mechanism, direct elimination of chlorine from the enediolate yields the enol-tautomer of the final product, which undergoes ketonization to yield malonate semialdehyde (Scheme 1.5, B). Distinguishing between these two different fates of the enediolate intermediate is an experimental challenge; however, a clue about which of the two routes is more likely is offered by the promiscuous PPT activity of CaaD. In this reaction, the enol form of phenylpyruvate is accepted as a substrate for enol-keto tautomerization. These findings have led the authors to conclude that the active site of CaaD is equipped to ketonize enol compounds which makes the occurance of enol-intermediates in the dehalogenase mechanism more likely. Considering both mechanistic routes for the dehalogenation of *trans*-3-CAA (Scheme 1.5), route B thus appears to be more likely in view of the enol-intermediate that is generated by α,β-elimination of HCl from the initial enediolate. This enediolate in itself can also be viewed as an enol-type functionality, making it a plausible initial intermediate in view of the mechanistic similarities between the native dehalogenase and the promiscuous PPT activity of CaaD.
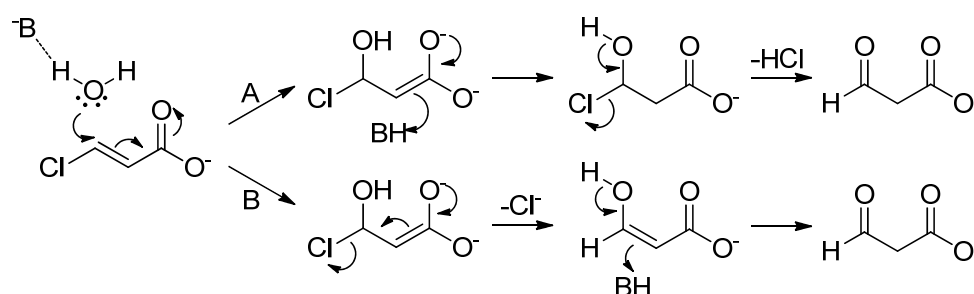


**Scheme 1.5.** Two proposed mechanistic routes for the CaaD-catalyzed dehalogenation of *trans*-3-CAA to malonate semialdehyde.

The origin of catalytic promiscuity can often be found in the diverse chemical nature of the residues that make up the active site. In the tautomerase superfamily, Pro-1 is implicated as a key catalytic residue for all superfamily members characterized to this date.[5b] Indeed, Pro-1 is implicated in both the native and promiscuous activities found among superfamily members and in many cases, Pro-1 plays a different mechanistic role. This residue may act as a general acid, a general base, or as a nucleophile. The diverse chemical nature of the reactions catalyzed by the tautomerase superfamily stems, in part,

from these three distinct chemical roles that the secondary amine functionality of the amino-terminal proline has to offer. A few examples will be reviewed here.
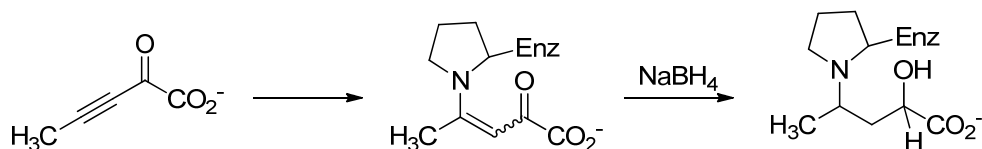


**Scheme 1.6.** Modification of Pro-1 by 2-oxo-3-pentynoate (2-OP) yielding a covalent enamine-adduct. Subsequent reduction by $NaBH_4$ irreversibly traps the adduct.

The enzyme MsCCH2 was discussed in the previous section due to its relevance in unraveling the evolutionary origins of *cis*-CaaD. However, during the characterization of MsCCH2 an interesting phenomenon was observed in the reactivity of the enzyme towards the acetylene compound 2-OP.[28] Previous studies on 4-OT and CaaD (and *cis*-CaaD) have shown that 2-OP can be used as a probe to determine the protonation state of Pro-1. In the case of 4-OT, 2-OP forms a covalent adduct with Pro-1 at neutral pH. The low p$K_a$ of Pro-1 (~6.4) renders it mainly deprotonated at neutral pH, whereby Pro-1 can act as a nucleophile towards 2-OP leading to covalent modification. Subsequent reduction by $NaBH_4$ irreversibly traps the covalent adduct (Scheme 1.6).[32] The p$K_a$ of the amino-terminal Pro-1 residue of CaaD was determined to be ~9.3 and as a result the enzyme is not covalently modified but instead catalyzes a hydration reaction converting 2-OP to acetopyruvate (Scheme 1.4).[33] At neutral pH, Pro-1 of CaaD is largely cationic and is likely involved in the hydration mechanism as a general acid. For MsCCH2, the p$K_a$ of the Pro-1 residue has not been determined, and, hence, the exact mechanistic role of Pro-1 at neutral pH is unknown. However, interesting results were obtained when incubating the enzyme with 2-OP at different pH values.[28] At pH 8.5, 2-OP inactivates MsCCH2 via 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 (similar to 4-OT). At pH 6.5, however, MsCCH2 exhibits hydratase activity, analogous to CaaD, and converts 2-OP to acetopyruvate, which implies that Pro-1 is cationic at pH 6.5 (similar to CaaD). At the intermediate pH of 7.5, the hydratase and inactivation reactions occur simultaneously. The reactivity of MsCCH2 towards 2-OP therefore reflects the protonation state of Pro-1 in this pH range. Hence, the p$K_a$ of the amino-terminal proline of the enzyme can be estimated to be about 7.5 as reflected by the dual fate of 2-OP at this pH. These results illustrate that the mechanistic role of Pro-1 is dependent on its p$K_a$ and that by altering the pH of the experimental conditions, one mechanism may be favored over the other.
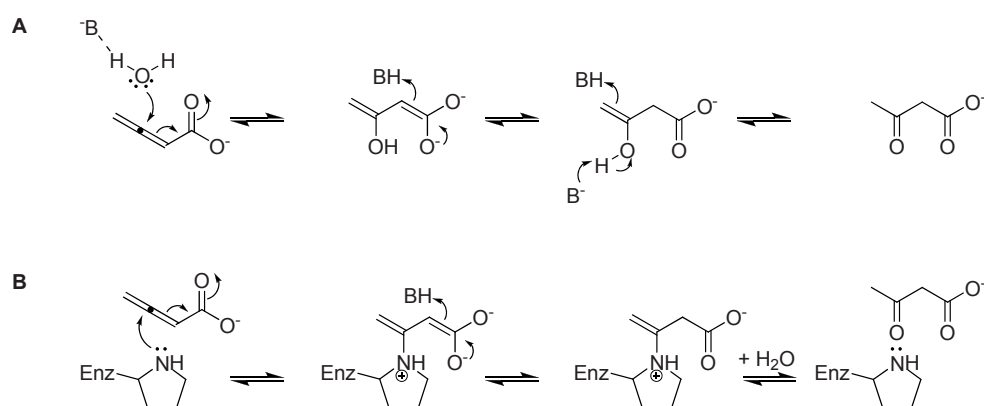
**Scheme 1.7.** The proposed mechanism of the *cis*-CaaD-catalyzed conversion of 2,3-butadienoate to acetoacetate via water addition (A) or via a covalent intermediate (B).

Recently Schroeder *et al.* identified a previously unknown promiscuous activity of *cis*-CaaD.[14] In an attempt to gain insight into the mechanistic details of the native dehalogenase activity of *cis*-CaaD through the study of substrate analogues, it was found that *cis*-CaaD catalyzes the conversion of 2,3-butadienoate to acetoacetate (Table 1.1, Scheme 1.7). It was anticipated that the enzyme would add water to the electrophilic C-3 carbon atom of the substrate (Scheme 1.7A), analogous to the initial step in the hydration of *cis*-3-CAA.[18] In an attempt to study the stereochemistry of this presumed reaction, it was carried out in $D_2O$ followed by reduction using $NaBH_4$. Surprisingly, this reduction step inactivated the enzyme. ESI-MS and MALDI-MS analysis of the inactivated enzyme revealed that the amino-terminal Pro-1 residue was covalently modified by a species with a mass consistent with that of the decarboxylated imine of 2,3-butadienoate. This finding, combined with the analysis of pre-steady state kinetics, is consistent with a mechanism involving covalent catalysis (Scheme 1.7B) that is favored over a mechanism involving the conjugate addition of water (Scheme 1.7A). The question was raised whether the native dehalogenation of *cis*-3-CAA catalyzed by *cis*-CaaD could also proceed through a covalent intermediate in an addition-elimination type mechanism (Scheme 1.8). Preliminary results from the analysis of pre-steady state kinetics revealed a branched pathway for the *cis*-CaaD-catalyzed dehalogenation of *cis*-3-CAA.[14] This leads to the possibility that covalent catalysis and direct hydration of the substrate might take place simultaneously. However, a covalent intermediate has not yet been observed in the *cis*-CaaD-catalyzed dehalogenation of *cis*-3-CAA. These findings show how the study of a promiscuous activity may yield new insights into the mechanism of the primary activity of an enzyme. Additionally, it emphasizes the need of studying the mechanism of the native reaction more closely, which may ultimately yield a more refined image of the actual mechanistic capabilities of *cis*-CaaD.
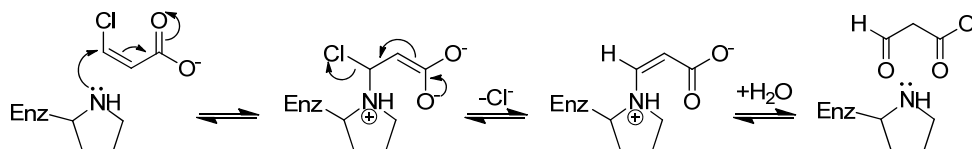
**Scheme 1.8.** A proposed alternative mechanism of the *cis*-CaaD-catalyzed dehalogenation of *cis*-3-CAA involving a covalent intermediate.

### 'Predicting' catalytic promiscuity

By studying catalytic promiscuity within enzyme (super)families, in this specific case within the tautomerase superfamily, many insights are gained not only on enzyme evolution and enzyme mechanisms, but also on the chemical versatility of an enzyme's active site.[10] In turn, this knowledge can be used to predict new unnatural activities for the same enzyme scaffold, which can be exploited as a springboard for the generation of new biocatalysts. More specifically, the characteristic amino-terminal proline is a key catalytic residue in all tautomerase superfamily members, and provides a unique chemical functionality in the tautomerase active site. The secondary amine of Pro-1 can act as a general acid, a general base, or as a nucleophile; which is illustrated by the diverse roles of Pro-1 in the various activities described in the previous sections.

In the natural tautomerase activity of 4-OT, Pro-1 functions as a general base.[12] Intriguingly, since the p$K_a$ of Pro-1 is relatively low (~6.4), it might also be able to act as a nucleophile at neutral pH. Nucleophilic secondary amines, like Pro-1, can react with carbonyl compounds to give enamine intermediates, which can undergo a wide range of further reactions with appropriate electrophiles. The remaining chapters of this thesis describe the exploration of 4-OT for its ability to catalyze carbon-carbon bond-formation via enamine formation between carbonyl compounds and Pro-1.

# References

1. a) A. Babtie, N. Tokuriki, F. Hollfelder, *Curr Opin Chem Biol.* **2010**, *14*, 200-207; b) O. Khersonsky, S. Malitsky, I. Rogachev, D. S. Tawfik, *Biochemistry* **2011**, *50*, 2683-2690.

2. A. Sakai, A. A. Fedorov, E. V. Fedorov, A. M. Schnoes, M. E. Glasner, S. Brown, M. E. Rutter, K. Bain, S. Chang, T. Gheyi, J. M. Sauder, S. K. Burley, P. C. Babbitt, S. C. Almo, J. A. Gerlt, J.A., *Biochemistry* **2009**, *48*, 1445-1453.

3. R. J. Kazlauskas, *Curr. Opin. Chem. Biol.* **2005**, *9*, 195-201.

4. a) O. Khersonsky, D. S. Tawfik, *Annu Rev Biochem.* **2010**, *79*, 471-505; b) P. C. Babbitt, M. S. Hasson, J. E. Wedekind, D. R. J. Palmer, W. C. Barrett, G. H. Reed, I. Rayment, D. Ringe, G. L. Kenyon, J. A. Gerlt, *Biochemistry.* **1996**, *35*, 16489-16501.

5. a) C. Roodveldt, D. S. Tawfik, *Biochemistry* **2005**, *44*, 12728-12736; b) G. J. Poelarends, V. Puthan Veetil, C. P. Whitman, *Cell. Mol. Life Sci.* **2008**, *65*, 3606-3618.

6. a) S. Ohno, *Evolution by gene duplication*, **1970**, Allen & Unwin/Springer-Verlag, London-NY; b) R. A. Jensen, *Ann. Rev. Microbiol.* **1976**, *30*, 409-425; c) F. Jacob, *Science* **1977**, *196*, 1161-1166.

7. a) P. J. O'Brien, D. Herschlag, *Chem Biol.* **1999**, *6*, R91-R105; b) L. C. James, D. S. Tawfik, *Protein Sci.* **2001**, *10*, 2600-2607; c) M. E. Glasner, J. A. Gerlt, P. C. Babbitt, *Curr Opin Chem Biol.* **2006**, *10*, 492-497; d) A. Aharoni, L. Gaidukov, O. Khersonsky, S. McQ Gould, C. Roodveldt, D. S. Tawfik, *Nat Genet.* **2005**, *37*, 73-76.

8. a) G. J. Poelarends, H. Serrano, W. H. Johnson, Jr., D. W. Hoffman, C. P. Whitman, *J. Am. Chem. Soc.* **2004**, *126*, 15658-15659; b) D. E. Künzler, S. Sasso, M. Gamper, D. Hilvert, P. Kast, *J. Biol. Chem.* **2005**, *280*, 32827-32834; c) K. Hult, P. Berglund, *Trends Biotechnol.* **2007**, *25*, 231-238.

9. S. D. Copley, *Curr Opin Chem Biol.* **2003**, *7*, 265-272.

10. F. P. Seebeck, D. Hilvert, *J. Am. Chem. Soc.* **2003**, *125*, 10158-10159.

11. C. P. Whitman, *Arch Biochem Biophys.* **2002**, *402*, 1-13.

12. a) A. G. Murzin, *Curr. Opin. Struct. Biol.* **1996**, *6*, 386–394; b) G. J. Poelarends, H. Serrano, M. D. Person, W. H. Johnson, Jr., A. G. Murzin, C. P. Whitman, *Biochemistry,* **2004**, *43*, 759–772; c) G. J. Poelarends, W. H. Johnson, Jr., A. G. Murzin, C. P. Whitman, *J. Biol. Chem.* **2003**, *278*, 48674–48683; d) C. P. Whitman, *Comprehensive Natural Products Chemistry*, **1999**, Elsevier, Oxford, Vol. 5, 31–50; e) J. B. Lubetsky, M. Swope, C. Dealwis, P. Blake, E. Lolis, *Biochemistry* **1999**, *38*, 7346–7354.

13. G. J. Poelarends, W. H. Johnson, Jr., H. Serrano, C. P. Whitman, *Biochemistry* **2007**, *46*, 9596-9604.

14. G. K. Schroeder, W. H. Johnson, Jr., J. P. Huddleston, H. Serrano, K. A. Johnson, C. P. Whitman, *J. Am. Chem. Soc.* **2012**, *134*, 293-304.

15. a) G. J. Poelarends, C. P. Whitman, *Bioorg. Chem.* **2004**, *32*, 376–392; b) G. J. Poelarends, M. Wilkens, M. J. Larkin, J. D. Van Elsas, D. B. Janssen, *Appl. Environ. Microbiol.* **1998**, *64*, 2931-2936.

16. S. Hartmans, M. W. Jansen, M. J. van der Werf, J. A. de Bont, *J. Gen. Microbiol.* **1991**, *137*, 2025-2032.

17. a) R. M. de Jong, W. Brugman, G. J. Poelarends, C. P. Whitman, B. W. Dijkstra, *J. Biol. Chem.* **2004**, *279*, 11546-11552; b) S. D. Pegan, H. Serrano, C. P. Whitman, A. D. Mesecar, *Acta. Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64*, 1277-1282.

18. R. M. de Jong, P. Bazzacco, G. J. Poelarends, W. H. Johnson, Jr., Y. J. Kim, E. A. Burks, H. Serrano, A. M. W. H. Thunnissen, C. P. Whitman, B. W. Dijkstra, *J. Biol. Chem.* **2007**, *282*, 2440-2449.

19. C. M. Horvat, R. V. Wolfenden, *Proc. Natl. Acad. Sci.* **2005**, *102*, 16199-16202.

20. G. J. Poelarends, R. Saunier, D. B. Janssen, *J. Bacteriol.* **2001**, *183*, 4269-4277.

21. S. C. Wang, W. H. Johnson, Jr., C. P. Whitman, *J. Am. Chem. Soc.* **2003**, *125*, 14282-14283.

22. L. H. Chen, G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour, C. P. Whitman, *J. Biol. Chem.* **1992**, *267*, 17716-17721.

23. H. F. Azurmendi, S. C. Wang, M. A. Massiah, G. J. Poelarends, C. P. Whitman, A. S. Mildvan, *Biochemistry* **2004**, *43*, 4082-4091.

24. J. T. Stivers, C. Abeygunawardana, A. S. Mildvan, G. Hajipour, C. P. Whitman, *Biochemistry* **1996**, *35*, 814-823.

25. G. J. Poelarends, J. J. Almrud, H. Serrano, J. E. Darty, W. H. Johnson, Jr., M. L. Hackert, C. P. Whitman, *Biochemistry* **2006**, *45*, 7700-7708.

26. S. C. Wang, W. H. Johnson, Jr., R. M. Czerwinski, S. L. Stamps, C. P. Whitman, *Biochemistry* **2007**, *46*, 11919-11929.

27. G. J. Poelarends, H. Serrano, M. D. Person, W. H. Johnson, Jr., C. P. Whitman, *Biochemistry* **2008**, *47*, 8139-8147.

28. B. J. Baas, E. Zandvoort, A. A. Wasiel, W. J. Quax, G. J. Poelarends, *Biochemistry* **2011**, *50*, 2889-2899.

29. M. Silberbach, M. Schäfer, A. T. Hüser, J. Kalinowski, A. Pühler, R. Krämer, A. Burkovski, *Appl. Environ. Microbiol.* **2005**, *71*, 2391-2402.

30. A. A. Wasiel, B. J. Baas, E. Zandvoort, W. J. Quax, G. J. Poelarends, *Chembiochem* **2012**, *13*, 1270-1273.

31. E. Rosengren, P. Åman, S. Thelin, C. Hansson, S. Ahlfors, P. Björk, L. Jacobsson, H. Rorsman, *FEBS Lett.* **1997**, *417*, 85-88.

32. W. H. Johnson, Jr., R. M. Czerwinski, M. C. Fitzgerald, C. P. Whitman, *Biochemistry* **1997**, *36*, 15724–15732.

33. S. C. Wang, M. D. Person, W. H. Johnson, Jr., C. P. Whitman, *Biochemistry* **2003**, *42*, 8762-8773.

# 2

## Systematic Screening for Catalytic Promiscuity in 4-Oxalocrotonate Tautomerase: Enamine Formation and Aldolase Activity

*E. Zandvoort, B. J. Baas, W. J. Quax, and G. J. Poelarends*

# Abstract

The enzyme 4-oxalocrotonate tautomerase (4-OT) is part of a catabolic pathway for aromatic hydrocarbons in *Pseudomonas putida* mt-2, where it catalyzes the conversion of 2-hydroxyhexa-2,4-dienedioate (**1**) to 2-oxohex-3-enedioate (**2**). In the mechanism of 4-OT, Pro-1 is a general base that abstracts the 2-hydroxyl proton of **1** for delivery to the C-5 position to yield **2**. In the present study, 4-OT was systematically explored for nucleophilic catalysis based on the mechanistic reasoning that its Pro-1 residue has the correct protonation state ($pK_a$ ~6.4) to be able to act as a nucleophile at pH 7.3. This led to the discovery of a new promiscuous activity in wild-type 4-OT: the enzyme catalyzes the aldol condensation of acetaldehyde with benzaldehyde to form cinnamaldehyde. This low-level aldolase activity can be improved 16-fold with a single point mutation (L8R) in 4-OT's active site. Further systematic screening of 4-OT and related proline-based biocatalysts may prove to be a useful approach to discover new promiscuous carbonyl transformation activities that could be exploited to develop new biocatalysts for carbon-carbon bond formation.
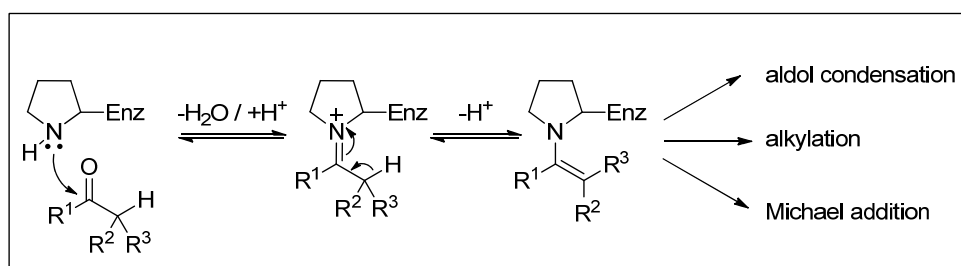
# Introduction

The notion that many enzymes are catalytically promiscuous and can catalyze one or more alternative reactions in addition to the one they evolved for has become a source of inspiration for the design of useful biocatalysts.[1] However, in order to exploit the promiscuous activities of existing enzymes to develop new biocatalysts, an important consideration is how the usually low-level promiscuous activities of enzymes might be systematically characterized. So far, most promiscuous activities have been discovered either by chance or by looking for a specific reaction based on an enzyme's close relatives.[2,3] Here, we have used a systematic screening strategy to discover new promiscuous activities in 4-oxalocrotonate tautomerase (4-OT). It is based on the mechanistic reasoning that the catalytic amino-terminal proline of this enzyme provides a unique chemical functionality in the active site that might be suitable for enamine catalysis.
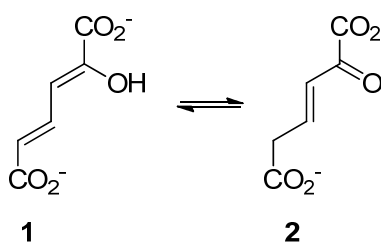


**Scheme 2.1.** The tautomerization reaction catalyzed by 4-OT.

4-OT is part of a catabolic pathway for aromatic hydrocarbons in *Pseudomonas putida* mt-2, in which it catalyzes the conversion of 2-hydroxyhexa-2,4-dienedioate (**1**) to 2-oxohex-3-enedioate (**2**) (Scheme 2.1).[4] The enzyme is a member of the tautomerase superfamily, a group of homologous proteins that are characterized by a conserved catalytic amino-terminal proline embedded within a β-α-β structural fold.[5] In the mechanism of 4-OT, Pro-1 is a general base that abstracts the 2-hydroxyl proton of **1** for delivery to the C-5 position, yielding **2**.[6] Pro-1 can function as a general base because the prolyl nitrogen has a p$K_a$ of ~6.4 and exists largely as the uncharged species at pH 7.3. Two other key catalytic residues are Arg-11 and Arg-39.[7] Arg-39 is proposed to interact with the 2-hydroxyl group of **1** and a C-1 carboxylate oxygen, whereas Arg-11 is proposed to interact with the C-6 carboxylate group in a bidentate fashion. The latter interaction might draw electron density toward C-5 for protonation by Pro-1.

Dawson and coworkers demonstrated that by mutating Pro-1 to an alanine, 4-OT can be changed from an acid/base tautomerase into an enzyme that decarboxylates oxaloacetate through a nucleophilic imine mechanism.[8] This change in catalytic activity and mechanism is consistent with the expectation that the active site primary amine of Ala-1

would be less basic and more conformationally flexible than the secondary amine of Pro-1. This enables the nucleophilic addition of Ala-1 to oxaloacetate to form an iminium ion intermediate (i.e., a protonated Schiff base), which facilitates decarboxylation. The decarboxylated product, pyruvate, is released from Ala-1 by hydrolysis.[8]

Different from this earlier work on mutant 4-OT, we set out to test whether the active site secondary amine of Pro-1 in wild-type 4-OT might be suitable for nucleophilic catalysis. We reasoned that although Pro-1 functions as a base catalyst in 4-OT's natural activity, it has the correct protonation state ($pK_a$ ~6.4) to be able to act as a nucleophile at pH 7.3.[9] Furthermore, being the only natural amino acid with a secondary amine group, Pro-1 might facilitate the reversible generation of enamines from carbonyl compounds (Scheme 2.2). Indeed, secondary amines react with carbonyl compounds to favor formation of enamine intermediates, whereas primary amines favor the imine tautomers.[10] The enzymatically generated enamine then could undergo a wide range of further reactions by using different electrophiles, which might allow for the discovery of several new activities within the active site of wild-type 4-OT. This mechanistic reasoning is strengthened by a large body of literature that highlights the success of proline and related cyclic secondary amines as organocatalysts in asymmetric enamine catalysis.[11]
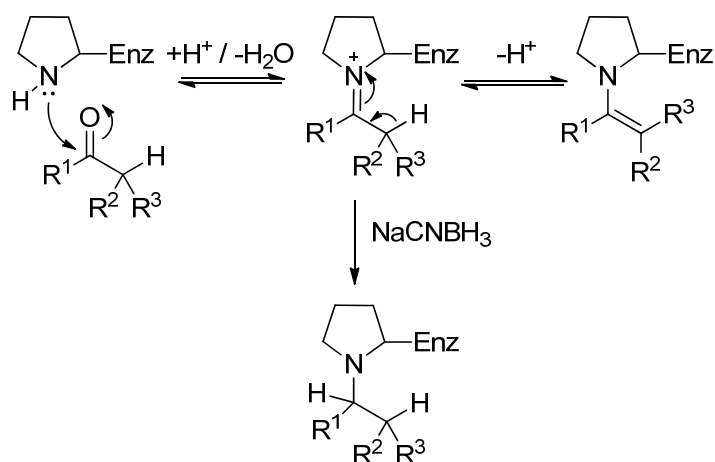


**Scheme 2.2.** Proposed mechanism of enamine formation in the active site of 4-OT. Reduction of the imine intermediate by NaCNBH$_3$ leads to irreversible covalent modification and inactivation of the enzyme.

## Results and Discussion

To identify reactive carbonyl compounds that might give rise to useful enamines, 4-OT was incubated with various aldehydes and ketones (in separate reactions) and the mixtures were treated with NaCNBH$_3$. Reduction of the imine intermediate will covalently link the carbonyl compound to the enzyme and result in its inactivation (Scheme 2.2).[8, 12] When 4-OT was treated with NaCNBH$_3$ in the presence of the selected aldehydes **3**-**7** (Scheme 2.3), enzymatic activity was almost completely lost (Table 2.1). Gel filtration chromatography did not restore activity, indicative of irreversible covalent modification. Treatment of the enzyme with the aldehyde or NaCNBH$_3$ alone did not result in the loss of enzymatic activity. These observations suggest that an imine can form between 4-OT and aldehydes **3**-**7**. Although these compounds are reactive aldehydes with the potential of forming imines or enamines with nearby amines, it is significant that inactivation and covalent modification of 4-OT also occurs with the less reactive ketones **8** and **9** (Scheme 2.3, Table 2.1). The higher reactivity of **9** likely reflects its structural resemblance to the pyruvoyl moiety of 4-OT's natural substrate **2**.

Carbonyl compounds tested as enamine donor
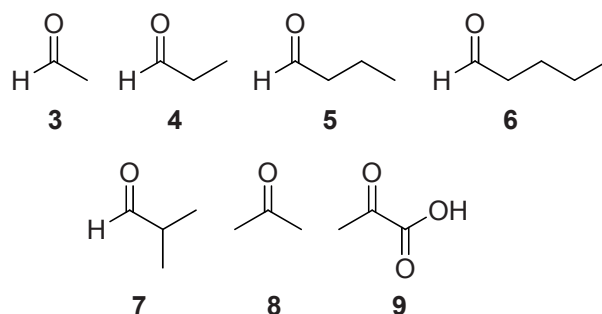


Electrophiles tested in enamine catalysis
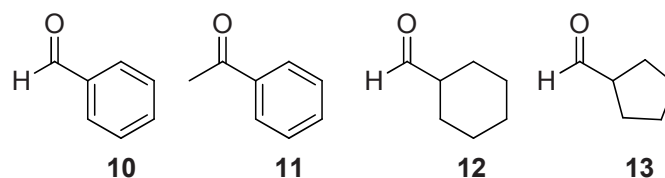


**Scheme 2.3.** Carbonyl compounds tested as enamine donor (**3**-**9**) or acceptor (**10**-**13**).

**Table 2.1.** Inactivation and covalent modification of 4-OT by selected carbonyl compounds in the presence of NaCNBH$_3$.

| Enamine donor | Inactivation (%) | Mass$_{calcd}$ (Da)[a] | Mass$_{obs}$ (Da) | Covalent labeling (%)[b] |
|:---:|:---:|:---:|:---:|:---:|
| **3** | 95 | 6839 | 6839 | >99 |
| **4** | >99 | 6853 | 6853 | >99 |
| **5** | >99 | 6867 | 6867 | >99 |
| **6** | 97 | 6881 | 6811[c],6881 | 60 |
| **7** | 79 | 6867 | 6867 | 95 |
| **8** | 23 | 6854 | 6811[c],6854 | 29 |
| **9** | >99 | 6883 | 6883 | >99 |

[a] Mass corresponding to reduced imine complex of donor substrate and 4-OT.
[b] Estimated from relative peak heights in ESI-MS spectra.
[c] Mass of unmodified 4-OT.

The inactivated protein samples were analyzed by ESI-MS to determine whether the mass is consistent with the mechanism shown in Scheme 2.2 and to ascertain whether single or multiple modifications had occurred. The deconvoluted spectrum of the 4-OT sample inactivated by either **3**-**5**, **7** or **9** displayed one major peak corresponding to the mass expected for the enzyme modified by a single molecule of the respective compound and reduced by NaCNBH$_3$ (Table 2.1). Analysis of the 4-OT sample inactivated by either **6** or **8** showed two peaks, one corresponding to the mass of unmodified 4-OT and the other to the mass expected for the enzyme modified by a single molecule of **6** or **8**, respectively, and reduced by NaCNBH$_3$ (Table 2.1). These results indicate that the reaction between 4-OT and the carbonyl compounds is specific (and not the consequence of multiple nonspecific encounters between 4-OT and carbonyl compound) and consistent with the proposed mechanism of enamine formation in the enzyme active site (Scheme 2.2).

We next investigated whether imine/enamine formation is an active site process that involves Pro-1. A 4-OT sample inactivated by **3** and a control sample (unmodified 4-OT) were digested with endoproteinase Glu-C, and the resulting peptide mixtures were analyzed by nano-LC-MS. A comparison of the peaks of the 4-OT sample inactivated by **3** in the presence of NaCNBH$_3$ to those of the unmodified 4-OT sample revealed a single modification by a species having a mass of 28 Da on the fragment Pro-1 to Glu-9 (Table **2.2**). The mass of this species corresponds to labeling by one acetaldehyde molecule. Analysis of the remaining peaks showed no modification of other fragments.

**Table 2.2.** Identification of Pro-1 as the sole site of labeling by **3** using Glu-C digestion and nano-LC-MS and MS/MS analyses.

| Peptide fragment | Mass$_{calcd}$ (Da)[a] | Mass$_{obs}$ (Da)[b] |
|---|---|---|
| PIAQIHILE | 1033.6[c] | 1032.5[c] |
| PIAQIH | 660.4 | 660.4 |
| PIAQ | 410.2 | 410.2 |
| PI | 211.1 | 211.1 |
| $C_2H_4$-PIAQIHILE | 1061.6[c] | 1060.6[c] |
| $C_2H_4$-PIAQIH | 688.4 | 688.4 |
| $C_2H_4$-PIAQ | 438.2 | 438.2 |
| $C_2H_4$-PI | 239.1 | 239.1 |

[a] These values are calculated using the average molecular mass.
[b] The reported masses correspond to the b-ions.
[c] These values correspond to the total mass of the parent ion.

Within the amino-terminal fragment Pro-1 to Glu-9, the most likely targets for alkylation are Pro-1 and His-6. To identify the alkylated amino acid residue, the unmodified and modified peptides were subjected to nano-LC-MS/MS analysis. The spectrum of the ion corresponding to the unlabeled peptide (PIAQIHILE) displayed characteristic b-ions resulting from the peptide fragments PIAQIH, PIAQ, and PI. MS/MS analysis of the ion corresponding to the modified peptide ($C_2H_4$-PIAQIHILE) revealed an increase in mass of 28 Da for these b-ions (Table **2.2**). Because one of these fragment ions is generated by the dipeptide Pro-Ile, we conclude that the active site Pro-1 residue is the sole site of modification by **3**. While the precise site of labeling of the other reactive carbonyl compounds (**4**-**9**) has not been determined by mass spectrometry, it can be inferred from these results that compounds **4**-**9** likely form imines/enamines with Pro-1.

Encouraged by these findings, the chemical potential of the smallest enamine (that generated from **3**) was explored for further reactions by using four selected electrophiles (**10**-**13**, Scheme 2.3). These electrophiles were chosen as model substrates for screening because they show structural resemblance to the phenyl portion of phenylpyruvate, a known substrate of 4-OT.[6b] Accordingly, 4-OT (0.4 mg; 90 µM) was incubated with **3** (50 mM) and each of the electrophiles (50 mM), and the four reactions were followed by $^1$H NMR spectroscopy. After incubation of **3** and **10** with 4-OT for 14 days at 22 °C (in 20 mM NaH$_2$PO$_4$ buffer, pH 7.3), the intensity of the signals corresponding to these aldehyde

substrates diminished and new signals corresponding to cinnamaldehyde (**15**, Scheme 2.4) appeared (Figure 2.1A). Integration of the signals indicates that ~10% of **3** and **10** had been converted to **15**. Whereas **15** has characteristic ^{1}H NMR signals (at 6.70 and 9.44 ppm), its identity in the incubation mixture described above was further confirmed by GC/MS analysis (Figure 2.2). In addition to the two substrate molecules, a major product was observed and identified as **15** based both on retention time comparison and the detection of fragment ions with masses identical to those found with an authentic standard.
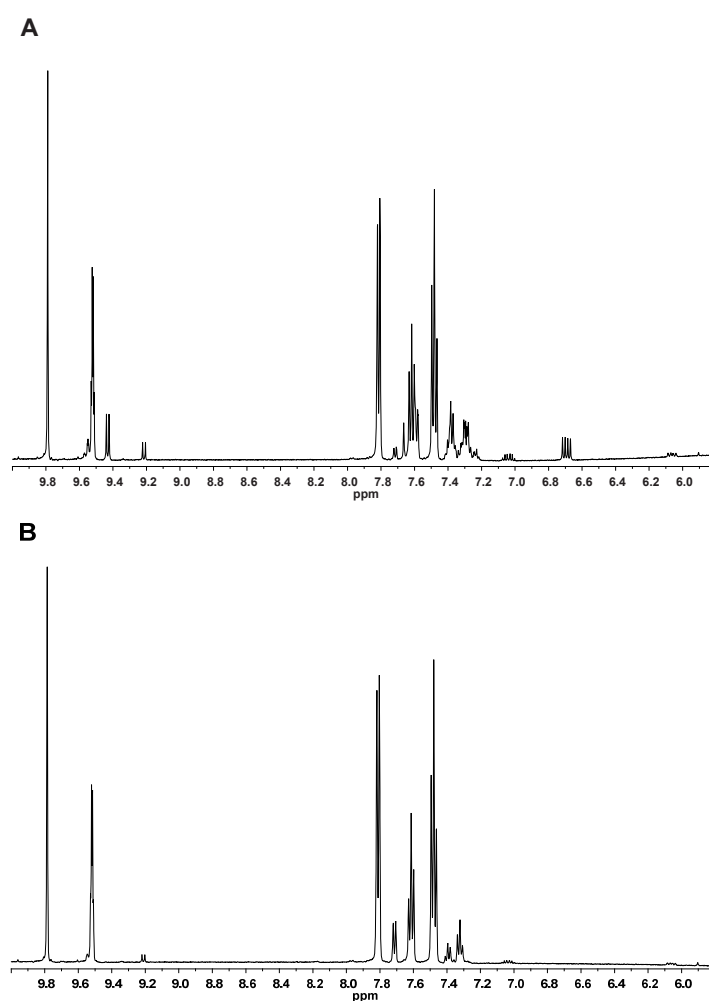


**Figure 2.1.** Partial ^{1}H NMR spectra of the 4-OT-catalyzed (A) and uncatalyzed (B) aldol condensation of **3** and **10**. Signals corresponding to **15** are only observed in spectrum A, wheras signals corresponding to **16** are detected in both spectra. The ^{1}H NMR signals for **3**, **10**, **15** and **16** are reported in the experimental procedures.
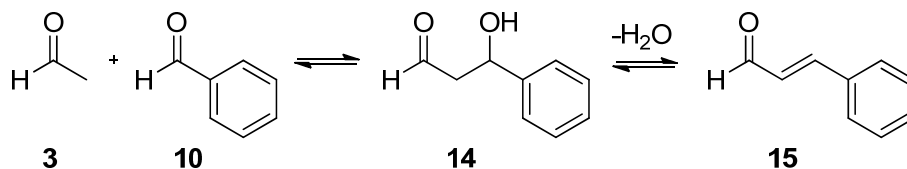
**Scheme 2.4.** The aldol condensation reaction catalyzed by 4-OT.

The effect of enzyme concentration on the amount of **15** produced was also investigated. Accordingly, **3** and **10** (each at 20 mM) were incubated with either 90 μM or 180 μM of 4-OT, and the two reactions were followed by $^{1}$H NMR spectroscopy. By using 90 μM of 4-OT, ~4% and ~10% of **3** and **10** had been converted to **15** after 7 and 14 days, respectively. By using 180 μM of 4-OT, ~9% and ~16% of **3** and **10** had been converted to **15** after 7 and 14 days, respectively. These results show that doubling the enzyme concentration doubles the amount of product formed, strongly suggesting that the reaction is enzyme-catalyzed. Notably, no significant enzyme inhibition occurs at these aldehyde concentrations.



**Figure 2.2.** GC-MS fragmentation pattern of the product formed in the incubation of **3** and **10** with 4-OT (A). Product identification was based on the detection of fragment ions with masses highly similar to those found in MS/MS spectra of **15** in the NIST107 mass spectral reference library (B).

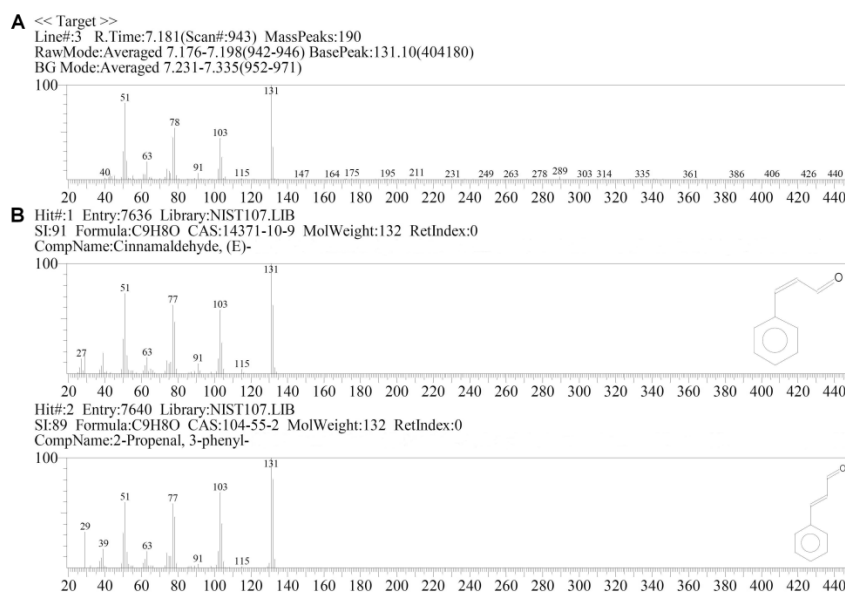Control experiments further demonstrate that the aldol condensation reaction between **3** and **10** is an enzyme-catalyzed process. GC/MS and ^{1}H NMR analyses of **3** and **10** (each at 50 mM) incubated in 20 mM $NaH_2PO_4$ buffer (pH 7.3) for 14 days at 22 °C showed no formation of **15**; this rules out a nonenzymatic aldol condensation (Figure 2.1B). Hence, the results show that the 4-OT-catalyzed aldol condensation of **3** with **10** generates **15**. A likely scenario for the formation of **15** from these compounds involves the initial enzymatic cross coupling of **3** and **10** to generate the aldol product **14** (Scheme 2.4). Subsequent chemical or enzymatic dehydration of **14** yields **15**. The incubations of **3** and each of the electrophiles **11**-**13** with 4-OT showed no detectable conversion, demonstrating the high selectivity of 4-OT for substrate **10**. As expected, incubation of only **10** (without **3**) with 4-OT in 20 mM $NaH_2PO_4$ buffer (pH 7.3) for 14 days at 22 °C also showed no conversion, which demonstrates that both substrates (**3** and **10**) are required for product formation.

The preparation of 4-OT used in these experiments was highly purified, but it remained possible that a contaminating enzyme from the *E. coli* BL21(DE3) expression host could be responsible for the observed aldolase activity. To eliminate this concern, three control experiments were performed. First, incubation of 4-OT with 3-bromopyruvate, an active-site directed irreversible inhibitor that alkylates Pro-1,[13] led to single-site modification of 4-OT (as shown by ESI-MS) and the concomitant loss of the aldolase activity (Figure 2.3C). Second, a mock purification was performed from BL21(DE3) cells harboring an empty pET20b(+) vector. Incubation of an aliquot of this purification with **3** and **10** for 14 days at 22 °C did not result in the formation of **15**. Third, a 4-OT sample free of contaminating cellular enzymes was prepared by total chemical synthesis.[14] Incubation of **3** and **10** with synthetically prepared 4-OT for 14 days at 22 °C yielded product **15** in an amount comparable to that formed in the reaction catalyzed by purified recombinant 4-OT (Figure 2.3D).

We next investigated the importance of Pro-1, Arg-11 and Arg-39 to 4-OT's aldolase activity by analyzing the kinetic properties of the corresponding alanine mutants. ^{1}H NMR spectroscopic analysis revealed that after a 14-day incubation period at 22 °C, the reaction catalyzed by 4-OT R39A results in ~8% of **15**, whereas the reaction catalyzed by 4-OT P1A and 4-OT R11A gave only a trace amount of **15** (<1%) (Figure 2.3). These results provide further evidence indicating that 4-OT is responsible for the observed aldolase activity and that Pro-1 and Arg-11 are critical for the activity, but that Arg-39 is not essential for the activity.

It has previously been reported that 4-OT exhibits a promiscuous dehalogenase activity that can be significantly increased (50-fold in terms of $k_{cat}/K_m$) by the replacement of the
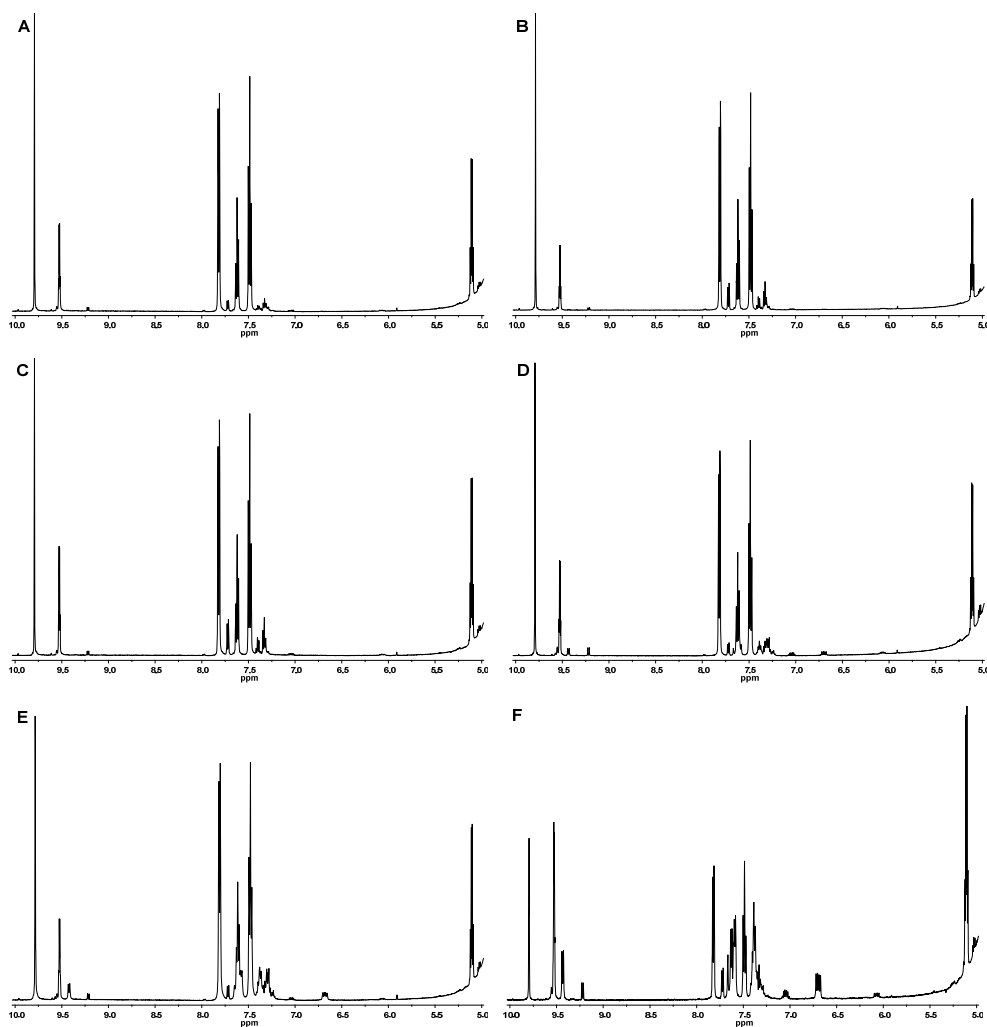
**Figure 2.3.** Partial ^{1}H NMR spectra of the condensation of **3** with **10** catalyzed by 4-OT P1A (A), 4-OT R11A (B), 4-OT inactivated with 3-bromopyruvate (C), synthetic 4-OT (D), 4-OT R39A (E), and 4-OT L8R (F). Signals corresponding to the cross-condensation product cinnamaldehyde (**15**) are observed only in spectra D-F, whereas signals corresponding to the self-condensation product 2-butenal (**16**) are observed in all spectra. The ^{1}H NMR signals for **3**, **10**, **15** and **16** are reported in the experimental procedures.

active site residue Leu-8 with an arginine.[15] This observation prompted us to test whether this mutation may also affect the promiscuous aldolase activity of 4-OT. Accordingly, the 4-OT L8R mutant was constructed, purified, and tested for its ability to catalyze the aldol condensation of **3** with **10**. To perform accurate rate measurements, an UV spectroscopy assay was developed that follows the decrease in substrate absorbance at 250 nm and the increase in product absorbance at 290 nm. Incubation of **3** and **10** with 4-OT L8R resulted in a decrease in absorbance corresponding to **10** ($\lambda_{max}$ = 250 nm), accompanied by the appearance of a new absorbance peak at 290 nm, which corresponds to **15** (Figure 2.4B). When the rate of this reaction is compared to that catalyzed by wild-type 4-OT (Figure 2.4A), it is clear that the 4-OT L8R mutant has an improved aldolase activity. Analysis of the reaction by ^{1}H NMR spectroscopy verified that the product of the aldol condensation of **3** with **10**, catalyzed by 4-OT L8R, is **15**, as established for wild-type 4-OT (Figure 2.3F). Again, no product formation was detected for the control reaction without enzyme.
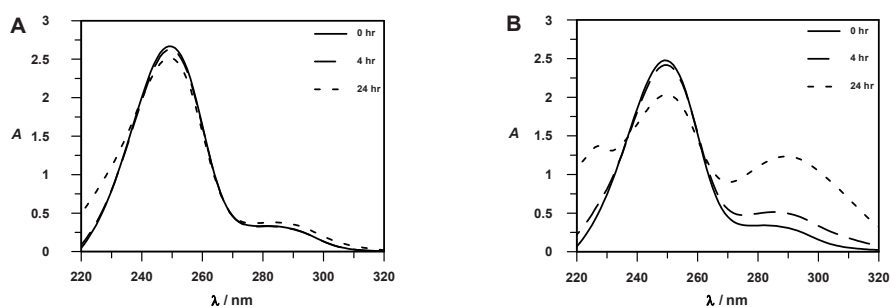


**Figure 2.4.** UV spectra monitoring the aldol condensation of **3** with **10** catalyzed by either 4-OT (A) or 4-OT L8R(B).

Having established that the 4-OT L8R mutant has a significantly improved aldolase activity, kinetic parameters were determined. The rates of the reactions catalyzed by 4-OT and 4-OT L8R were dependent on the concentrations of both **3** and **10** (Figure 2.5). Apparent $k_{cat}/K_m$ values were estimated at a fixed concentration of **3** (50 mM) and varying concentrations of **10**. A comparison of these values shows that the 4-OT L8R mutant ($k_{cat}/K_m$ = $1.4 \times 10^{-2}$ M$^{-1}$ s$^{-1}$) is 16-fold more efficient in catalyzing the aldol condensation of **3** with **10** than wild-type 4-OT ($k_{cat}/K_m$ = $8.5 \times 10^{-4}$ M$^{-1}$ s$^{-1}$). The improved aldolase activity of 4-OT L8R can also be clearly visualized using a phloroglucinol-based colorimetric assay,[16] which monitors the formation of **15** (Figure 2.6).

**Figure 2.5.** Kinetic analysis of the condensation of **3** with **10** catalyzed by 4-OT L8R. A) A series of Lineweaver Burk plots of the reaction of **3** (AA) and **10** (BZ) catalyzed by 4-OT L8R (292 µM). The concentration of **10** is varying from 1 to 8 mM and the concentration of **3** is varying from 10 to 50 mM. B) Michaelis-Menten plot of the reaction of **3** and **10** catalyzed by 4-OT L8R (145 µM). The initial rates were measured at a fixed concentration of **3** (50 mM) and varying concentrations of **10** (0.5 to 2 mM). The slope of this plot 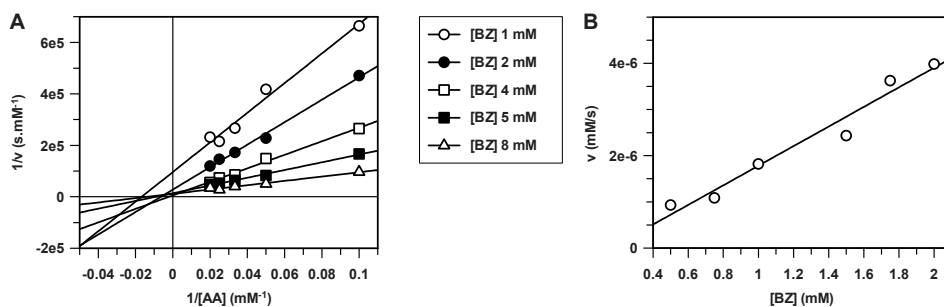is a straight line that equals ($k_{cat} \times$ [E])/$K_m$. Dividing the slope by the enzyme concentration results in a value for the apparent $k_{cat}/K_m$ (for **10**). For wild-type 4-OT, the apparent $k_{cat}/K_m$ was determined in the same way.
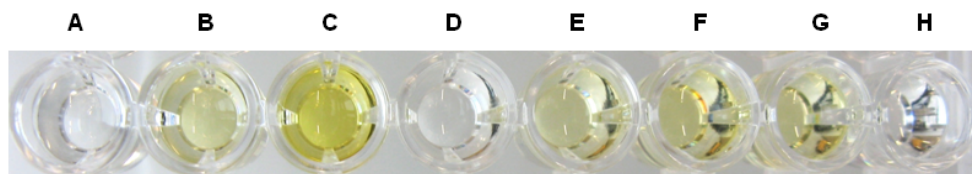


**Figure 2.6.** Phloroglucinol-based colorimetric assay for the detection of cinnamaldehyde (**15**) in reaction mixtures of **3** and **10** incubated for 3 days with either A) an aliquot from a mock purification, B) 4-OT, C) 4-OT L8R, D) 4-OT P1A, E) 4-OT R11A, F) 4-OT R39A, G) synthetic 4-OT, and H) no enzyme added. A yellow color indicates the presence of **15**.

In a possible mechanism that might explain the observed aldolase activity, Pro-1 functions as a nucleophile and attacks the carbonyl carbon of **3** to form an iminium ion (Scheme 2.5). Deprotonation of this ion leads to the generation of the reactive enamine intermediate, which is equivalent to a nucleophilic carbanion. Nucleophilic addition of the enamine to the carbonyl carbon of **10** gives a modified iminium ion. Dehydration and hydrolysis of this iminium ion intermediate yields the final product **15**. The negative charge that develops on the carbonyl oxygen of **10** (and/or **3**) could be stabilized by the active site arginine (i.e., Arg-11), which is known to be critical for the activity. This arginine may also

serve as the source for a proton that would assist in formation of the hydroxyl group of the iminium ion intermediate. The additional arginine residue in the 4-OT L8R mutant may assist Arg-11 in charge stabilization, explaining the increased aldolase activity of this mutant enzyme.[17]
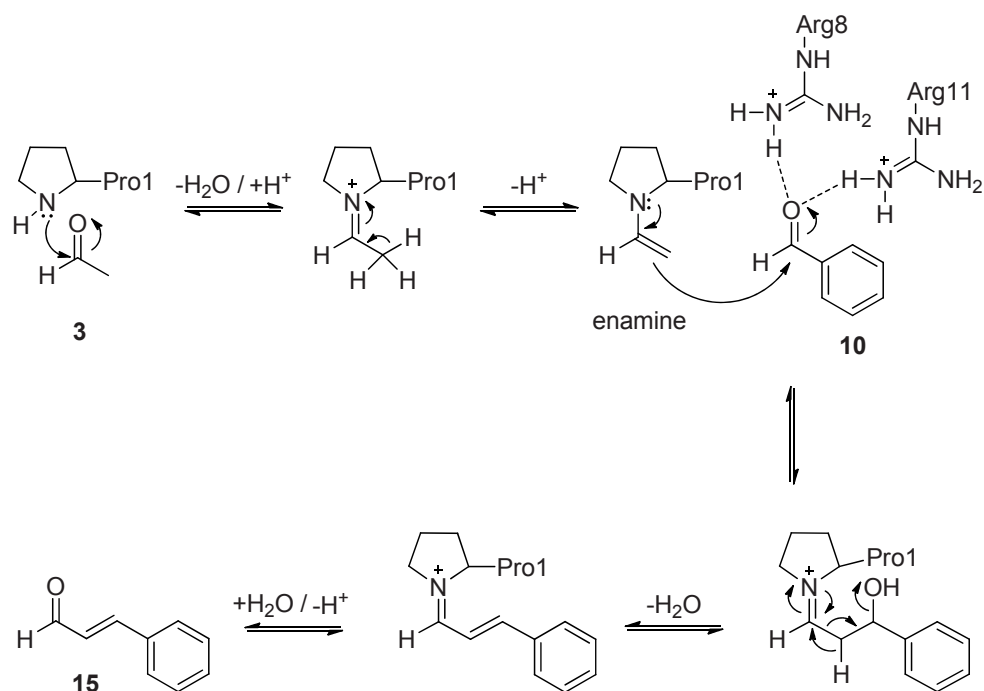


**Scheme 2.5.** Proposed mechanism for the 4-OT-catalyzed condensation of **3** with **10** to yield **15**.

Lineweaver-Burk plots at varying concentrations of **3** and different fixed values of **10** give a series of intersecting lines (Figure 2.5A). This pattern is characteristic of a sequential bi-substrate mechanism. Because only substrate **3** can serve as enamine donor (**10** does not have enolisable hydrogens and can only act as enamine acceptor in the aldol reaction), the mechanism is expected to be sequential ordered rather than sequential random. This is supported by the relative rate of imine formation between 4-OT and **3** or **10**. After incubation of 4-OT with either **3** or **10** (each at 1 mM) for 1 h at 22 °C, and subsequent reduction with NaCNBH$_3$, 4-OT is found to be labeled only partially by **10**, whereas covalent modification by **3** is almost complete (Figure 2.7). Taken together, the kinetic and labeling results are consistent with a sequentially ordered mechanism for the 4-OT-catalyzed condensation of **3** with **10**.
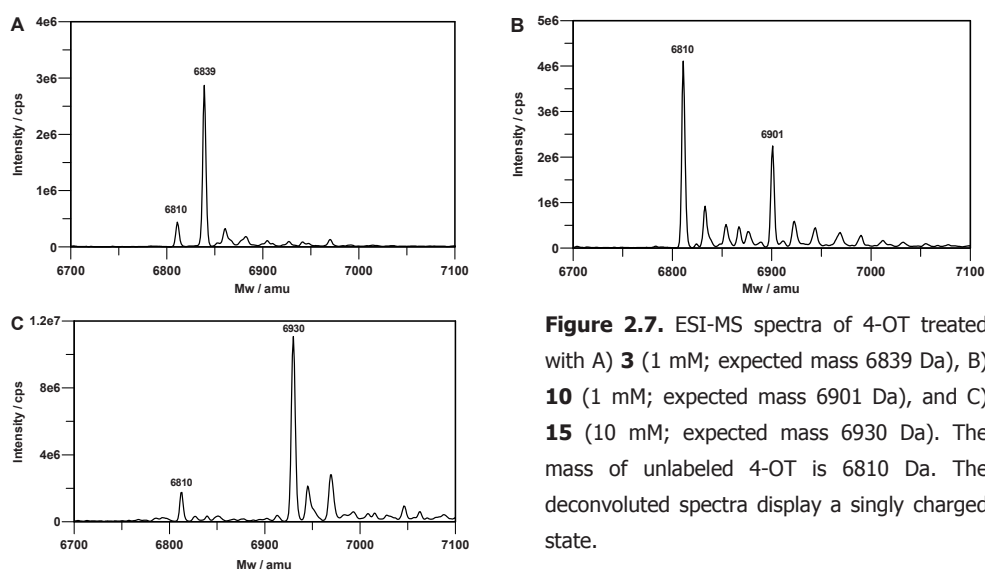
**Figure 2.7.** ESI-MS spectra of 4-OT treated with A) **3** (1 mM; expected mass 6839 Da), B) **10** (1 mM; expected mass 6901 Da), and C) **15** (10 mM; expected mass 6930 Da). The mass of unlabeled 4-OT is 6810 Da. The deconvoluted spectra display a singly charged state.

Notably, the UV spectra monitoring the aldol reactions (Figure 2.4) also show a time-dependent increase in absorbance at ~227 nm. Control experiments in which the nonenzymatic self-condensation of **3** was followed by UV and ^{1}H NMR spectroscopy demonstrate that this absorbance corresponds to the formation of but-2-enal (**16**, Scheme 2.6). Although 4-OT might weakly catalyze this reaction, the activity appears to be more significant for the 4-OT L8R mutant (Figure 2.4). Indeed, additional signals corresponding to **16** were also observed in the ^{1}H NMR spectra monitoring the incubations of **3** and **10** with either 4-OT (Figure 2.1) or the 4-OT L8R mutant (Figure 2.3F). Additional experiments with synthetically prepared 4-OT and 4-OT L8R are underway to verify whether these enzymes also catalyze the self-condensation of **3** (rather than a contaminating protein).
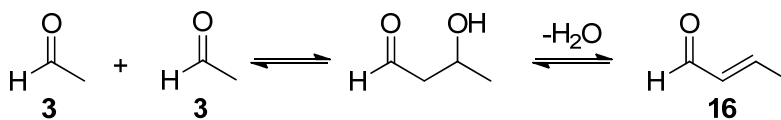


**Scheme 2.6.** The self-condensation of acetaldehyde (**3**) to yield but-2-enal (**16**).

## Conclusion

In conclusion, we have used a systematic screening strategy to discover new promiscuous activities in wild-type 4-OT. The key strength of this strategy is the observation that the same amino-terminal proline that is used to catalyze proton transfer (in 4-OT's natural activity) can be used to generate a range of structurally distinct enamines that might undergo reactions with various electrophiles. As a proof of principle experiment, we have shown that the aldol condensation of **3** with **10** is catalyzed by 4-OT, and that this activity can be improved (16-fold) by a single point mutation (L8R) in the enzyme's active site. The proposed mechanism of this reaction mimicks that used by natural class I aldolases,[18] catalytic aldolase antibodies,[19] and designed aldolase peptides.[20] However, an important mechanistic difference is that these catalysts use the primary amine of a lysine residue to form enamines with carbonyl substrates, whereas 4-OT uses the secondary amine of an active site proline as the nucleophile catalyst.

Systematic screening of a protein scaffold in which an active site proline residue is present as a nucleophile and that has the reactivity to form enamines might prove to be a useful approach to discover new promiscuous carbonyl transformation activities that could serve as starting activities to develop new biocatalysts for carbon-carbon bond formation. This approach does not copy something in Nature (i.e., proline-based enamine catalysis is not known to occur in Nature), but is based on the synthetic requirements of the desired bond-forming reactions.[21] Work is in progress to explore new enamine donors for 4-OT and to exploit the chemical potential of the various enamines that can be formed in 4-OT's active site in order to find new promiscuous aldol, alkylation and Michael addition reactions in this fascinating enzyme. In addition, we have initiated studies aimed at screening known and putative tautomerase superfamily members (that share a nucleophilic active site proline) with various aldehyde and ketone probes for the presence of intrinsic, promiscuous carbonyl transformation activities. Directed evolution experiments could then be used to enhance the desired activities to a practical level.

# Experimental Procedures

*Materials:* All chemicals were obtained from Sigma-Aldrich Chemical Co. (St. Louis, MO), unless stated otherwise. The sources for the biochemicals, buffers, solvents, components of Luria-Bertani (LB) media as well as the materials, enzymes, and reagents used in the molecular biology procedures are reported elsewhere.[22] Sequencing grade endoproteinase Glu-C (protease V-8) was purchased from Roche Diagnostics (Mannheim, Germany). High-purity synthetic 4-OT (lyophilized, with Met-45 replaced by Nle to prevent oxidation upon sample handling) was obtained from GenScript USA Inc. (Piscataway, NY) and folded into the active homohexamer as described before.[23] 2-Hydroxyhexa-2,4-dienedioate (commonly known as 2-hydroxymuconate) was kindly provided by Prof. Dr. Christian P. Whitman (University of Texas at Austin, TX).

*General methods:* Techniques for restriction enzyme digestions, ligation, transformation, and other standard molecular biology manipulations were based on methods described elsewhere.[24] The PCR was carried out in a DNA thermal cycler (model GS-1) obtained from Biolegio (Nijmegen, The Netherlands). DNA sequencing was performed by Macrogen (Seoul, Korea). Protein was analyzed by polyacrylamide gel electrophoresis (PAGE) using sodium dodecyl sulfate (SDS) gels containing polyacrylamide (10%). The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the Waddell method.[25] Kinetic data were obtained on a V-650 spectrophotometer from Jasco (IJsselstein, The Netherlands). $^1$H 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) and are referenced to $H_2O$ (δ = 4.67). GC-MS spectra were recorded on an in-house Shimadzu GC-17A/GCMS-QP5000 system. The masses of 4-OT and 4-OT mutants were determined by ESI-MS using a Sciex API 3000 triple quadrupole mass spectrometer (AB Sciex, Concord, Ontario, Canada). The masses of peptide fragments in samples of Glu-C-digested 4-OT were determined by nano-LC-MS/MS using a QSTAR XL mass spectrometer (AB/MDS-Sciex, Toronto, Canada) coupled to an Agilent 1100 nanoflow system (Waldbronn, Germany). These mass spectrometers are housed in the Mass Spectrometry Facility Core in the Department of Pharmacy at the University of Groningen.

*Construction of the expression vector for the production of 4-OT:* The gene encoding wild-type 4-OT was amplified from plasmid pET3b(4-OT)[15] using the following primers: 5'- A TAG CAG GTA <u>CAT ATG</u> CCT ATT GCC CAG ATC CAC AT-3' (primer F, *Nde*I site underlined) and 5'- G TGA TGT TAT <u>GGA TCC</u> TCA GCG TCT GAC CTT GCT GGC CAG TTC G-3' (primer R, *Bam*HI site underlined). The PCR product was gel-purified, digested with *Nde*I and *Bam*HI, and cloned in frame with the ATG start codon of the pET20b(+) vector. The newly constructed expression vector was named pET20b(4-OT).

*Construction of the 4-OT mutants P1A, R11A, R39A and L8R:* The four 4-OT mutants were constructed by PCR using the coding sequence for 4-OT in plasmid pET20b(4-OT) as the template. To introduce the P1A, L8R, and R11A mutations, the following oligonucleotides were used as forward primers: 5'-A TAG CAG GTA <u>CAT ATG</u> **GCC** ATT GCC CAG ATC CAC ATC CTT G-3', 5'-A TAG CAG GTC <u>CAT ATG</u> CCT ATT GCC CAG ATC CAC ATC **CGC** GAA GGC CGC AGC-3', and 5'- A TAG CAG GTA <u>CAT</u>

ATG CCT ATT GCC CAG ATC CAC ATC CTT GAA GGC **GCC** AGC GAC GAG CAG-3′, respectively, where the *Nde*I site is underlined and the mutated codon is in bold. These primers correspond to the 5′-end of the wild-type coding sequence and were used in combination with primer R. For the construction of the R39A mutant the following oligonucleotide was used as the reverse primer: 5′-A TGT TAT GGA TCC TCA GCG TCT GAC CTT GCT GGC CAG TTC GCC GCC GAT GCC GAA GTG GCC CTT GGC CAT CTC CGT GAT AAT CAC **GGC** CAC GCT GG-3′, where the *Bam*HI site is underlined and the mutated codon is in bold. This primer corresponds to the 3′-end of the wild-type coding sequence and was used in combination with primer F. The resulting PCR products were gel-purified, digested and cloned in frame with the ATG start codon of the pET20b(+) vector. All genes were fully sequenced to assure that only the desired mutations were introduced.

*Expression and purification of 4-OT wild-type and mutants:* The 4-OT enzyme, either wild-type or mutant, was produced in *E. coli* BL21(DE3) using the pET20b(+) expression system. LB medium (5 mL) containing ampicillin (Ap, 100 µg/mL) was inoculated with cells from a glycerol stock of *E. coli* BL21(DE3) containing the appropriate expression vector using a sterile loop. After overnight growth at 37 ℃, this culture was used to inoculate fresh LB/Ap medium (1 L) in a 3 L Erlenmeyer flask. Cultures were grown overnight at 37 ℃ with vigorous shaking to an $OD_{600}$ of ~4.5. Cells were harvested by centrifugation (20 min, 4500 rpm), washed with 10 mM $NaH_2PO_4$ buffer (pH 8.0) (Buffer A) and stored at -20 ℃.

The 4-OT protein, either wild-type or mutant, was purified using disposable hand-packed columns. Typically, in this protocol, cells from 0.5 L of culture were suspended in ~10 mL of Buffer A, sonicated and centrifuged as described above. Subsequently, the supernatant was loaded onto a DEAE-sepharose column (10 × 1.0 cm filled with ~8 mL of resin) that had been previously equilibrated with Buffer A. The column was first washed with Buffer A (3 × 10 mL) and then the protein was eluted by gravity flow using Buffer A containing 0.5 M $Na_2SO_4$ (8 mL). Fractions (~1.5 mL) were collected and 4-OT was identified by SDS-PAGE. The appropriate fractions were pooled and made 1.6 M in $(NH_4)_2SO_4$. After stirring for 2 hr at 4 ℃, the precipitate was removed by centrifugation (20 min at 13,200 rpm), and the supernatant was filtered and loaded onto a phenyl-sepharose column (10 × 1.0 cm filled with ~8 mL of resin) that had been previously equilibrated with Buffer A containing 1.6 M $(NH_4)_2SO_4$. The column was first washed with the loading buffer (3 × 10 mL) and then the protein was eluted by gravity flow using Buffer A (8 mL). Fractions (~1.5 mL) were collected and analyzed as described above, and those that contained purified 4-OT protein were combined and the buffer was exchanged against 20 mM $NaH_2PO_4$ buffer (pH 7.3) using a pre-packed PD-10 Sephadex G-25 gelfiltration column. The purified protein was stored at -80 °C until further use.

Wild-type 4-OT and the P1A and R11A mutant proteins had little interaction with the phenyl-sepharose column and eluted as homogeneous proteins (>95% purity as assessed by SDS-PAGE) in the first washing step. The R39A mutant interacted more strongly with the phenyl-sepharose column and eluted in the second and third washing step, as well as in one of the first elution fractions. The R39A protein that eluted in the washing steps (~85% purity) was used in the assays. The L8R mutant eluted from the DEAE-sepharose column in the second washing step as homogenous protein (>95% purity), and no further purification on the phenyl-sepharose column was performed.

The masses of the purified wild-type and mutant proteins were determined with ESI-MS to confirm that the proteins had been processed correctly and the initiating methionine had been removed. The observed monomer mass for wild-type 4-OT was 6811 Da (calc. 6810 Da). The observed monomer mass for the P1A mutant was 6786 Da (calc. 6785 Da), that of the R11A mutant was 6727 Da (calc. 6726 Da), and that of the R39A mutant was 6727 Da (calc. 6726 Da).

In addition, a mock purification was performed from BL21(DE3) cells harboring an empty pET20b(+) vector according to the procedure described above for wild-type 4-OT. An aliquot of this sample was used as a control in the colorimetric and $^1$H NMR assays for aldolase activity (see below).

*Spectrophotometric assay for tautomerase activity:* The ketonization of 2-hydroxymuconate (**1**) by 4-OT was monitored by following the depletion of the absorbance of **1** at 288 nm ($\varepsilon = 20$ mM$^{-1}$ cm$^{-1}$) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3). A small aliquot (1 µL) of a 1.1 µM (monomer concentration) stock solution of 4-OT was added to a cuvette containing 1 mL of buffer. To initiate the assay, 2 µL of a stock solution of **1** (50 mM in ethanol) was added.

*Labeling of 4-OT with 3-bromopyruvate:* An amount of 4-OT (2 mg) was incubated with 3-bromopyruvate (3-BP; final concentration 20 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) for 1 h at 22 °C (total volume of 1 mL). Subsequently, NaBH$_4$ was added to a final concentration of 25 mM and the sample was incubated at 22 °C for 1 h. In a separate control experiment, the same quantity of 4-OT was incubated without inhibitor under otherwise identical conditions. The two incubation mixtures were loaded onto separate PD-10 Sephadex G-25 gel filtration columns, which had previously been equilibrated with 20 mM NaH$_2$PO$_4$ buffer (pH 7.3). The proteins were eluted by gravity flow using the same buffer. Fractions (~0.5 mL) were analyzed for the presence of protein by UV absorbance at 214 nm. The appropriate fractions containing the purified proteins were combined and assayed for tautomerase activity as described above. The 4-OT sample treated with 3-BP had no residual tautomerase activity. Incubation of 4-OT without inhibitor under the same conditions had no effect on activity. The covalent modification of 4-OT resulting from the incubation with 3-BP was confirmed by ESI-MS. The observed monomer mass for the covalently modified 4-OT was 6901 Da (calc. 6900 Da). This 4-OT protein inactivated by 3-BP was used as a control in the $^1$H NMR assay for aldolase activity (see below).

*Sodium cyanoborohydride treatment of 4-OT in the presence of carbonyl compounds (**3**-**10**, **15**):* An amount of 4-OT (1 mg) was incubated with 1 mM of compounds **3**-**7**, **10**, or **15** or 10 mM of compounds **8** or **9** in a final volume of 1 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) for 1-3 h at 22 °C (in separate reactions). Subsequently, an aliquot of a 100 mM stock solution of NaCNBH$_3$ in water was added to give a final concentration of 25 mM. After incubation for 1 h at 22 °C, the buffer was exchanged against 5 mM NH$_4$HCO$_2$ buffer (pH 8.0) using a pre-packed PD-10 Sephadex G-25 gelfiltration column. The purified 4-OT proteins were assayed for residual tautomerase activity using the spectrophotometric assay described above. To assess the extent and specificity of the covalent labeling, the purified proteins were also analyzed by ESI-MS.

Control reactions containing enzyme, buffer, and carbonyl compound, or enzyme, buffer, and NaCNBH$_3$ were carried out under identical conditions. These mixtures did not lead to inactivation of 4-OT.

*Mass spectral analysis of modified 4-OT and peptide mapping:* An amount of 4-OT (0.5 mg) was incubated with 1 mM of acetaldehyde (**3**) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) for 1 h at 22 °C (total volume of 1 mL). A second 4-OT sample was not treated with **3** and was used as the control sample. An aliquot from a 100 mM stock solution of NaCNBH$_3$ in water was added to both samples to give a final concentration of 25 mM. After incubation for 1 h at 22 °C, the buffer was exchanged against 10 mM NaH$_2$PO$_4$ buffer (pH 8.0) using a pre-packed PD-10 Sephadex G-25 gelfiltration column. The two purified 4-OT proteins were assayed for tautomerase activity, analyzed by ESI-MS, and used in the following peptide mapping experiments.

For the peptide mapping studies, a quantity (50 µg) of unmodified 4-OT and 4-OT modified by **3** was vacuum-dried. The individual protein pellets from the two samples were dissolved in 10 µL of 10 M guanidine-HCl and incubated for 2 hrs at 37 °C. Subsequently, the samples were diluted by addition of 90 µL of 100 mM NH$_4$HCO$_3$ buffer (pH 8.0) and incubated for 2 days at 37 °C with protease GluC (2.5 µL from a 10 mg/mL stock solution in water). These two digested samples were analyzed by nano-LC/MS to identify the labeled peptide fragment. Selected ions of both samples were subjected to MS/MS analysis.

*$^1$H NMR spectroscopic screening for carbonyl transformations by 4-OT:* The $^1$H NMR spectra montoring the reactions between **3** and the four selected electrophiles were recorded as follows. A reaction mixture of **3** (50 mM) and either benzaldehyde (**10**), acetophenone (**11**), cyclohexanecarboxyaldehyde (**12**) or cyclopentanecarboxyaldehyde (**13**) (each at ~50 mM) in 20 mM NaH$_2$PO$_4$ buffer (0.55 mL, pH 7.3) was placed in an NMR tube, along with D$_2$O (0.05 mL) and 0.4 mg 4-OT (0.05 mL from a 8 mg/mL solution). Similar mixtures without 4-OT (the control samples) were prepared as well in order to analyze the non-enzymatic (uncatalyzed) reaction. $^1$H NMR spectra were recorded directly after mixing, and then after 1, 7 and 14 days.

The $^1$H NMR signals for **3**, **10**, **11**, **12**, and **13** are as follows. $^1$H NMR (500 MHz, 20 mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3) of **3** and its hydrate: δ = 9.52 (q, $J$ = 3.0 Hz, 1H), 5.11 (q, $J$ = 5.0 Hz, 1H), 2.09 (d, $J$ = 3.0 Hz, 3H), 1.19 (d, $J$ = 4.5 Hz, 3H). $^1$H NMR (500 MHz, 20 mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3) of **10**: δ = 9.77 (s, 1H), 7.79 (d, $J$ = 8.0 Hz, 2H), 7.60 (t, $J$ = 7.5 Hz, 1H), 7.46 (t, $J$ = 7.5 Hz, 2H). $^1$H NMR (500 MHz, 20 mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3) of **11**: δ = 7.85 (d, $J$ = 8.5 Hz, 2H), 7.55 (t, $J$ = 7.5 Hz, 1H), 7.42 (t, $J$ = 7.5 Hz, 2H), 2.52 (s, 3H). $^1$H NMR (500 MHz, 20 mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3) of **12**: δ = 9.41 (s, 1H), 1.77 (m, 1H), 1.50-1.70 (m, 4H), 1.06-1.40 (m, 6H). $^1$H NMR (500 MHz, 20 mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3) of **13**: δ = 9.45 (d, $J$ = 2.0 Hz, 1H), 2.27 (m, 1H), 1.58-1.76 (m, 2H), 1.20-1.54 (m, 6H).

In the incubation containing **3**, **10**, and 4-OT, extra signals were detected after 1 day, indicating the formation of cinnamaldehyde (**15**). The $^1$H NMR signals for **15** are as follows: $^1$H NMR (500 MHz, 20 mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3) of *trans*-**15**: δ = 9.44 (d, $J$ = 8 Hz, 1H), 7.66 (d, $J$ = 16 Hz, 1H), 7.59 (d, $J$ = 6.5 Hz, 2H), 7.38 (m, 2H), 7.29 (m, 1H), 6.70 (dd, $J$ = 8 Hz, 1H). Furthermore, in all samples (with or without enzyme) the following extra signals were detected after 1 week: $^1$H NMR (500 MHz, 20

mM NaH$_2$PO$_4$/D$_2$O buffer, pH 7.3): δ = 9.22 (d, *J* = 8.5 Hz, 1H), 7.03 (m, 1H), 6.07 (dd, *J* = 8.0 Hz, 1H), 1.91 (d, *J* = 7.0 Hz, 3H). These signals correspond to the formation of but-2-enal (**16**). This self-condensation product of **3** is formed in slightly higher amounts in the incubations containing 4-OT as compared to the control reactions without enzyme. Reference spectra of **3** (and its hydrate), **10**, **15**, and **16** are given below in Figure 2.8.
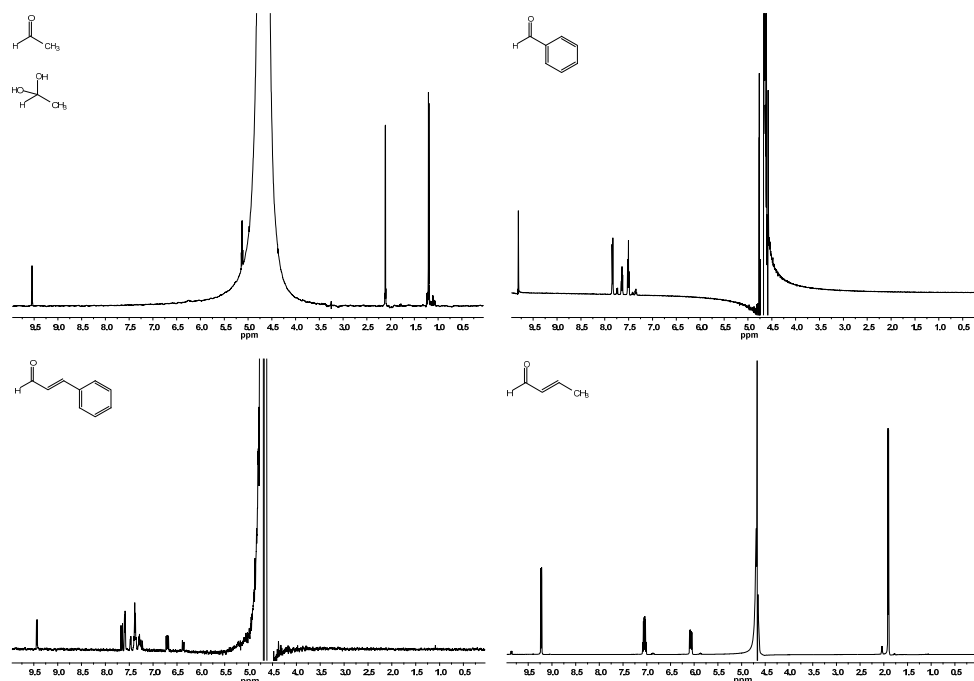


**Figure 2.8.** $^1$H NMR reference spectra of compounds **3**, **10**, **15** and **16**.

*Detection of **15** by UV spectroscopy and GC/MS analysis:* While **15** has characteristic $^1$H NMR signals, its identity in the incubation mixture described above was further confirmed by UV spectroscopy and GC/MS analysis. Accordingly, after 14 days of incubation, a small aliquot was removed from the mixture containing **3**, **10**, and 4-OT and diluted 200-fold in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3). In addition, a small aliquot was removed from the control mixture (**3** and **10** incubated without enzyme) and also diluted 200-fold in the same buffer. Subsequently, UV/Vis spectra were recorded from the diluted samples. Apart from the characteristic absorbance peak of **10** ($\lambda_{max}$ = 250 nm, $\varepsilon_{250}$= 15 mM$^{-1}$ cm$^{-1}$), the sample containing 4-OT showed extra peaks at $\lambda_{max}$ = 290 nm (**15**, $\varepsilon_{290}$= 26.7 mM$^{-1}$ cm$^{-1}$) and around $\lambda_{max}$ = 227 nm (**16**, $\varepsilon_{227}$= 19 mM$^{-1}$ cm$^{-1}$). The absorbance peak corresponding to **15** was lacking in the control sample without enzyme.

For detection of **15** by GC/MS analysis, the remaining part (~0.6 mL) of the reaction mixture containing **3**, **10**, and 4-OT was removed from the NMR tube and extracted with 1.8 mL of ethylacetate. The ethylacetate layer was dried over MgSO$_4$ and subsequently analyzed by GC/MS. The control sample (**3** and **10** incubated without enzyme) was prepared and analyzed in the same way, but did not show the presence of **15**.

*$^1$H NMR spectroscopy assay for aldolase activity:* $^1$H NMR spectra monitoring the aldol condensation of **3** with **10** catalyzed by either wild-type 4-OT, 4-OT L8R, 4-OT P1A, 4-OT R11A, 4-OT R39A, 4-OT inactivated by 3-BP, or synthetic 4-OT, were recorded as follows. In an NMR tube, the enzyme (90 µM) was incubated with **3** and **10** (50 mM each, unless stated otherwise) in 0.6 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) at 22 °C. In addition, to each tube 0.05 mL of D$_2$O was added. In two additional control experiments, **3** and **10** were incubated without enzyme or with an aliquot from a mock purification under otherwise identical conditions. The first $^1$H NMR spectrum was recorded immediately after mixing, and then after 7 and 14 days. The formation of **15** is indicative of the presence of aldolase activity. The $^1$H NMR signals for **3** (and its hydrate), **10** and **15** are described above.

*Colorimetric assay for aldolase activity:* Purified wild-type 4-OT, 4-OT mutants, and synthetic 4-OT were assayed for aldolase activity by monitoring production of cinnamaldehyde (**15**) upon incubation with acetaldehyde (**3**) and benzaldehyde (**10**). Accordingly, an amount (200 µg) of wild-type 4-OT, 4-OT P1A, 4-OT R11A, 4-OT R39A, 4-OT L8R, or synthetic 4-OT was incubated (in separate vials) with **3** and **10** (30 mM each) in 1.2 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) at 22 °C. In two separate control experiments, **3** and **10** were incubated without enzyme or with an aliquot of a mock purification. After incubation of the reaction mixtures at 22 °C for 3 days, a sample of 50 µL was removed and mixed with 150 µL of 0.2% (w/v) phloroglucinol in 25/75 % (v/v) HCl/EtOH. Compound **15** forms a short-lived yellow-colored complex with phloroglucinol, indicative of the presence of aldolase activity.

*UV spectrophotometric assay for aldolase activity:* The kinetic assays were performed at 22 °C by following the increase in absorbance at 290 nm, which corresponds to the formation of **15**. An aliquot of 4-OT was added to 0.3 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) in a 1 mm cuvette, yielding a final enzyme concentration of 587 µM (146 µM for the 4-OT L8R mutant). The enzyme activity was assayed by the addition of **3** (at a fixed concentration of 50 mM) and **10** (in concentrations varying from 0.5 to 10 mM). The initial rates (mM/s) were plotted versus the concentration of **10** (mM) (Figure 2.5). The slope of this plot is a straight line that equals $(k_{cat} \times [E])/K_m$. Dividing the slope by the enzyme concentration results in a value for the apparent $k_{cat}/K_m$ (for **10**).

Additional kinetic assays using the 4-OT L8R mutant were performed to demonstrate that the initial rate is dependent on both substrate concentrations, and to verify that the kinetic mechanism is sequential. An aliquot of 4-OT L8R was added to 0.3 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) in a 1 mm cuvette, yielding a final enzyme concentration of 292 µM . The enzyme activity was assayed by the addition of **3** (in concentrations varying from 10 to 50 mM) and **10** (in concentrations varying from 1 to 8 mM). The resulting data were plotted as the reciprocal value of the initial rate (s/mM) versus the reciprocal value of the concentration of **3** (mM$^{-1}$) (Figure 2.5).

## Acknowledgements

## References

1. a) U. T. Bornscheuer, R. J. Kazlauskas, *Angew. Chem. Int. Ed.* **2004**, *43*, 6032-6040; b) R. J. Kazlauskas, *Curr. Opin. Chem. Biol.* **2005**, *9*, 195-201.

2. A. Yarnell, *CENEAR* **2003**, *81* (49), 33-35.

3. a) D. R. J. Palmer, J. B. Garrett, V. Sharma, R. Meganathan, P. C. Babbitt, J. A. Gerlt, *Biochemistry* **1999**, *38*, 4252-4258; b) L. C. James, D. S. Tawfik, *Protein Sci.* **2001**, *10*, 2600-2607; c) P. J. O'Brien, D. Herschlag, *Biochemistry* **2001**, *40*, 5691-5699; d) S. C. Wang, W. H. Johnson, Jr., C. P. Whitman, *J. Am. Chem. Soc.* **2003**, *125*, 14282-14283; e) D. M. Schmidt, E. C. Mundorff, M. Dojka, E. Bermudez, J. E. Ness, S. Govindarajan, P. C. Babbitt, J. Minshull, J. A. Gerlt, *Biochemistry* **2003**, *42*, 8387-8393; f) G. J. Poelarends, H. Serrano, W. H. Johnson, Jr., D. W. Hoffman, C. P. Whitman, *J. Am. Chem. Soc.* **2004**, *126*, 15658-15659; g) W. S. Yew, J. Akana, E. L. Wise, I. Rayment, J. A. Gerlt, *Biochemistry* **2005**, *44*, 1807-1815; h) C. Roodveldt, D. S. Tawfik, *Biochemistry* **2005**, *44*, 12728-12736; i) O. Khersonsky, C. Roodveldt, D. S. Tawfik, *Curr. Opin. Chem. Biol.* **2006**, *10*, 498-508; j) L. F. Olguin, S. E. Askew, A. C. O'Donoghue, F. Hollfelder, *J. Am. Chem. Soc.* **2008**, *130*, 16547–16555; k) B. van Loo, S. Jonas, A. C. Babtie, A. Benjdia, O. Berteau, M. Hyvönen, F. Hollfelder, *Proc. Natl. Acad. Sci. U S A.* **2010**, *107*, 2740-2745.

4. a) S. Harayama, M. Rekik, K. L. Ngai, L. N. Ornston, *J. Bacteriol.* **1989**, *171*, 6251-6258; b) C. P. Whitman, B. A. Aird, W. R. Gillespie, N. J. Stolowich, *J. Am. Chem. Soc.* **1991**, *113*, 3154-3162; c) L. H., Chen, G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour, C. P. Whitman, *J. Biol. Chem.* **1992**, *267*, 17716-17721.

5. a) C. P. Whitman, *Arch. Biochem. Biophys.* **2002**, *402*, 1-13; b) G. J. Poelarends, V. Puthan Veetil, C. P. Whitman, *Cell. Mol. Life Sci.* **2008**, *65*, 3606-3618.

6. a) S. C. Wang, W. H. Johnson, R. M. Czerwinski, S. L. Stamps, C. P. Whitman, *Biochemistry* **2007**, *46*, 11919-11929; b) E. A. Burks, C. D. Fleming, A. D. Mesecar, C. P. Whitman, S. D. Pegan, *Biochemistry* **2010**, *49*, 5016-5027.

7. T. K. Harris, R. M. Czerwinski, W. H. Jr. Johnson, P. M. Legler, C. Abeygunawardana, M. A. Massiah, J. T. Stivers, C. P. Whitman, A. S. Mildvan, *Biochemistry* **1999**, *38*, 12343-12357.

8. A. Brik, L. J. D'Souza, E. Keinan, F. Grynszpan, P. E. Dawson, *ChemBioChem* **2002**, *3*, 845-851.

9. J. T. Stivers, C. Abeygunawardana, A. S. Mildvan, G. Hajipour, C. P. Whitman, *Biochemistry* **1996**, *35*, 814-823.

10. P.Y. Bruice, *J. Am. Chem. Soc.* **1989**, *111*, 962-970.

11. S. Mukherjee, J. W. Yang, S. Hoffmann, B. List, *Chem. Rev.* **2007**, *107*, 5471-5569, and references therein.

12. I. Fridovich, F. H. Westheimer, *J. Am. Chem. Soc.* **1962**, *84*, 3208-3209.

13. J. T. Stivers, C. Abeygunawardana, A. S. Mildvan, G. Hajipour, C. P. Whitman, L. H. Chen, *Biochemistry* **1996**, *35*, 803-813.

14. 4-OT was synthesized by GenScript USA Inc. (Piscataway, NY).

15. G. J. Poelarends, J. J. Almrud, H. Serrano, J. E. Darty, W. H. Johnson, Jr., M. L. Hackert, C. P. Whitman, *Biochemistry* **2006**, *45*, 7700-7708.

16. a) F. Pomar, F. Merino, A. Ros Barceló, *Protoplasma* **2002**, *220*, 17–28; b) W. Zha, S. B. Rubin-Pitel, H. Zhao, *Mol. BioSyst.* **2008**, *4*, 246–248.

17. An alternative mechanism in which acetaldehyde will react via the derived enolate anion rather than via an enamine intermediate cannot be excluded. In such a mechanism, the critical Pro-1 residue would function as a general base catalyst removing a proton from acetaldehyde to give the reactive enolate. However, this alternate mechanism seems unlikely in view of the ability of Pro-1 to rapidly react with acetaldehyde to form an iminium ion.

18. T. D. Machajewski, C.-H. Wong, *Angew. Chem.* **2000**, *112*, 1406-1430; *Angew. Chem. Int. Ed.* **2000**, *39*, 1352-1374, and references therein.

19. J. Wagner, R. A. Lerner, C. F. Barbas III, *Science* **1995**, *270*, 1797-1800.

20. F. Tanaka, R. Fuller, C. F. Barbas III, *Biochemistry* **2005**, *44*, 7583-7592.

21. B. List, *Acc. Chem. Res.* **2004**, *37*, 548-557.

22. H. Raj, B. Weiner, V. Puthan Veetil, C. R. Reis, W. J. Quax, D. B. Janssen, B. L. Feringa, G. J. Poelarends, *ChemBioChem* **2009**, *10*, 2236-2245.

23. M. C. Fitzgerald, I. Chernushevich, K. G. Standing, S. B. H. Kent, C. P. Whitman, *J. Am. Chem. Soc.* **1995**, *117*, 11075-11080.

24. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., **1989**, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

25. W. J. Waddell, *J. Lab. Clin. Med.* **1956**, *48*, 311-314.

# 3

# Enhancement of the Promiscuous Aldolase and Dehydration Activities of 4-Oxalocrotonate Tautomerase by Protein Engineering

E. Zandvoort[#], E. M. Geertsema[#], W. J. Quax, G. J. Poelarends

## Abstract

The enzyme 4-oxalocrotonate tautomerase (4-OT) is catalytically promiscuous and catalyzes, in addition to its natural tautomerization reaction, the aldol condensation of acetaldehyde with benzaldehyde to yield cinnamaldehyde (Chapter 2). This reaction consists of two steps: the coupling of acetaldehyde with benzaldehyde to give 3-hydroxy-3-phenylpropanal, and the dehydration of this aldol compound resulting in cinnamaldehyde. It was not known whether both steps are catalyzed by 4-OT. Here, it is described that 4-OT not only catalyzes the initial cross-coupling of acetaldehyde with benzaldehyde to yield 3-hydroxy-3-phenylpropanal, but is also responsible for the subsequent dehydration of this aldol compound to yield cinnamaldehyde as final product. Furthermore, mechanism-inspired engineering provided an active site mutant (F50A) with strongly enhanced aldol condensation activity.
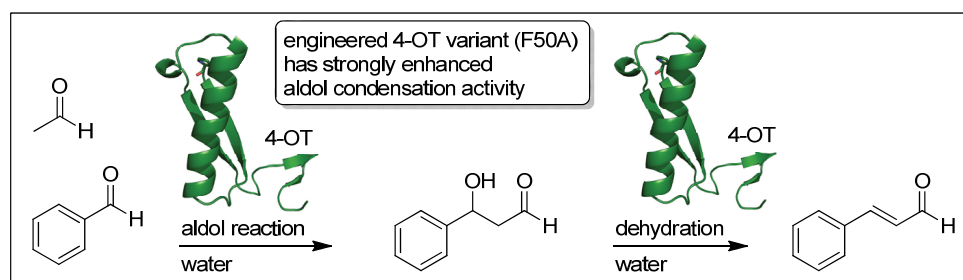
# Introduction

The enzyme 4-oxalocrotonate tautomerase (4-OT) catalyzes the conversion of 2-hydroxyhexa-2,4-dienedioate (**1**) to 2-oxohex-3-enedioate (**2**) (Scheme 3.1) as part of a degradation pathway for aromatic hydrocarbons in *Pseudomonas putida* mt-2.[1] 4-OT belongs to the tautomerase superfamily, a group of homologous proteins that share a characteristic β-α-β structural fold and a unique catalytic amino-terminal proline.[2] In the tautomerase activity of 4-OT, residue Pro-1 functions as a general base (p$K_a$ ~6.4) that shuttles the 2-hydroxyl proton of **1** to the C-5 position to give **2**.[3] In the course of screening for catalytic promiscuity in 4-OT, we recently found that this tautomerase promiscuously catalyzes the aldol condensation of acetaldehyde (**3**) with benzaldehyde (**4**) ($k_{cat}/K_m$ = 8.5 × $10^{-4}$ $M^{-1}$ $s^{-1}$) to yield cinnamaldehyde (**6**) (Scheme 3.2).[4] In this promiscuous aldolase activity of 4-OT, Pro-1 likely acts as a nucleophile and forms an enamine intermediate with **3** (Scheme 3.3). A reaction between this nucleophilic intermediate and electrophile **4** results in carbon-carbon bond formation. The presumed scenario for the formation of **6** involves the initial cross-coupling of **3** and **4** to yield 3-hydroxy-3-phenylpropanal (**5**, Scheme 3.2), which then undergoes dehydration to yield **6**.[5] However, when the 4-OT-catalyzed reaction between **3** and **4** was monitored by $^1$H NMR spectroscopy, the signals corresponding to **6** were observed, but those corresponding to **5** were not.[4] The absence of the latter signals suggests either that **5** is not sufficiently stable to accumulate in quantities detectable by $^1$H NMR spectroscopy, or that 4-OT catalyzes the dehydration of **5** to yield **6**.
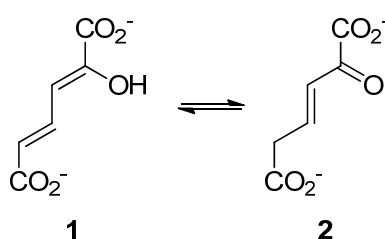


**Scheme 3.1.** The natural tautomerization reaction catalyzed by 4-OT.
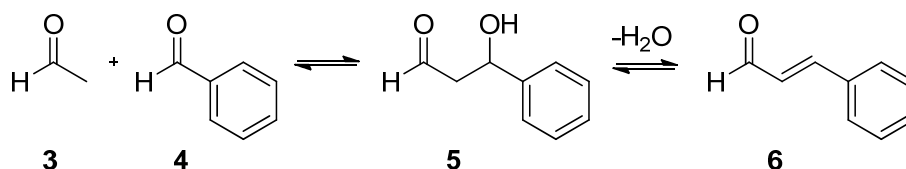


**Scheme 3.2.** The aldol reaction catalyzed by 4-OT to yield **5**, as well as subsequent dehydration of **5** to give **6**.

We have previously shown that a single amino acid substitution (L8R) in the active site of 4-OT results in a slightly enhanced (16-fold in terms of $k_{cat}/K_m$) aldolase activity ($k_{cat}/K_m$ = $1.4 \times 10^{-2}$ M$^{-1}$ s$^{-1}$).[4] Here, we describe the construction and characterization of a 4-OT variant (F50A) with strongly improved aldolase activity (600-fold in terms of $k_{cat}/K_m$). Furthermore, we report the chemical synthesis and use of compound **5** to demonstrate that both wild-type 4-OT and the F50A mutant accept **5** as a substrate, and catalyze its dehydration (to give **6**) and retro-aldol reaction (to give **3** and **4**) simultaneously (Scheme 3.4). In addition to the large body of literature on natural aldolases,[6] a few studies on promiscuous enzyme-catalyzed aldol reactions have been reported.[7] Whereas these biocatalysts use either an active site base or the primary amine of a lysine residue to generate reactive enolate anion or enamine intermediates, respectively, 4-OT is a unique example of an enzyme that uses the secondary amine of an active site proline as the nucleophile catalyst to give enamines with carbonyl substrates.
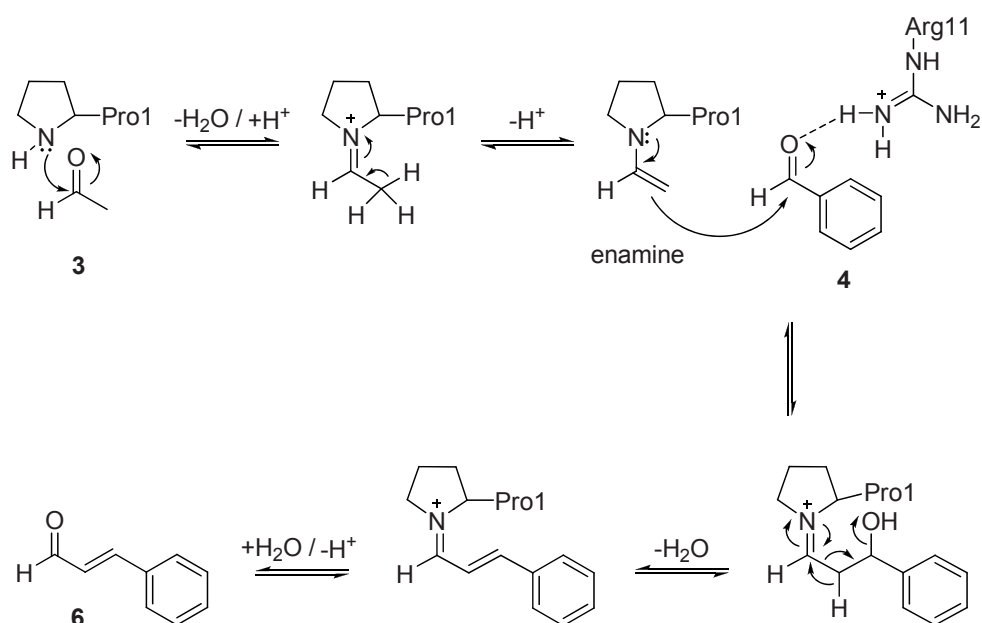


**Scheme 3.3.** Proposed mechanism for the 4-OT-catalyzed aldol condensation of **3** with **4** to yield **6**. The aldol dehydration step is visualized as being catalyzed by 4-OT. The other possibility of the initial aldol product being expelled from 4-OT's active site by hydrolysis prior to chemical dehydration (i.e. non-4-OT-catalyzed dehydration) is not shown.

In the proposed mechanism for 4-OT's aldolase activity (Scheme 3.3), the first step is the formation of an enamine between **3** and residue Pro-1. The highly reactive nucleophilic enamine intermediate subsequently attacks electrophile **4**. Inhibition and mass spectrometry

experiments have shown rapid covalent labeling of Pro-1 by **3**, indicating that enamine formation is relatively fast.[4] We therefore hypothesized that the final hydrolysis step, in which product **6** is released from Pro-1, might be rate-limiting. Interestingly, Whitman and co-workers have reported that the replacement of residue Phe-50 by an alanine makes the active site of 4-OT more accessible to the external aqueous environment, causing an increase in the local dielectric constant and in the p$K_a$ of Pro-1 (from 6.4 to 7.3).[8] The F50A mutation hence appears to make the active site more amenable for hydrolysis, without raising the p$K_a$ of Pro-1 too much, and might thus enhance the aldolase activity of 4-OT.
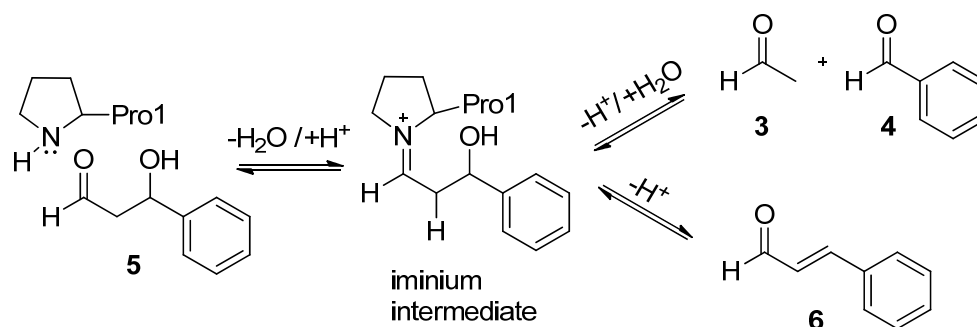


**Scheme 3.4.** The 4-OT-catalyzed conversion of **5** proceeds by two possible routes. Dehydration of the iminium intermediate gives **6**, whereas a retro-aldol reaction gives **3** and **4**.

## Results and Discussion

The 4-OT F50A mutant was constructed, purified to homogeneity, and tested for its ability to catalyze the aldol condensation reaction. Incubation of this mutant enzyme (149 μM) with **3** (50 mM) and **4** (2 mM) in $NaH_2PO_4$ buffer (20 mM, pH 7.3) resulted in a relatively fast increase in absorbance at 290 nm ($A_{290}$), which corresponds to the formation of **6** (Figure 3.1). When the rate of this reaction is compared to the rate of the same reaction catalyzed by wild-type 4-OT or the 4-OT L8R mutant, it is clear that the 4-OT F50A mutant has a strongly improved aldolase activity (Figure 3.1). To verify that **6** is the product of an aldol condensation catalyzed by 4-OT F50A, the reaction was also monitored by ^{1}H NMR spectroscopy. Indeed, upon incubation of **3** (20 mM) and **4** (20 mM) with 4-OT F50A (0.5 mg, 114 μM, 0.6 mol%) in $NaH_2PO_4$ buffer (20 mM, pH 7.3), characteristic ^{1}H NMR signals (at 6.83 and 9.56 ppm) corresponding to product **6** were detected (Figure 3.2). Integration of the signals indicated ~16% conversion of **3** and **4** into **6** after 24 h. It is important to note that no ^{1}H NMR signals corresponding to aldol product **5** were observed in the course of the reaction.
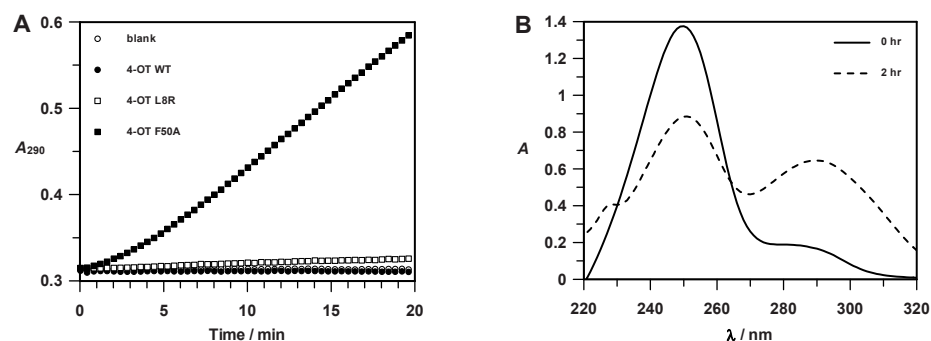
**Figure 3.1.** A) UV traces monitoring the formation of cinnamaldehyde (**6**) from acetaldehyde (**3**) and benzaldehyde (**4**) in the presence of 4-OT WT, 4-OT L8R, 4-OT F50A, or without 4-OT (blank). B) UV spectra monitoring the aldol condensation of acetaldehyde (**3**) with benzaldehyde (**4**) ($\lambda_{max}$ = 250 nm) to yield cinnamaldehyde (**6**) ($\lambda_{max}$ = 290 nm) catalyzed by 4-OT F50A.
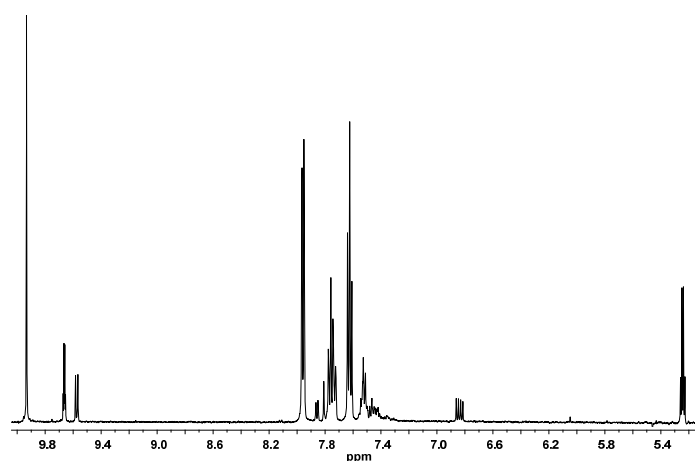


**Figure 3.2.** Partial $^1$H NMR spectrum monitoring the aldol condensation of **3** with **4** to yield **6** catalyzed by 4-OT F50A (t = 1 day). Signals corresponding to substrates **3** and **4** and product **6** are observed.

After we established that 4-OT F50A catalyzes the aldol condensation of **3** with **4**, apparent kinetic parameters were determined by using a fixed concentration of **3** (50 mM) and varying concentrations of **4** (0.5-15 mM). 4-OT F50A catalyzes the aldol condensation with a $k_{cat}$ of 2.2 × 10$^{-3}$ s$^{-1}$ and a $K_m$ of 4.3 mM, yielding a $k_{cat}/K_m$ of ~0.5 M$^{-1}$ s$^{-1}$. With 36- and 600-fold increases in $k_{cat}/K_m$ relative to the reactions catalyzed by 4-OT L8R and 4-OT WT, respectively, 4-OT F50A has a greatly enhanced aldolase activity. The construction, purification, and analysis of the double mutant 4-OT P1A F50A, which displays no aldol condensation activity (Figure 3.3), showed that Pro-1 is essential for this activity of 4-OT F50A.
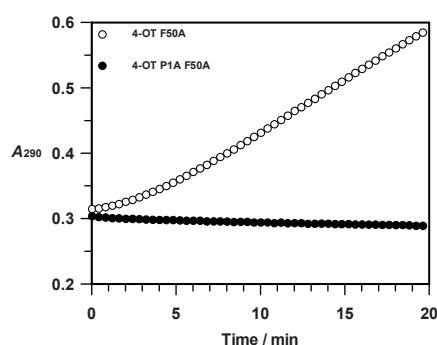
**Figure 3.3.** UV traces monitoring the formation of cinnamaldehyde (**6**) from acetaldehyde (**3**) and benzaldehyde (**4**) in the presence of 4-OT F50A or 4-OT P1A F50A.

The fact that the initial product **5** was not observed by spectroscopic analysis of the aldol condensation catalyzed by 4-OT F50A implies that the dehydration of **5** to give **6** (Scheme 3.2) is significantly faster than the initial addition of **3** to **4**. The question whether this conversion is 4-OT-catalyzed or proceeds via a chemical process remained. In order to address this question, compound **5** was chemically synthesized [9] and tested for its stability in NaD$_2$PO$_4$ buffer (20 mM, pD 7.8, [**5**] = 8.8 mM). $^1$H NMR analysis indicated 10% conversion after 2 days and ~50% conversion after 9 days (Figure 3.4). Compound **5** mainly undergoes dehydration to yield **6** (>98%), but small amounts of **3** and **4**, resulting from the retro-aldol reaction of **5** (<2%), were observed as well.
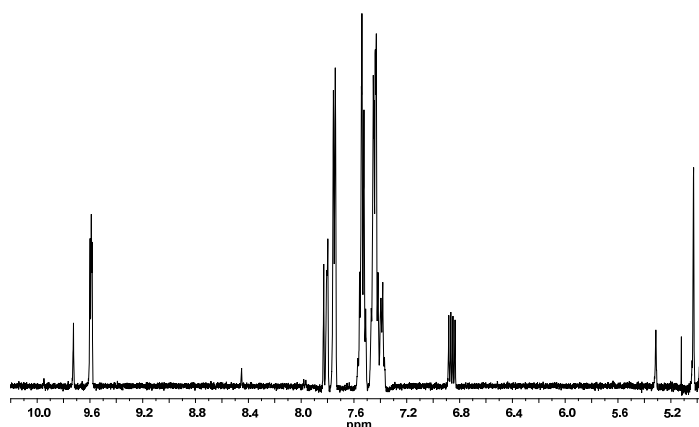


**Figure 3.4.** Partial $^1$H NMR spectrum monitoring the stability of **5** in NaD$_2$PO$_4$ buffer (20 mM, pD 7.8) (t = 9 days). Signals corresponding to substrate **5** and products **4** and **6** are observed.
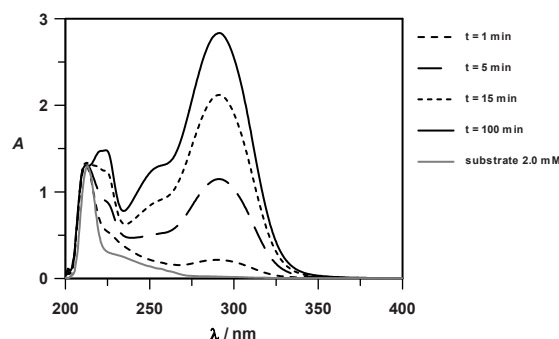
**Figure 3.5.** UV spectra monitoring the conversion of **5** catalyzed by 4-OT F50A. The increase in $A_{290}$ (formation of **6**) indicates the dehydration of **5**. The increase in $A_{250}$ (formation of **4**) indicates the retro-aldol reaction of **5**.

We next examined the enzymatic conversion of **5**. In separate reactions, 4-OT WT, 4-OT L8R, and 4-OT F50A (150 µM each) were incubated with **5** (2 mM) in $NaH_2PO_4$ buffer (20 mM, pH 7.3). In all cases, UV spectroscopic analysis (200-400 nm) revealed increases in absorbance at 250 nm and 290 nm (Figure 3.5). The increase in the $A_{290}$ value corresponds to the formation of **6**, which results from the dehydration of **5**. The increase in the $A_{250}$ value corresponds to the formation of **4**, which results from the retro-aldol reaction of **5**. Interestingly, the product distribution of **6** and **4** differs for the three enzyme-catalyzed reactions. Molar ratios of **6**/**4** of 74:26, 88:12, and 61:39 were found for 4-OT WT, 4-OT L8R, and 4-OT F50A, respectively (Table 3.1). Although the absorbance spectra of compounds **4**, **5**, and **6** partially overlap, ratios were calculated on basis of the UV measurements and individually determined ε values of **4**, **5**, and **6** (Experimental Procedures). Control experiments in which 4-OT (WT, L8R, or F50A) was incubated with **6** showed no formation of **4** nor of **5**, which suggests that enzymatic hydration of **6** does not occur under the assay conditions used.

**Table 3.1.** Comparison of the dehydration and retro-aldol reactions of 3-hydroxy-3-phenylpropanal (**5**) in the absence of enzyme (blank) or catalyzed by 4-OT WT, 4-OT L8R, 4-OT F50A, or 4-OT L8R F50A.

| Enzyme | $t_{reaction}$ (h) | Conversion (%) | Dehydration (%) | Retro-aldol (%) |
|---|---|---|---|---|
| blank | 40 | 9 | >98 | <2 |
| 4-OT WT | 40 | 54 | 74 | 26 |
| 4-OT L8R | 4 | 83 | 88 | 12 |
| 4-OT F50A | 2.3 | 93 | 61 | 39 |
| 4-OT L8R F50A | 8 | 93 | 82 | 18 |

Conditions: compound **5** (2 mM), enzyme (150 µM), $NaH_2PO_4$ (20 mM), pH 7.3.
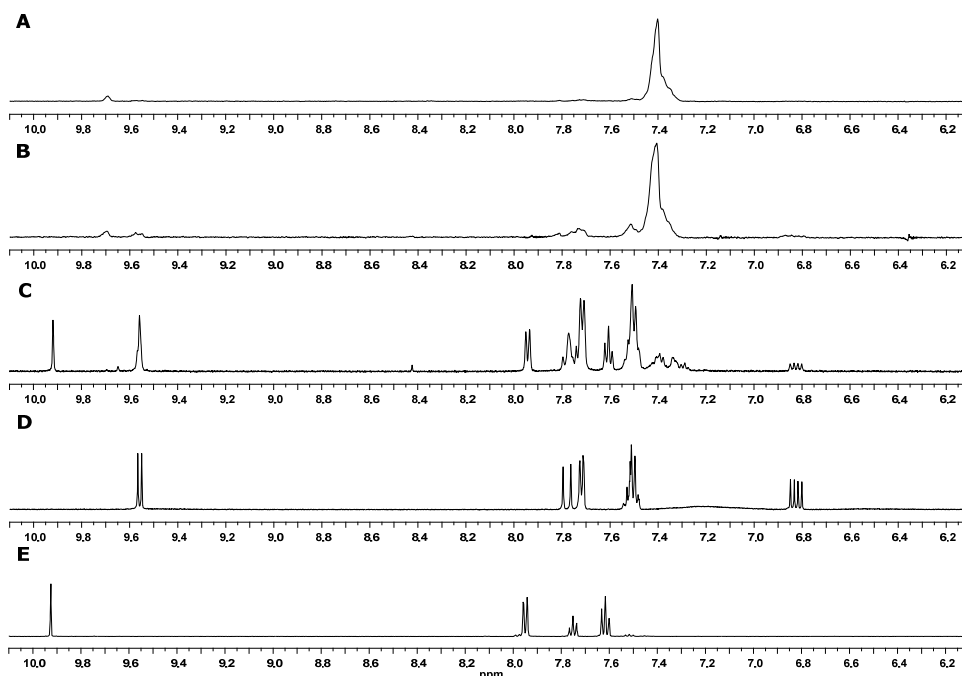
**Figure 3.6.** Stack plot of partial $^1H$ NMR spectra of A) 3-hydroxy-3-phenylpropanal (**5**) at t = 0, B) **5** at t = 2 days, C) **5** incubated with 4-OT F50A (t = 2 days), D) cinnamaldehyde (**6**), and E) benzaldehyde (**4**). During the reaction, approximately 70% of the C-2 hydrogens of **5** exchanged with deuterium from the reaction medium, so the integration value of the signals of the C-2 hydrogen of final product **6** (6.83 ppm) is reduced to 30% relative to all other integration values of signals of **6**. As a result, the signals of the aldehyde proton (9.56 ppm) and the C-3 hydrogen (7.78 ppm) of **6** mainly appear as singlets instead of doublets (spectrum C).

To verify that **4** and **6** are the products of the conversion of **5** catalyzed by 4-OT F50A, the reaction was also monitored by $^1H$ NMR spectroscopy. Indeed, upon incubation of **5** (8.8 mM) with 4-OT F50A (0.25 mg, 57 μM, 0.6 mol%) in $NaD_2PO_4$ buffer (20 mM, pD 7.8), characteristic $^1H$ NMR signals corresponding to **4** and **6** were detected (Figure 3.6).[10] After 2 days, a conversion of 95% had been reached and a ratio of **6/4** of 68:32 was observed. This ratio is similar to what had been found by UV spectroscopy, thus validating our calculation method for product ratios based on UV absorbance values.[11]

To quantify the dehydration and retro-aldol activities of 4-OT WT, 4-OT L8R and 4-OT F50A, the three enzymes (150 μM) were incubated separately with **5** (2 mM; in 20 mM $NaH_2PO_4$ buffer, pH 7.3) and the formation of **6** ($A_{320}$, dehydration) and **4** ($A_{249}$, retro-aldol) were followed over time. Initial rates for the dehydration and retro-aldol reactions were calculated for all three enzymes by the above-mentioned calculation method, which is based on UV measurements (Table 3.2). The initial dehydration and retro-aldol reaction rates

display the same trend as the rates for the enzyme-catalyzed aldol condensation reactions, with increasing activity from 4-OT WT to 4-OT L8R to 4-OT F50A. The importance of Pro-1 for the dehydration and retro-aldol activities was demonstrated by the analysis of the double mutant 4-OT P1A F50A, which showed greatly reduced activity in relation to 4-OT F50A (Table 3.2).

**Table 3.2.** Initial rates of enzymatic dehydration or retro-aldol reaction with **5** as substrate.

| Enzyme | Initial rate dehydration (mM/s)[a] | Relative | Initial rate retro-aldol (mM/s)[b] | Relative |
|---|---|---|---|---|
| blank | $2.5 \times 10^{-3}$ | 0.13 | 0 | 0 |
| 4-OT WT | $1.8 \times 10^{-2}$ | 1 | $6.5 \times 10^{-3}$ | 1 |
| 4-OT L8R | 0.60 | 33 | $3.8 \times 10^{-2}$ | 6 |
| 4-OT F50A | 1.8 | 100 | 0.58 | 89 |
| 4-OT L8R F50A | 0.28 | 16 | $3.5 \times 10^{-2}$ | 5 |
| 4-OT P1A F50A | 0.14 | 8 | $1.2 \times 10^{-2}$ | 2 |

[a] Based on formation of cinnamaldehyde (**6**).
[b] Based on the formation of benzaldehyde (**4**).
Conditions: compound **5** (2 mM), enzyme (150 µM), $NaH_2PO_4$ (20 mM), pH 7.3.

Apparent kinetic parameters were estimated for the dehydration and retro-aldol reactions catalyzed by 4-OT F50A (Table 3.3). 4-OT F50A (18.9 µM) was incubated with varying concentrations of **5** (0.25-20 mM), and the initial rates of dehydration into **6** and retro-aldol reaction into **3** and **4** were determined simultaneously by measuring UV-spectra from 200-400 nm over time. At higher substrate concentrations, substrate inhibition (apparent $K_i$ = 10.8 mM) was observed for both the dehydration and the retro-aldol reaction (Figure 3.7). The $k_{cat, dehydration}/k_{cat, retro-aldol}$ ratio ($2.0 \times 10^{-2}$:$6.7 \times 10^{-3}$ is ~ 3; Table 3.3) is the same as the ratio of the determined initial rates (1.8:0.58 is ~ 3; Table 3.2).

**Table 3.3.** Apparent kinetic parameters for the dehydration and retro-aldol reaction of **5** catalyzed by 4-OT F50A.

| Reaction | Product | $k_{cat}$ (s$^{-1}$) | $K_m$ (mM) | $K_i$ (mM) | $k_{cat}/K_m$ (M$^{-1}$ s$^{-1}$) |
|---|---|---|---|---|---|
| Dehydration | **6** | $2.0 \times 10^{-2}$ | 2.8 | 10.8 | 7.2 |
| Retro-aldol | **3** and **4** | $6.7 \times 10^{-3}$ | 1.7 | 10.8 | 3.9 |
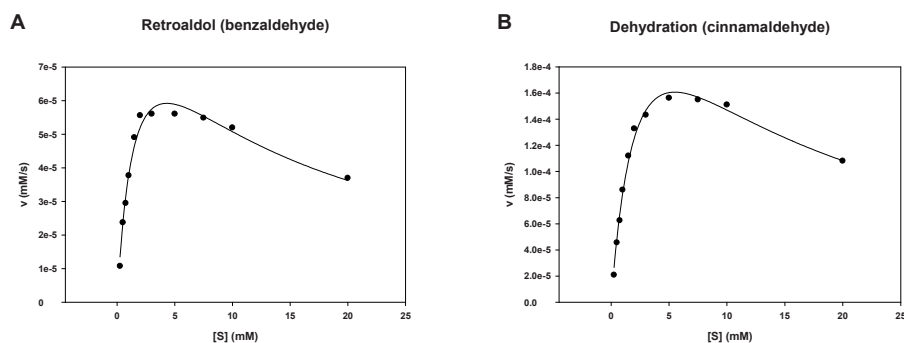
**Figure 3.7.** Michaelis-Menten plots of the retro-aldol (A) and dehydration (B) reactions of **5** catalyzed by 4-OT F50A. 4-OT F50A shows substrate inhibition kinetics for both activities.

Unfortunately, combination of the L8R and F50A mutations into one enzyme (L8R F50A) did not have an additive or synergistic effect either on the aldol condensation activity or on the dehydration or retro-aldol activities. The activity levels of double mutant 4-OT L8R F50A were greatly decreased in relation to those of 4-OT F50A and were comparable to the activities of 4-OT L8R (Table 3.1, Table 3.2 and Figure 3.8).



**Figure 3.8.** UV traces monitoring the formation of cinnamaldehyde (**6**) from acetaldehyde (**3**) and benzaldehyde (**4**) in the presence of 4-OT F50A, 4-OT L8R, or 4-OT L8R F50A.

## Conclusion

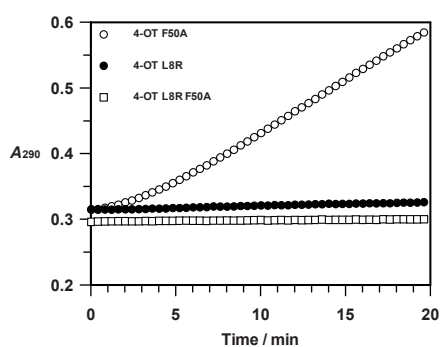In summary, we demonstrated that a single amino acid substitution (F50A) gives a large increase (600-fold in terms of $k_{cat}/K_m$) in the promiscuous aldolase activity of 4-OT. Furthermore, we provide evidence that the dehydration of aldol product **5** into **6** and the retro-aldol reaction of **5** into **3** and **4** are both catalyzed by 4-OT wild-type and 4-OT mutants, of which 4-OT F50A is most effective. With regard to our efforts to screen 4-OT systematically for promiscuous carbonyl transformation activities, we have currently established that 4-OT can catalyze three promiscuous carbonyl transformation reactions: the reversible cross-coupling of aldehydes **3** and **4** to yield the aldol product **5** (this work),[4] the dehydration of **5** to yield **6** (this work), and the Michael-type addition of **3** to β-nitrostyrenes.[12] The mechanistic and kinetic insights gained in this study will serve as a stepping stone to further explore the promiscuous reaction scope of 4-OT and to design additional 4-OT mutants that lack dehydration activity but possess further enhanced carbon-carbon bond-forming activities.

# Experimental Procedures

*Materials:* All chemicals were obtained from Sigma-Aldrich Chemical Co. (St. Louis, MO), unless stated otherwise. The sources for the biochemicals, buffers, solvents, components of Luria-Bertani (LB) media as well as the materials, enzymes, and reagents used in the molecular biology procedures are reported elsewhere.[13]

*General methods:* Techniques for restriction enzyme digestions, ligation, transformation, and other standard molecular biology manipulations were based on methods described elsewhere.[14] The PCR was carried out in a DNA thermal cycler (model GS-1) obtained from Biolegio (Nijmegen, The Netherlands). DNA sequencing was performed by Macrogen (Seoul, Korea). Protein was analyzed by polyacrylamide gel electrophoresis (PAGE) using sodium dodecyl sulfate (SDS) gels containing polyacrylamide (10%). The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the Waddell method.[15] Kinetic data were obtained on a V-650 or V-660 spectrophotometer from Jasco (IJsselstein, The Netherlands) and were fitted by nonlinear regression data analysis using the GraFit program (Erithacus Software Ltd., Horley, U.K.) obtained from Sigma Chemical Co or SigmaPlot 11.0 (Systat Software Inc., Chicago, U.S.A.). $^1$H NMR spectra were recorded on a Varian Inova 300 (300 MHz) or 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 ($\delta$ scale) and are referenced to $H_2O$ ($\delta$ = 4.80). The subunit masses of 4-OT and 4-OT mutants were determined by ESI-MS using a Sciex API 3000 triple quadrupole mass spectrometer (AB Sciex, Concord, Ontario, Canada) which is housed in the Mass Spectrometry Facility Core in the Department of Pharmacy at the University of Groningen.

*Construction of the 4-OT mutants F50A, P1A F50A and L8R F50A:* For construction of the three 4-OT mutants by PCR, the coding sequence for 4-OT in plasmid pET20b(4-OT)[4] was used as the template. To construct the 4-OT F50A mutant, the following oligonucleotide was used as forward primer: 5′-A TAG CAG GTA <u>CAT ATG</u> CCT ATT GCC CAG ATC CAC ATC CCT-3′. To introduce the P1A and L8R mutations, the following oligonucleotides were used as forward primers, respectively: 5′-A TAG CAG GTA <u>CAT ATG</u> **GCC** ATT GCC CAG ATC CAC ATC CTT G-3′ and 5′- A TAG CAG GTA <u>CAT ATG</u> CCT ATT GCC CAG ATC CAC ATC **CGC** GAA GGC-3′ (with the *Nde*I site underlined and the mutated codon in bold). These primers correspond to the 5′-end of the wild-type coding sequence and were used in combination with oligonucleotide 5′-A TGT TAT <u>GGA TCC</u> TCA GCG TCT GAC CTT GCT GGC CAG TTC GCC GCC GAT GCC **GGC** GTG GCC CTT-3′ (with the *Bam*HI site underlined and the mutated codon in bold) as the reverse primer to introduce the F50A mutation. The resulting PCR products were gel-purified, digested and cloned in frame with the ATG start codon of the pET20b(+) vector. All genes were fully sequenced to assure that only the desired mutations were introduced.

*Expression and purification of wild-type 4-OT and mutants:* 4-OT and 4-OT mutants were produced in *E. coli* BL21(DE3) cells harboring the appropriate expression vector (based on the pET20b(+) plasmid). The construction of the expression vectors for 4-OT and the mutants 4-OT P1A and 4-OT L8R, and the expression and purification of 4-OT and 4-OT mutants, including 4-OT F50A, 4-OT P1A F50A and 4-OT L8R F50A, were performed according to previously reported protocols.[14] Mutant proteins 4-OT F50A, 4-OT P1A F50A and 4-OT L8R F50A had minor interaction with the phenyl-sepharose column and eluted in the first and second washing steps as homogeneous proteins (>95% purity as assessed by SDS-PAGE). To confirm that the proteins had been processed correctly and the initiating methionine had been removed, the masses of the purified mutant proteins were determined by ESI-MS. The observed monomer masses for the mutants are as follows: F50A, 6735 Da (calc. 6735 Da); P1A F50A, 6710 Da (calc. 6709 Da); L8R F50A, 6779 Da (calc. 6778 Da).

*General UV spectroscopic assay for the aldol condensation of acetaldehyde (**3**) with benzaldehyde (**4**):* The aldolase activity of wild-type 4-OT and 4-OT mutants was monitored by following the increase in absorbance at 290 nm, which corresponds to the formation of cinnamaldehyde (**6**) ($\varepsilon = 26.7$ mM$^{-1}$ cm$^{-1}$). Enzyme (300 µg, 149 µM) was incubated in a 1 mm cuvette with **3** (50 mM) and **4** (2 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3; 0.3 mL final volume). Absorbance spectra were recorded from 200 to 400 nm.

*Kinetic assay for aldolase activity:* The kinetic assays were performed at 22 °C by following the increase in absorbance at 290 nm, which corresponds to the formation of **6**. An aliquot of 4-OT F50A was added to 0.3 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) in a 1 mm cuvette, yielding a final enzyme concentration of 149 µM. The enzyme activity was assayed by the addition of **3** (at a fixed concentration of 50 mM) and **4** (in concentrations varying from 0.5 to 15 mM). The initial rates (mM/s) were plotted versus the concentration of **4** (mM). GraFit was used to fit the data to Michaelis-Menten kinetics and to calculate the apparent kinetic parameters $k_{cat}$ and $K_m$ (for **4**).

*$^1$H NMR spectroscopic assay for aldolase activity:* $^1$H NMR spectra monitoring the aldol condensation of **3** with **4** catalyzed by 4-OT F50A were recorded as follows. The enzyme (0.5 mg; 114 µM; 0.6 mol%) was incubated with **3** and **4** (20 mM each) in 0.6 mL of 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) at 22 °C. In addition, 0.05 mL of D$_2$O, containing 0.65 µL DMSO as internal standard, was added. A control experiment without enzyme, but with otherwise identical conditions, was performed as well. The first $^1$H NMR spectrum was recorded immediately after mixing, and then after 1 and 7 days. The $^1$H NMR signals for **3** (and its hydrate), **4**, and **6** are reported below.

*Compound **3** and its hydrate:* $^1$H NMR (500 MHz, D$_2$O, 25 °C): δ = 9.65 (q, $J$ = 3.0 Hz, 1H, **3**), 5.22 (q, $J$ = 5.0 Hz, 1H, hydrate), 2.21 (d, $J$ = 3.0 Hz, 3H, **3**), 1.30 (d, $J$ = 5.0 Hz, 3H, hydrate).

*Compound 4:* ^{1}H NMR (500 MHz, $D_2O$, 25 °C): δ = 9.93 (s, 1H), 7.95 (d, *J* = 8.0 Hz, 2H), 7.76 (t, *J* = 7.5 Hz, 1H), 7.62 (t, *J* = 7.5 Hz, 2H).

*Compound 6:* ^{1}H NMR (500 MHz, $D_2O$, 25 °C): δ = 9.56 (d, *J* = 8.0 Hz, 1H), 7.78 (d, *J* = 15.5 Hz, 1H), 7.72 (d, *J* = 8.0 Hz, 2H), 7.51 (m, 3H), 6.83 (dd, *J* = 8.0, 8.0 Hz, 1H).

*Chemical synthesis of 3-hydroxy-3-phenylpropanal (5):* Compound **5** was prepared according to a modified literature procedure.[9] 4-Phenylbutane-1,2,4-triol[9] (25 mg, 0.14 mmol) was dissolved in $D_2O$ (3 mL). $NaIO_4$ (29.4 mg, 0.14 mmol) was added and the solution was stirred at room temperature for 1 h. 4-Phenylbutane-1,2,4-triol was completely converted into **5** and its hydrate, 3-phenylpropane-1,1,3-triol, in a 1:3 ratio as assessed by ^{1}H NMR spectroscopy. No further purification steps were performed since **5** is prone to dehydration in non-aqueous environments such as organic solvents. The in situ generated mixture of **5** and its hydrate was directly used in kinetic and ^{1}H NMR assays. ^{1}H NMR (300 MHz, $D_2O$, 25 °C): 3-hydroxy-3-phenylpropanal (**5**); δ = 9.72 (dd, *J* = 1.9, 1.9 Hz, 1H), 7.49 - 7.34 (m, 5H), 5.32 (dd, *J* = 7.8, 5.4 Hz, 1H), 3.04 (ddd, *J* = 17.1, 7.8, 1.9 Hz, 1H), 2.96 (ddd, *J* = 17.1, 5.4, 1.9 Hz, 1H); 3-phenylpropane-1,1,3-triol; δ = 7.49 - 7.34 (m, 5H), 5.03 (dd, *J* = 6.0, 5.4 Hz, 1H), 4.85 (dd, *J* = 8.4, 5.8 Hz, 1H), 2.15 (ddd, *J* = 13.9, 8.4, 5.4 Hz, 1H), 2.00 (ddd, *J* = 13.9, 6.0, 5.8 Hz, 1H).

*General UV spectroscopic assay for dehydration and retro-aldol activities of 4-OT and 4-OT mutants:* The dehydration and retro-aldol activities of wild-type 4-OT and 4-OT mutants were monitored by following the formation of **6** and **4** by UV spectroscopy. Enzyme (0.3 mg; 150 μM) was incubated with **5** (2 mM; 13.1 μL of the above-described reaction mixture to prepare **5**) in 20 mM $NaH_2PO_4$ buffer (pH 7.3; 0.3 mL final volume) in a 1 mm cuvette. Absorbance spectra were recorded from 200 to 400 nm. At different reaction times, the $\Delta A_{249}$ and $\Delta A_{320}$ values were determined and used to calculate the ratio of product **4** (retro-aldol) versus product **6** (dehydration). The calculation method is described below in the paragraph *Kinetic assay for dehydration and retro-aldol activities*.

*Preparation of $NaD_2PO_4$ buffer:* $NaD_2PO_4$ buffer was prepared by lyophilizing 20 mM $NaH_2PO_4$ buffer (pH 7.3, 20 mL). The white residue was dissolved in $D_2O$ (2 mL) and stirred for 60 min. After lyophilizing (2nd), dissolving the residue in $D_2O$ (2 mL), stirring for 60 min and lyophilizing (3rd), the residue was dissolved in 10 mL of $D_2O$ giving 40 mM $NaD_2PO_4$ (pD 7.8). Half of this solution (5 mL) was diluted with another portion of 5 mL of $D_2O$ to give 10 mL of 20 mM $NaD_2PO_4$ (pD 7.8).

*Dissolving 4-OT F50A in $NaD_2PO_4$ buffer:* For ^{1}H NMR spectroscopic experiments, 4-OT F50A was dissolved in $NaD_2PO_4$ buffer (20 mM, pD 7.8). First, a Vivaspin concentrator (2 mL, cut-off 5 kDa, Sartorius Stedim) was washed four times with $D_2O$ by centrifugation. Subsequently, a solution of 4-OT F50A (300 μL, 12.5 mg/mL in 20 mM $NaH_2PO_4$ buffer, pH 7.3) was loaded and centrifuged at 4,000 rpm

for 30 min. The concentrated enzyme was diluted with 20 mM NaD$_2$PO$_4$ (200 µL, pD 7.8) and concentrated again (4,000 rpm, 30 min). The concentrated enzyme was diluted in 20 mM NaD$_2$PO$_4$ (300 µL, pD 7.8) and the final concentration was determined (10.8 mg/mL).

*$^1$H NMR spectroscopic assay for dehydration and retro-aldol activities:* $^1$H NMR spectra monitoring the conversion of **5** by 4-OT F50A were recorded as follows. The enzyme (23 µL of a 10.8 mg/mL stock solution in 20 mM NaD$_2$PO$_4$ buffer (pD 7.8), 0.25 mg; 57 µM, 0.6 mol%) was incubated with **5** (8.8 mM, 0.12 mL of the reaction mixture described above for the preparation of **5**) in 0.12 mL of 40 mM NaD$_2$PO$_4$ buffer (pD 7.8) and 0.36 mL 20 mM NaD$_2$PO$_4$ buffer (pD 7.8) at 22 °C. The total volume was 623 µL. A control experiment without enzyme but with otherwise identical conditions was performed as well. The first $^1$H NMR spectrum was recorded immediately after mixing, and then after 2 and 9 days. The $^1$H NMR signals for retro-aldol products acetaldehyde (**3**) (and its hydrate) and benzaldehyde (**4**), and the dehydration product cinnamaldehyde (**6**), were in accordance with authentic standards.

*Kinetic assay for dehydration and retro-aldol activities:* The kinetic assays were performed at 22 °C by following the increase in absorbance at 249 and 320 nm, which corresponds to the formation of **4** and **6**, respectively. A stock solution of 4-OT F50A (16.7 µM, 0.112 mg/ml) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) was prepared. Stock solutions of **5** of various concentrations were prepared by dilution of the reaction mixture to prepare **5** (*vide supra*, total concentration of **5** and its hydrate is 45.5 mM) in D$_2$O. To assay the enzymatic activity, an aliquot (267 µL) of the enzyme solution was incubated in a 1 mm cuvette and **5** was added (to a final volume of 0.3 mL), yielding final concentrations of 18.9 µM 4-OT F50A and 0.25-20 mM **5**. Absorbance spectra were recorded in duplo from 200-400 nm for ten minutes per reaction (one spectrum each minute). The A$_{249}$ and A$_{320}$ values were read out and plotted versus time. The slopes $\Delta$A$_{249}$ and $\Delta$A$_{320}$ were determined and used in the following calculation:

$\Delta$A$_{320}$ is solely caused by the formation of **6** (cinnamaldehyde = CA) since **4** and **5** do not absorb at 320 nm, therefore:

$$\Delta[CA] = \Delta A_{320}/\varepsilon_{CA,320} \tag{1}$$

$\Delta$A$_{249}$ is caused by substrate **5** (S) and products **4** (benzaldehyde = BZ) and **6**:

$$\Delta A_{249} = \Delta A_{CA,249} + \Delta A_{BZ,249} + \Delta A_{S,249} = \varepsilon_{CA,249}*\Delta[CA] + \varepsilon_{BZ,249}*\Delta[BZ] + \varepsilon_{S,249}*\Delta[S] \tag{2}$$

$$\Delta[BZ] = (\Delta A_{249} - \varepsilon_{CA,249}*\Delta[CA] - \varepsilon_{S,249}*\Delta[S])/\varepsilon_{BZ,249} \qquad \text{rewrite [2]} \tag{3}$$

$$\Delta[S] = - \Delta[CA] - \Delta[BZ] \hspace{5cm} [4]$$

$$\Delta[BZ] = (\Delta A_{249} - \varepsilon_{CA,249}*\Delta[CA] + \varepsilon_{S,249}*\Delta[CA] + \varepsilon_{S,249}*\Delta[BZ])/\varepsilon_{BZ,249} \hspace{1cm} \text{insert [4] in [3]} \hspace{0.5cm} [5]$$

$$\Delta[BZ]*\varepsilon_{BZ,249} = \Delta A_{249} - \varepsilon_{CA,249}*\Delta[CA] + \varepsilon_{S,249}*\Delta[CA] + \varepsilon_{S,249}*\Delta[BZ] \hspace{1cm} \text{rewrite [5]} \hspace{0.5cm} [6]$$

$$\Delta[BZ]*\varepsilon_{BZ,249} - \varepsilon_{S,249}*\Delta[BZ] = \Delta A_{249} - \varepsilon_{CA,249}*\Delta[CA] + \varepsilon_{S,249}*\Delta[CA] \hspace{1cm} \text{rewrite [6]} \hspace{0.5cm} [7]$$

$$\Delta[BZ] = (\Delta A_{249} - \varepsilon_{CA,249}*\Delta[CA] + \varepsilon_{S,249}*\Delta[CA])/(\varepsilon_{BZ,249} - \varepsilon_{S,249}) \hspace{1cm} \text{rewrite [7]} \hspace{0.5cm} [8]$$

$$\Delta[BZ] = (\Delta A_{249} - (\varepsilon_{CA,249} - \varepsilon_{S,249})*\Delta[CA])/(\varepsilon_{BZ,249} - \varepsilon_{S,249}) \hspace{1cm} \text{rewrite [8]} \hspace{0.5cm} [9]$$

$\Delta[CA]$ was calculated with formula [1] and used in formula [9] to determine $\Delta[BZ]$. The change in concentration was divided by the time to give rates in mM/s. The rates were plotted against the substrate concentration (of **5**) and Sigmaplot was used to fit the data to substrate inhibition kinetics and to calculate the apparent kinetic parameters $k_{cat}$, $K_m$, and $K_i$.

The ε values of the three components were determined in 20 mM $NaH_2PO_4$ buffer (pH = 7.3):

$\varepsilon_{CA,320} = 0.64$ mM$^{-1}$ mm$^{-1}$

$\varepsilon_{CA,249} = 0.26$ mM$^{-1}$ mm$^{-1}$

$\varepsilon_{BZ,249} = 1.25$ mM$^{-1}$ mm$^{-1}$

$\varepsilon_{S,249} = 0.07$ mM$^{-1}$ mm$^{-1}$

*Determination of initial rates of dehydration and retro-aldol reaction:* In separate reactions, 4-OT WT, 4-OT L8R, 4-OT F50A, 4-OT L8R F50A, and 4-OT P1A F50A (each 0.3 mg; 150 μM) were mixed with **5** (2 mM = 13.1 μL of reaction mixture to prepare **5**, *vide supra*) in 20 mM $NaH_2PO_4$ buffer (pH 7.3). The final volume was 300 μL. The dehydration reaction was monitored by the formation of **6** (time course UV measurement at 320 nm), while the retro-aldol reaction was monitored by the formation of **4** (time course UV measurement at 249 nm). The slopes of the UV traces were determined ($\Delta A_{320}$ and $\Delta A_{249}$) and the initial rates of formation of **6** and **4** were calculated using formulas [1] and [9].

## Acknowledgements

# References

1.  a) S. Harayama, M. Rekik, K. L. Ngai, L. N. Ornston, *J. Bacteriol.* **1989**, *171*, 6251-6258; b) C. P. Whitman, B. A. Aird, W. R. Gillespie, N. J. Stolowich, *J. Am. Chem. Soc.* **1991**, *113*, 3154-3162; c) L. H., Chen, G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour, C. P. Whitman, *J. Biol. Chem.* **1992**, *267*, 17716-17721.

2.  C. P. Whitman, *Arch. Biochem. Biophys.* **2002**, *402*, 1-13.

3.  E. A. Burks, C. D. Fleming, A. D. Mesecar, C. P. Whitman, S. D. Pegan, *Biochemistry* **2010**, *49*, 5016-5027 and references therein.

4.  E. Zandvoort, B. J. Baas, W. J. Quax, G. J. Poelarends, *ChemBioChem* **2011**, *12*, 602-609.

5.  P. Yurkanis Bruice, *Essential Organic Chemistry*, 2nd ed., **2010**, Pearson, Upper Saddle River,NJ, Chapter 6.

6.  a) M. Brovetto, D. Gamenara, P. Saenz Méndez, G. A. Seoane, *Chem. Rev.* **2011**, *111*, 4346-4403; b) P. Clapés, W.-D. Fessner, G. A. Sprenger, A. K. Samland, *Curr. Opin. Chem. Biol.* **2010**, *14*, 154-167; c) T. D. Machajewski, C.-H. Wong, *Angew. Chem.* **2000**, *112*, 1406-1430; *Angew. Chem. Int. Ed.* **2000**, *39*, 1352-1374, and references therein; d) P. Valentin-Hansen, F. Boetius, K. Hammer-Jespersen, I. Svendsen, *Eur. J. Biochem.* **1982**, *125*, 561-566; e) A. Heine, J. G. Luz, C.-H. Wong, I. A. Wilson, *J. Mol. Biol.* **2004**, *343*, 1019-1034.

7.  a) C. Branneby, P. Carlqvist, A. Magnusson, K. Hult, T. Brinck, P. Berglund, *J. Am. Chem. Soc.* **2003**, *125*, 874-875; b) C. Li, Y.-J. Zhou, N. Wang, X.-W. Feng, K. Li, X.-Q. Yu, *J. Biotechnol.* **2010**, *150*, 539-545.

8.  R. M. Czerwinski, T. K. Harris, M. A. Massiah, A. S. Mildvan, C. P. Whitman, *Biochemistry* **2001**, *40*, 1984-1995.

9.  See *Experimental Procedures* for a modified synthesis procedure from: S. Couty, C. Meyer, J. Cossy, *Tetrahedron*, **2009**, *65*, 1809–1832.

10. Only a trace amount of acetaldehyde (**3**) was detected due to its evaporation from the reaction mixture.

11. The product ratio in the UV-experiment (**6/4** 61:39 after 2.3 h) is slightly lower than the product ratio found in the NMR-experiment (**6/4** 68:32 after 40 h). This is due to the longer reaction time for the NMR-experiment as compared to the UV-experiment. The non-4-OT-catalyzed background dehydration reaction therefore contributes more significantly to formation of **6** during the NMR-experiment than during the UV-experiment. There is hardly any non-4-OT-catalyzed background retro-aldol reaction (see Table 3.1).

12. E. Zandvoort, E. M. Geertsema, B. J. Baas, W. J. Quax, G. J. Poelarends, *Angew. Chem.* **2012**, *124*, 1266-1269; *Angew. Chem. Int. Ed.* **2012**, *51*, 1240-1243.

13. H. Raj, B. Weiner, V. Puthan Veetil, C. R. Reis, W. J. Quax, D. B. Janssen, B. L. Feringa, G. J. Poelarends, *ChemBioChem* **2009**, *10*, 2236-2245.

14. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., **1989**, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

15. W. J. Waddell, *J. Lab. Clin. Med.* **1956**, *48*, 311-314.

# 4

## Asymmetric Addition of Acetaldehyde to β-Nitrostyrenes Catalyzed by a Promiscuous Proline-Based Tautomerase

*E. Zandvoort, E. M. Geertsema, B. J. Baas, W. J. Quax, G. J. Poelarends*

## Abstract

The enzyme 4-oxalocrotonate tautomerase (4-OT) belongs to the tautomerase superfamily. The members of this superfamily all contain a conserved catalytic amino-terminal proline residue (Pro-1), that naturally acts as a general acid or general base catalyst. However, because of its low p$K_a$ of ~6.4, residue Pro-1 of 4-OT is able to act as a nucleophile at neutral pH. This allows 4-OT to promiscuously catalyze aldol condensation and dehydration reactions (Chapters 2 and 3). In our search for other promiscuous carbon-carbon bond-forming activities in 4-OT, we found that 4-OT catalyzes the Michael-type addition of acetaldehyde to *trans*-nitrostyrene, yielding 4-nitro-3-phenylbutanal. Catalysis proceeds via enamine formation of the amino-terminal proline residue with acetaldehyde, in reminiscence to organocatalysis. High stereoselectivity, low catalyst loading and water as reaction medium characterize this new, environmentally benign methodology.
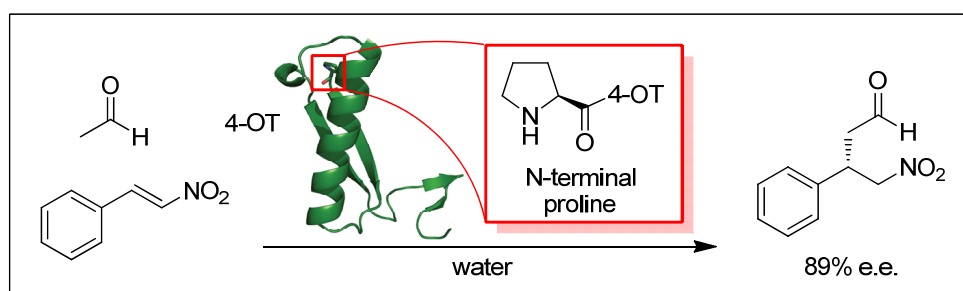
## Introduction

In recent years, organocatalysis has become one of the main areas in asymmetric catalysis of carbon-carbon bond-forming reactions.[1] The fast evolution of the organocatalysis field has been particularly fueled by aminocatalysis, in which secondary and primary amines react with carbonyl compounds to give enamine and iminium ion intermediates. The field was given a full face during the last two decades by the seminal contributions of List,[2] MacMillan,[3] Yamaguchi,[4] and co-workers. The natural, chiral amino acid proline and derivatives thereof were found to be powerful organocatalysts. These secondary amines are applied in substoichiometric quantities and afford high product yields and enantioselectivities in fundamental carbon-carbon bond-forming reactions such as aldolizations,[1,2,3b] Michael additions,[1,4,5] Mannich reactions[1,6] and Knoevenagel condensations.[1,7]

Inspired by the versatile success of proline and its derivatives as organocatalysts, we examined whether the enzyme 4-oxalocrotonate tautomerase (4-OT),[8] which carries a catalytic amino-terminal proline, might be suitable to promiscuously catalyze carbon-carbon bond-forming reactions. Herein, we describe the discovery and characterization of two 4-OT-catalyzed asymmetric carbon-carbon Michael-type addition reactions. Considering our previously reported 4-OT-catalyzed aldolizations,[9] this work is a pivotal step forward towards our aim to bridge organocatalysis and biocatalysis by developing a new class of biocatalysts that use the powerful proline-based enamine mechanism of organocatalysts,[1] but that take advantage of the water solubility and relatively high catalytic rates available with enzymes. A few elegant studies on promiscuous enzyme-catalyzed carbon-carbon Michael additions have been reported, but most of these reactions proceed in organic solvents and with moderate stereocontrol.[10]

4-OT is a stable enzyme composed of six identical subunits of only 62 amino acid residues each.[11] It belongs to the tautomerase superfamily, a group of homologous proteins that share a conserved catalytic amino-terminal proline and a characteristic β-α-β structural fold.[8,12] 4-OT takes part in a degradation pathway for aromatic hydrocarbons in *Pseudomonas putida* mt-2, where it catalyzes the tautomerization of 2-hydroxyhexa-2,4-dienedioate (**1**) into 2-oxohex-3-enedioate (**2**) (Scheme 4.1).[13] The key catalytic residues of 4-OT are Pro-1, Arg-11 and Arg-39. Residue Pro-1 functions as a general base (p$K_a$ ~6.4) that transfers the 2-hydroxyl proton of **1** to the C-5 position to give **2**.[14] Arg-11 and Arg-39 are important for binding of **1** and interact with respectively the C-6 and C-1 carboxylate group of **1**.[15] 4-OT also accepts (phenyl)enolpyruvate (**3**) and (*p*-hydroxyphenyl)enolpyruvate (**5**) as substrates for tautomerization, yielding (phenyl)pyruvate (**4**) and (*p*-hydroxyphenyl)pyruvate (**6**) as products, respectively (Scheme 4.1).[16]
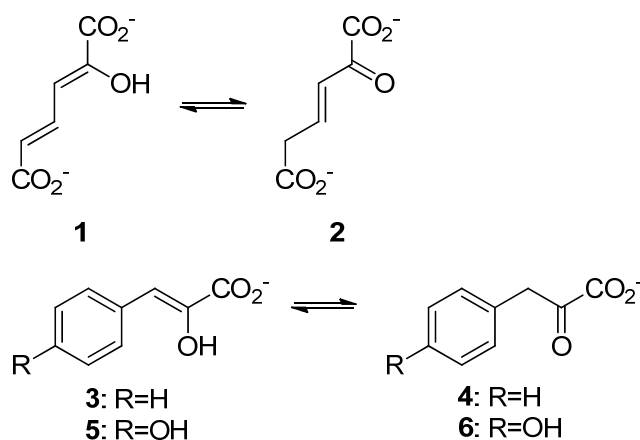
**Scheme 4.1.** Natural tautomerization reactions catalyzed by 4-OT.

## Results and Discussion

To investigate whether 4-OT exhibits promiscuous carbon-carbon bond-forming activities, the Michael-type addition of acetaldehyde (**7**) to *trans*-nitrostyrene (**8**) and *p*-hydroxy-*trans*-nitrostyrene (**10**) were selected as model reactions (Scheme 4.2) for a number of reasons. First, previous labeling experiments have indicated that residue Pro-1 of 4-OT rapidly attacks the carbonyl carbon of **7**.[9] The resulting enamine has nucleophilic character and may act as a donor. Second, the possible acceptors **8** and **10** show structural resemblance to compounds **3** and **5** (Scheme 4.1), two known substrates of 4-OT.[14,16] Third, the presumed product of the reaction between **7** and **8** is 4-nitro-3-phenylbutanal (**9**), an important precursor for the anti-depressant Phenibut (4-amino-3-phenylbutanoic acid).[5c,d]
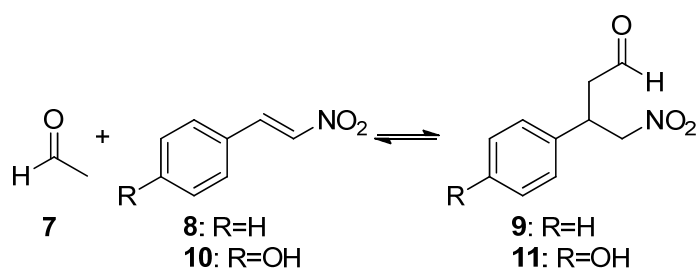


**Scheme 4.2.** Non-natural Michael-type addition reactions catalyzed by 4-OT.

The Michael-type addition reaction of **7** and **8** ($\lambda_{max}$ = 320 nm) was monitored by UV spectroscopy. Incubation of 4-OT (73 μM) with **7** (50 mM) and **8** (1.3 mM), resulted in a decrease in the absorbance at 320 nm ($A_{320}$), indicating the disappearance of **8** (Figure 4.1A and Figure 4.2A). To ensure that the depletion of **8** is enzyme-catalyzed, **7** and **8** were also

incubated without 4-OT. This experiment did not result in a similar decrease in the $A_{320}$ value (Figure 4.1B and Figure 4.2A) as observed for the incubation with 4-OT. The 4-OT sample used in the assay was highly purified. To eliminate the concern that a contaminating enzyme from the expression strain could be responsible for the observed activity, a 4-OT sample free of cellular enzymes was prepared by total chemical synthesis.[17] Indeed, incubation of **7** and **8** with chemically synthesized 4-OT resulted in the same decrease in the $A_{320}$ value as observed for the incubation with recombinant 4-OT (Figure 4.1C and Figure 4.2A). To exclude the possibility of nucleophilic addition to **8** of water or ethanol, present as main solvent and co-solvent, respectively, 4-OT was incubated with **8** in the absence of **7**. The small decrease in the $A_{320}$ value over time (Figure 4.1D) was similar to that observed for the incubation of **7** and **8** without enzyme (Figure 4.1B), thus demonstrating that 4-OT does not catalyze the addition of water or ethanol to **8**, but solely catalyzes the addition of **7** to **8**.
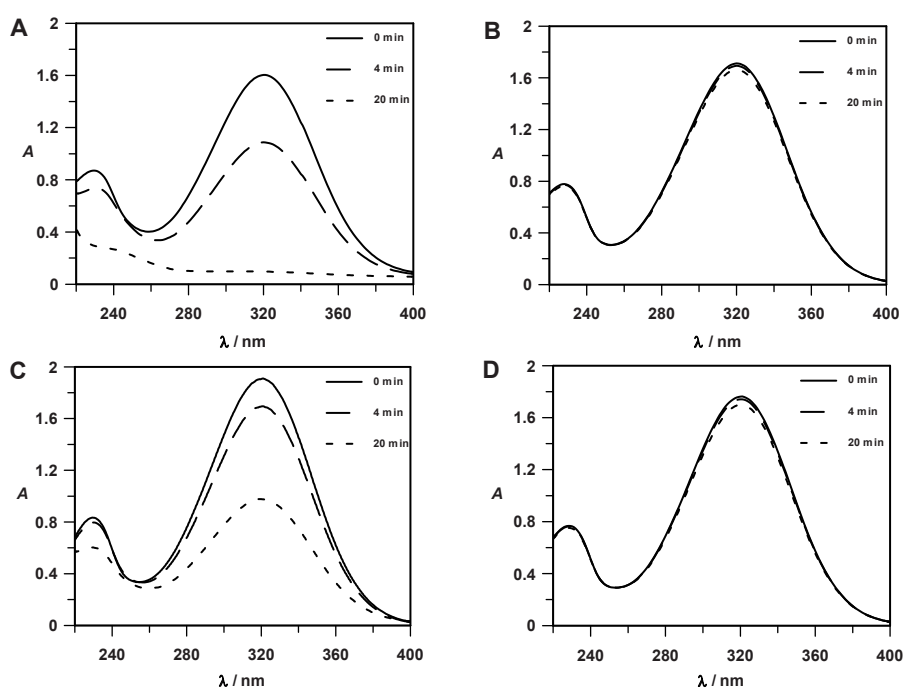


**Figure 4.1.** UV spectra monitoring the depletion of nitrostyrene (**8**) in the presence of A) acetaldehyde (**7**) and recombinant 4-OT, B) acetaldehyde (**7**), C) acetaldehyde (**7**) and synthetic 4-OT, and D) recombinant 4-OT.
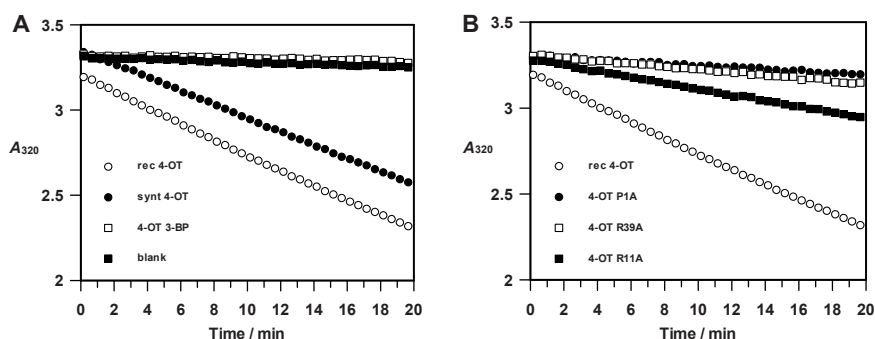
**Figure 4.2.** A) UV traces monitoring the depletion of nitrostyrene (**8**) in the presence of acetaldehyde (**7**) and recombinant 4-OT (rec 4-OT), synthetic 4-OT (synt 4-OT), 4-OT labeled with 3-bromopyruvate (4-OT 3-BP), or without 4-OT (blank). B) UV traces monitoring the depletion of nitrostyrene (**8**) in the presence of acetaldehyde (**7**) and recombinant 4-OT, 4-OT P1A, 4-OT R39A, or 4-OT R11A.

To identify the product of the 4-OT-catalyzed Michael-type reaction between **7** and **8**, a preparative scale reaction was performed with **7** (50 mM), **8** (2 mM, 18 mg), and 4-OT (0.7 mol%, 14.7 μM) in $NaH_2PO_4$ buffer (20 mM, pH 7.3, 60 mL). After 3 hours reaction time and a standard work up, $^1H$ NMR spectroscopy indicated complete depletion of **8** and 46% conversion into **9**. Subsequent purification of **9** (Figure 4.3) and analysis by HPLC on a chiral stationary phase (Figure 4.4) revealed that 4-OT is highly stereoselective, producing the (*S*)-enantiomer of **9** with 89% *ee*.
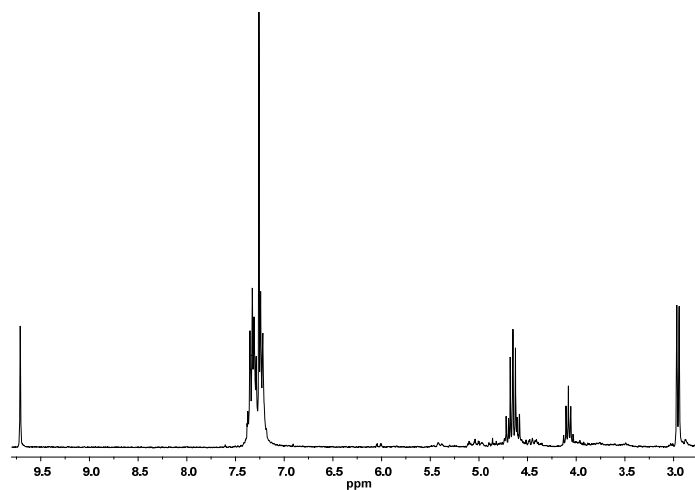


**Figure 4.3.** $^1H$ NMR spectrum of enzymatically prepared 4-nitro-3-phenylbutanal (**9**).
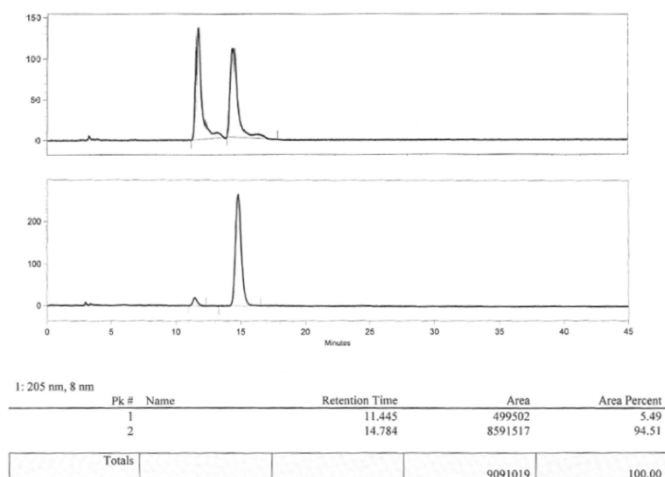
**Figure 4.4.** Determination of *ee* of 2-(3-nitro-2-phenylpropyl)-1,3-dioxolane (**12**), the ethylene glycol acetal of 4-nitro-3-phenylbutanal (**9**), by HPLC on a chiral stationary phase.

After we established that 4-OT catalyzes the addition of **7** to **8** to yield **9**, kinetic parameters were determined. The rate of the 4-OT-catalyzed reaction was dependent on the concentrations of both **7** and **8**. Apparent kinetic parameters were estimated at a fixed concentration (50 mM) of **7** and varying concentrations of **8**. 4-OT catalyzes the addition reaction with a $k_{cat}$ value of $1.7 \times 10^{-2}$ s$^{-1}$ and a $K_m$ value of 250 µM, resulting in a $k_{cat}/K_m$ value of 68 M$^{-1}$ s$^{-1}$.

We next investigated the importance of residue Pro-1 to the Michael-type addition activity of 4-OT by mutagenesis and labeling experiments. First, residue Pro-1 of 4-OT was mutated to an alanine. This 4-OT P1A mutant (29.4 µM) was incubated with **7** (50 mM) and **8** (2 mM) in NaH$_2$PO$_4$ buffer (20 mM, pH 7.3), and the depletion of **8** was followed by UV spectroscopy (Figure 4.2B). A similar experiment was performed with 4-OT inhibited by 3-bromopyruvate, an irreversible inhibitor that alkylates Pro-1 (Figure 4.2A).[18] Both mutant and alkylated 4-OT almost completely lost their activity, thereby indicating that Pro-1 is crucial for the enzymatic Michael-type addition of **7** to **8**. To evaluate the importance of the two active site arginines (Arg-11 and Arg-39) for the promiscuous Michael-type addition activity of 4-OT, the corresponding alanine mutants were prepared. In separate reactions, 4-OT R11A and 4-OT R39A (29.4 µM each) were incubated with **7** (50 mM) and **8** (2 mM) in NaH$_2$PO$_4$ buffer (20 mM, pH 7.3), and depletion of **8** was monitored by UV spectroscopy (Figure 4.2B). Mutant 4-OT R39A almost completely lost activity, while mutant 4-OT R11A showed a greatly reduced activity. This suggests that both arginine residues are crucial for the promiscuous activity of 4-OT. The importance of residues Pro-1, Arg-11, and Arg-39

strongly suggests that the addition of **7** to **8** takes place in the same active site as the natural tautomerization reactions (Scheme 4.1) catalyzed by 4-OT.[14-16]
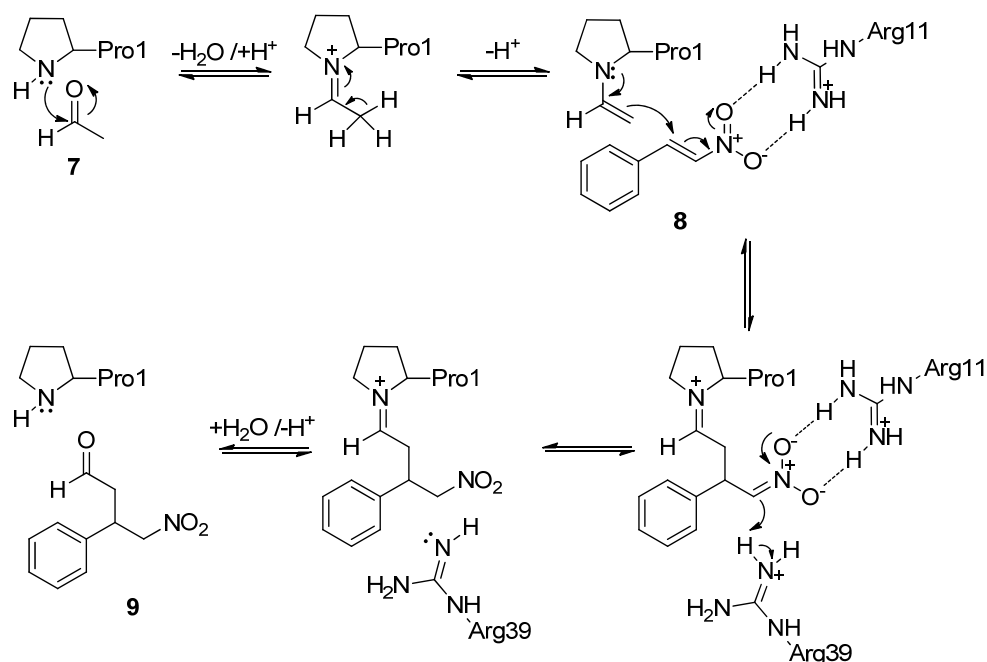


**Scheme 4.3.** Proposed mechanism for the 4-OT-catalyzed addition of **7** to **8** to yield **9**.

In a possible mechanism that may explain the Michael-type addition activity of 4-OT, residue Pro-1 acts as nucleophile and attacks the carbonyl carbon of **7** (Scheme 4.3).[1,9] The formed iminium ion is deprotonated, thereby resulting in an enamine intermediate. This nucleophilic species reacts with **8** in a Michael-type addition. Residue Arg-11 likely ensures correct substrate binding by coordinating the nitro functionality of **8** in analogy to the interaction of residue Arg-11 with the C-6 carboxylate group of the natural substrate **1**. Residue Arg-39 might act as a general acid catalyst, delivering a proton at the Cα carbon atom of **8**. The final product **9** is released from the Pro-1 residue of 4-OT by hydrolysis.

UV spectroscopic assays demonstrated that compound **10** (Scheme 4.2) is also accepted as substrate by 4-OT (Figure 4.5 and Figure 4.6). The 4-OT-catalyzed addition of **7** to **10** yields predominantly the (*S*)-enantiomer of 4-nitro-3-(4-hydroxyphenyl)butanal (**11**) (Figure 4.7 and Figure 4.8). A preparative scale reaction gave **11** in 65% yield and with 51% *ee*. The apparent kinetic parameters for the 4-OT-catalyzed addition of **7** to **10** ($K_m$ = 1.6 mM for **10**; $k_{cat}$ = 5.9 ×$10^{-2}$ $s^{-1}$; $k_{cat}/K_m$ = 37 $M^{-1}s^{-1}$) are similar to those determined for

the addition of **7** to **8**. As expected, the active site residues Pro-1, Arg-11 and Arg-39 are also important for the 4-OT-catalyzed addition of **7** to **10** (Figure 4.6).
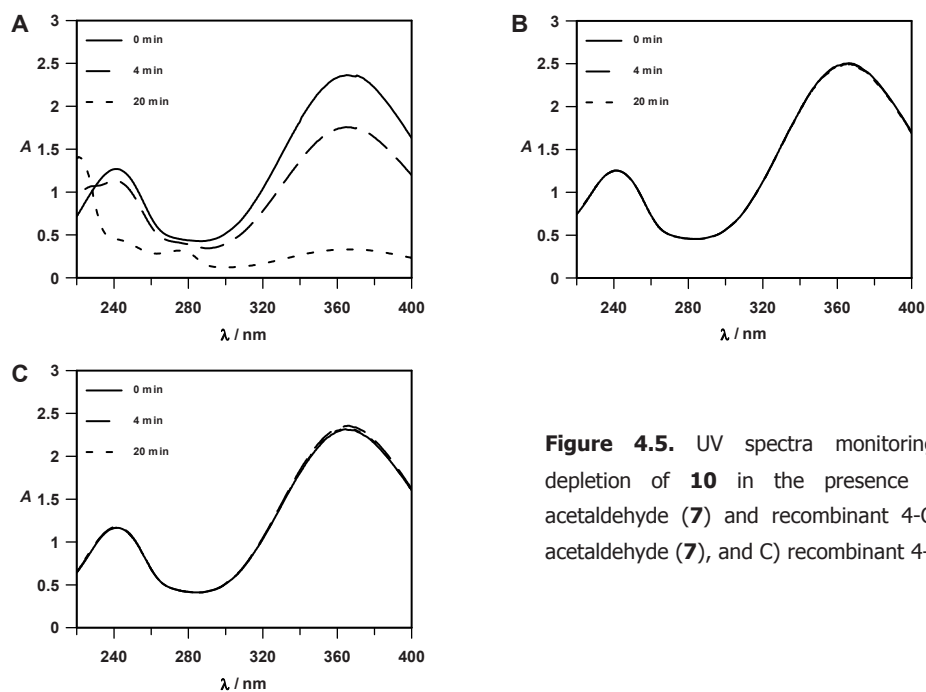


**Figure 4.5.** UV spectra monitoring the depletion of **10** in the presence of A) acetaldehyde (**7**) and recombinant 4-OT, B) acetaldehyde (**7**), and C) recombinant 4-OT.
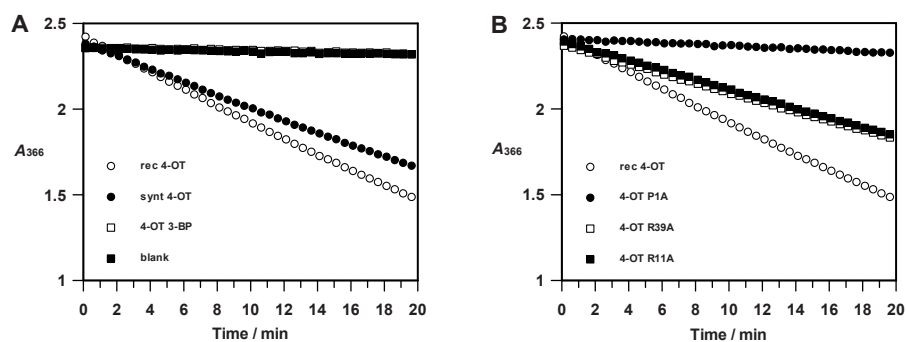


**Figure 4.6.** A) UV traces monitoring the depletion of **10** in the presence of acetaldehyde (**7**) and recombinant 4-OT, synthetic 4-OT, 4-OT labeled with 3-bromopyruvate (3-BP) and without 4-OT (blank). B) UV traces monitoring the depletion of **10** in the presence of acetaldehyde (**7**) and recombinant 4-OT, 4-OT P1A, 4-OT R39A and 4-OT R11A.
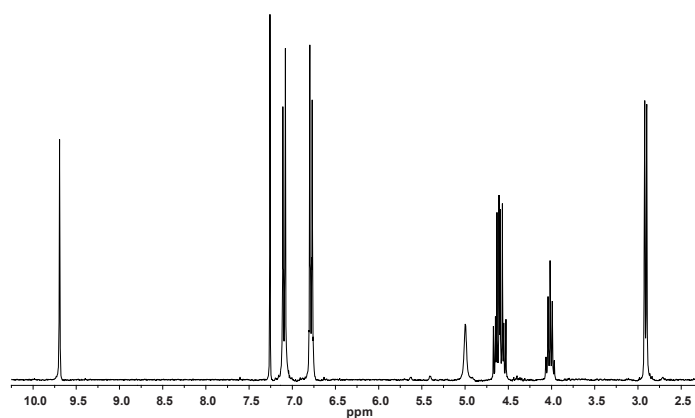
**Figure 4.7.** ^{1}H NMR spectrum of enzymatically prepared 4-nitro-3-(4-hydroxyphenyl)butanal (**11**).



| 1: 205 nm, 8 nm | | | | |
|---|---|---|---|---|
| Pk # | Name | Retention Time | Area | Area Percent |
| 1 | | 23.147 | 320169 | 24.70 |
| 2 | | 26.517 | 976277 | 75.30 |
| Totals | | | 1296446 | 100.00 |

**Figure 4.8.** Determination of *ee* of 4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**), the ethylene glycol acetal of 4-nitro-3-(4-hydroxyphenyl)butanal (**11**), by HPLC on a chiral stationary phase.

## Conclusion

In summary, we have shown that the proline-based tautomerase 4-OT is an efficient catalyst for the asymmetric Michael-type additions of acetaldehyde (**7**) 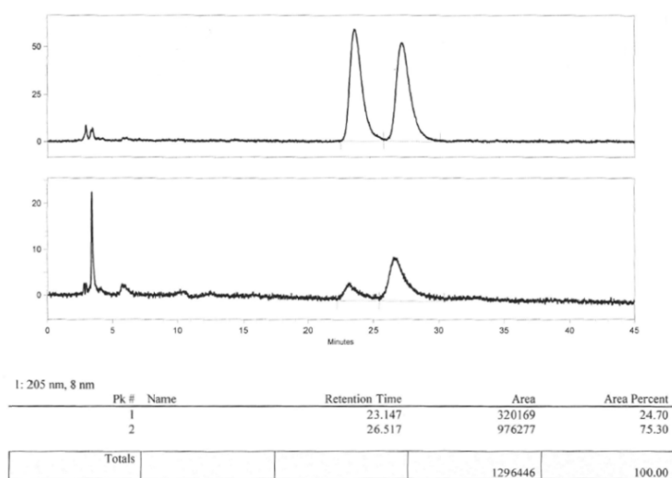to nitrostyrenes **8** and **10**. Product **9** was obtained in 46% yield and good enantiomeric excess (89% *ee*), while *p*-hydroxy-derivative **11** was furnished in 65% yield and with 51% *ee*. The characteristic Pro-1 residue of 4-OT is crucial for activity, and catalysis likely takes place through an enamine intermediate.[19] Our results compare well to the corresponding organocatalytic strategies reported (Table 4.1).[5c] Underivatized (*S*)-proline yielded **9** from **7** and **8** in low yield (10%) and *ee* (30%), whereas the more advanced proline-based catalyst diphenylprolinol silyl ether gave **9** in 51% yield and with 92% *ee*. Our enzymatic system complements the classic organocatalytic approach in two ways. First, when compared to the amount of catalyst generally used in organocatalysis (10-20 mol%), the catalyst loading of 0.7 mol% 4-OT in our set-up is low. Indeed, for promiscuous non-natural reactions, 4-OT exhibits pronounced catalytic efficiencies of $k_{cat}/K_m = 68$ M$^{-1}$s$^{-1}$ (addition of **7** to **8**) and 37 M$^{-1}$s$^{-1}$ (addition of **7** to **10**). Second, the enzymatic reactions were not performed in organic solvent, but in an environmentally benign aqueous buffer. Together with our reported 4-OT-catalyzed aldol condensations,[9] these results serve as an excellent starting point for directed evolution experiments to further broaden the substrate and reaction scope of 4-OT, and improve reaction rates. If successful, these efforts may help to increase the number and diversity of biocatalysts that can be used for (large industrial scale) asymmetric carbon-carbon bond-forming reactions.

**Table 4.1.** Comparison of organo- and bio-catalytic synthesis of **9** and **11**

| Substrates | Product | Catalyst | Catalyst loading (mol%) | Solvent | Reaction time (h) | Yield (%) | *ee* (%) |
|---|---|---|---|---|---|---|---|
| **7** and **8**[a] | **9** | (*S*)-Proline | 20 | MeCN | 15 | 10 | 30 |
| **7** and **8**[a] | **9** | Diphenylpro-linol silyl ether | 20 | MeCN | 15 | 51 | 92 |
| **7** and **8**[b] | **9** | 4-OT | 0.7 | H$_2$O | 3 | 46 | 89 |
| **7** and **10**[b] | **11** | 4-OT | 0.7 | H$_2$O | 3 | 65 | 51 |

[a] Data taken from reference.[5c]

[b] This work.

# Experimental Procedures

*Materials:* All chemicals were obtained from Sigma-Aldrich Chemical Co. (St. Louis, MO), unless stated otherwise. The sources for the biochemicals, buffers, solvents, components of Luria-Bertani (LB) media as well as the materials, enzymes, and reagents used in the molecular biology procedures are reported elsewhere.[20] High-purity synthetic 4-OT (lyophilized, with Met-45 replaced by Nle to prevent oxidation upon sample handling) was obtained from GenScript USA Inc. (Piscataway, NY) and folded into the active homohexamer as described before.[21]

*General methods:* Techniques for restriction enzyme digestions, ligation, transformation, and other standard molecular biology manipulations were based on methods described elsewhere.[22] The PCR was carried out in a DNA thermal cycler (model GS-1) obtained from Biolegio (Nijmegen, The Netherlands). DNA sequencing was performed by Macrogen (Seoul, Korea). Protein was analyzed by polyacrylamide gel electrophoresis (PAGE) using sodium dodecyl sulfate (SDS) gels containing polyacrylamide (10%). The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the Waddell method.[23] Kinetic data were obtained on a V-650 or V-660 spectrophotometer from Jasco (IJsselstein, The Netherlands) and were fitted by nonlinear regression data analysis using the GraFit program (Erithacus, Software Ltd., Horley, U.K.) obtained from Sigma Chemical Co. $^1$H NMR spectra were recorded on a Varian Inova 300 (300 MHz) spectrometer. Chemical shifts for protons are reported in parts per million scale (δ scale) and are referenced to CHCl$_3$ (δ = 7.26). $^{13}$C NMR spectra were recorded on a Varian-Gemini 200 (50 MHz) spectrometer. Chemical shifts for carbons are reported in parts per million scale (δ scale) and are referenced to CHCl$_3$ (δ = 77.00). The masses of 4-OT and 4-OT mutants were determined by ESI-MS using a Sciex API 3000 triple quadrupole mass spectrometer (AB Sciex, Concord, Ontario, Canada) which is housed in the Mass Spectrometry Facility Core in the Department of Pharmacy at the University of Groningen.

*Expression and purification of wild-type 4-OT and mutants:* 4-OT and 4-OT mutants were expressed in *E. coli* BL21(DE3) cells harboring the appropriate expression vector (based on the pET20b(+) plasmid). The construction of the vectors and the expression and purification of 4-OT and 4-OT mutants have been reported before.[9]

*Labeling of 4-OT with 3-bromopyruvate:* 4-OT was labeled with 3-bromopyruvate as described previously.[9]

*General UV spectroscopic assay for Michael-type addition of acetaldehyde (**7**) to trans-nitrostyrene (**8**) or p-hydroxy-trans-nitrostyrene (**10**):* The Michael-type addition activity of recombinant 4-OT, synthetic 4-OT, the 4-OT mutants, or 4-OT inactivated by 3-bromopyruvate was monitored by following the depletion of the absorbance of **8** at 320 nm (ε = 14.4 mM$^{-1}$ cm$^{-1}$) or **10** at 366 nm (ε = 11.4 mM$^{-1}$ cm$^{-1}$). Enzyme (150 μg, 73 μM, monomer concentration) was incubated in a 1 mm cuvette with **7** (50 mM) and **8** or **9** (1.3 mM, from a 13 mM stock solution in 100% EtOH) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3; 0.3 mL final volume). Absorbance spectra were recorded from 200 to 400 nm.

*Kinetic assays:* The kinetic assays were performed at 22 °C by following the decrease in absorbance at 320 or 366 nm, which corresponds to the disappearance of *trans*-nitrostyrene (**8**) and *p*-hydroxy-*trans*-nitrostyrene (**10**), respectively. A stock solution of 4-OT (0.25 mg/ml; 37 μM, monomer concentration) in 20 mM $NaH_2PO_4$ buffer (pH 7.3) was prepared. Fresh stock solutions of acetaldehyde (**7**) and **8** (or **10**) were made in 20 mM $NaH_2PO_4$ buffer (pH 7.3) and absolute ethanol, respectively. An aliquot (240 μL) of the enzyme solution was added to a 1 mm cuvette. The enzyme activity was assayed by the addition of 30 μL from each of the stock solutions of acetaldehyde (**7**) and **8** (or **10**), yielding final concentrations of 29.4 μM 4-OT, 50 mM acetaldehyde (**7**), and 0.05 to 1.5 mM *trans*-nitrostyrene (**8**) or 0.05 to 3 mM *p*-hydroxy-*trans*-nitrostyrene (**10**). The initial rates (μM/s) were plotted versus the concentration of **8** or **10** (μM). GraFit was used to fit the data to Michaelis-Menten kinetics and to calculate the kinetic parameters $k_{cat}$ and $K_m$.

*General procedure for preparative scale reaction and work-up:* A stock solution of 20 mM nitrostyrene (either **8** or **10**) in absolute ethanol was prepared. This stock solution (6 mL) was added to 54 mL of 20 mM $NaH_2PO_4$ (pH 7.3) buffer containing acetaldehyde (**7**) and 4-OT in a 100 mL flask. The final concentrations were 50 mM (3 mmol) acetaldehyde (**7**), 2 mM (120 μmol) nitrostyrene and 14.7 μM (monomer concentration; 6 mg, 0.1 mg/ml, 0.88 μmol, 0.7 mol%) 4-OT. The mixture was incubated at 22 °C for 3-4 hours. The progress of the reaction was monitored by UV/Vis analysis. Decrease of absorbance at 320 nm indicated the depletion of nitrostyrene **8** (at 366 nm for **10**). When the reaction was complete, 4-OT was filtered off using a Vivaspin centrifugal concentrator equipped with a 5000 Da molecular weight cut-off filter (Sartorius Stedim Biotech S.A., France). The flowthrough was collected and extracted with ethylacetate (3 × 40 mL). The combined organic layers were dried ($MgSO_4$) and concentrated *in vacuo*. ^{1}H NMR spectroscopy revealed complete conversion of nitrostyrene and the yield of the desired product was determined. Formation of byproducts was also observed, but these undesired products have not been identified.The residue was purified by column chromatography (silica gel, visualization: $KMnO_4$). The obtained product, γ-nitroaldehyde, was converted into its corresponding ethylene glycol acetal for enantiomeric excess (*ee*) determination by HPLC on a chiral stationary phase. As control experiments blank reactions of acetaldehyde and Michael acceptor without enzyme were performed under the exact same conditions. No formation of products was observed.

*4-nitro-3-phenylbutanal (**9**)*
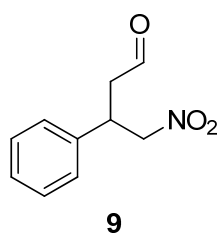


Yield crude: 46%; *trans*-nitrostyrene (**8**) (17.9 mg, 120 μmol) gave 4-nitro-3-phenylbutanal (**9**) (9.5 mg, 49 μmol, 41%) as a yellowish oil (silica gel, *n*-hexane/EtOAc 5/1, $R_f$ product = 0.20). HPLC (Chiralcel OD, heptane/*iso*-propanol 90/10, 1 mL/min.): $R_{t-1}$ ((*R*)-enantiomer) = 10.5 min. (minor); $R_{t-2}$ ((*S*)-enantiomer) = 12.9 min. (major); *ee* = 89% (Figure 4.4). Spectroscopical (Figure 4.3) and HPLC data are in accordance with the literature.[24,25]

### 4-nitro-3-(4-hydroxyphenyl)butanal (**11**)
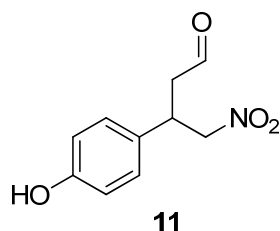
Yield crude: 65%; *p*-hydroxy-*trans*-nitrostyrene (**10**) (19.8 mg, 120 μmol) gave 4-nitro-3-(4-hydroxyphenyl)butanal (**11**) (14.8 mg, 71 μmol, 59%) as a yellowish oil (silica gel, hexanes/EtOAc 3/1, $R_f$ product = 0.40 with hexanes/EtOAc 1/1; visualization: KMnO$_4$). HPLC (Chiralcel OD, heptane/*iso*-propanol 90/10, 1 mL/min.): $R_{t-1}$ = 23.1 min. (minor); $R_{t-2}$ = 26.5 min. (major); *ee* = 51% (Figure 4.8). Spectroscopical (Figure 4.7) and HPLC data are in accordance with synthetically obtained racemic 4-nitro-3-(4-hydroxyphenyl)butanal (**11**) (*vide infra*).
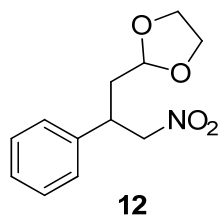
### Racemic 4-nitro-3-phenylbutanal (**9**)

Racemic 4-nitro-3-phenylbutanal (**9**) was synthesized according to a literature procedure.[25]

### Racemic 4-nitro-3-(4-hydroxyphenyl)butanal (**11**)

*p*-Hydroxy-*trans*-nitrostyrene (**10**) (100 mg, 0.61 mmol), acetaldehyde (**7**) (300 mg, 6.8 mmol) and piperidine (10 mg, 0.12 mmol) were dissolved in MeOH (2 mL). The solution was stirred for 24 h at room temperature and the reaction was monitored by TLC (silica, hexanes/EtOAc 1/1; $R_f$ starting material = 0.62; $R_f$ product = 0.40; visualization: KMnO$_4$). The solvent was evaporated and the residue purified by column chromatography (silica gel, hexanes/EtOAc 3/1) to give 4-nitro-3-(4-hydroxyphenyl)butanal (**11**) (35 mg, 0.17 mmol, 27%) as a yellowish oil. HPLC (Chiralcel OD, heptane/*iso*-propanol 90/10, 1 mL/min.): $R_{t-1}$ = 23.8 min.; $R_{t-2}$ = 27.5 min. (Figure 4.8); $^1$H NMR (300 MHz, CDCl$_3$, 25 °C): δ 9.68 (s, 1H), 7.09 (d, *J* = 8.5 Hz, 2H), 6.77 (d, *J* = 8.5 Hz, 2H), 5.24 (s, 1H), 4.64 (dd, *J* = 12.4, 7.2 Hz, 1H), 4.56 (dd, *J* = 12.4, 7.6 Hz, 1H), 4.01 (ddd, *J* = 7.6, 7.2, 7.2 Hz, 1H), 2.91 (d, *J* = 7.2 Hz, 2H) (Figure 4.7); $^{13}$C NMR (50 MHz, CDCl$_3$, 25 °C): δ 199.29, 155.37, 129.98, 128.66, 116.05, 79.67, 46.49, 37.33; HRMS (ESI): *m/z* = 208.0616 [M-H]$^-$ (calcd. 208.0610 for C$_{10}$H$_{10}$NO$_4$).

### General procedure for conversion of γ-nitroaldehyde into corresponding ethylene glycol acetal for enantiomeric excess (ee) determination by HPLC on a chiral stationary phase:

This conversion was done according to a modified literature procedure.[25] Ethylene glycol (4 eq) and *p*-TsOH (0.1 eq) were added to a 60 mM solution of γ-nitroaldehyde in CH$_2$Cl$_2$. The reaction mixture was stirred for 24 h at room temperature and monitored by TLC (silica, hexanes/EtOAc 1/1; visualization: KMnO$_4$). Solvents were removed *in vacuo* and the residue was purified by column chromatography.

### 2-(3-nitro-2-phenylpropyl)-1,3-dioxolane (**12**)

4-Nitro-3-phenylbutanal (**9**) (4 mg, 2.1×10$^{-2}$ mmol), as obtained from the enzyme-catalyzed reaction (*vide supra*), gave 2-(3-nitro-2-phenylpropyl)-1,3-dioxolane (**12**) (2 mg, 8.4×10$^{-3}$ mmol, 40%) as a yellowish oil after column chromatography (silica gel, *n*-hexane/EtOAc 6/1, $R_f$ product = 0.20); HPLC (Chiralcel OD, heptane/*iso*-propanol 90/10, 1 mL/min.): $R_{t-1}$ = 10.5 min. (minor); $R_{t-2}$ = 12.9 min. (major); *ee* = 89% (Figure 4.4). $^1$H NMR (300 MHz,

CDCl$_3$, 25 °C): δ 7.37-7.20 (m, 5H), 4.77 (dd, $J$ = 12.5, 6.5 Hz, 1H), 4.74 (dd, $J$ = 6.4, 3.2 Hz, 1H), 4.59 (dd, J = 12.5, 8.9 Hz, 1H), 4.01-3.92 (m, 2H), 3.89-3.72 (m, 3H), 2.13 (ddd, $J$ = 14.1, 7.8, 3.2 Hz, 1H), 1.98 (ddd, $J$ = 14.1, 6.5, 6.4 Hz, 1H) (Figure 4.9); $^{13}$C NMR (50 MHz, CDCl$_3$, 25 °C): δ 139.24, 128.96, 127.69, 127.38, 102.24, 80.35, 64.94, 64.85, 39.93, 37.27.
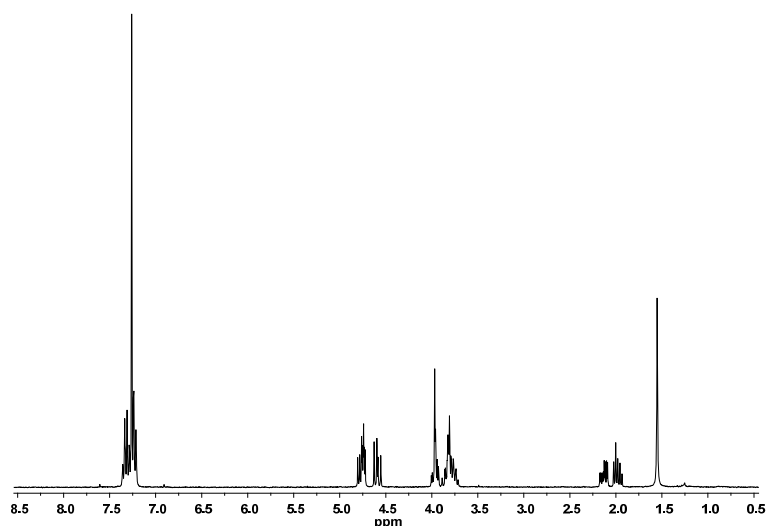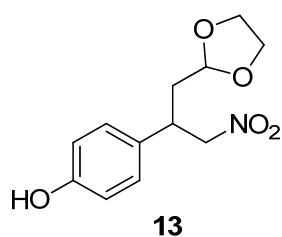


**Figure 4.9.** $^1$H NMR spectrum of 2-(3-nitro-2-phenylpropyl)-1,3-dioxolane (**12**), the ethylene glycol acetal of **9**.

*Racemic 2-(3-nitro-2-phenylpropyl)-1,3-dioxolane (**12**)*

Racemic 2-(3-nitro-2-phenylpropyl)-1,3-dioxolane (**12**) was synthesized according to a literature procedure.[25]

*4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**)*



4-nitro-3-(4-hydroxy-phenyl)butanal (**11**) (4 mg, 1.9×10$^{-2}$ mmol), as obtained from the enzyme-catalyzed reaction (*vide supra*), gave 4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**) (2 mg, 7.9×10$^{-3}$ mmol, 42%) as a yellowish oil after column chromatography (silica gel, hexanes/EtOAc 3/1); HPLC (Chiralcel OD, heptane/*iso*-propanol 90/10, 1 mL/min.): R$_{t-1}$ = 23.1 min. (minor); R$_{t-2}$ = 26.5 min. (major); *ee* = 51% (Figure 4.8). Spectroscopical and HPLC data in accordance with racemic 4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**) (*vide infra*).

*Racemic 4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**)*

Racemic 4-nitro-3-(4-hydroxy-phenyl)butanal (**11**) (25 mg, 0.12 mmol) gave racemic 4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**) (13 mg, 5.1×10$^{-2}$ mmol, 43%) as a yellowish oil after column chromatography (silica gel, hexanes/EtOAc 3/1); HPLC (Chiralcel OD, heptane/*iso*-propanol 90/10, 1 mL/min.): R$_{t-1}$ = 23.8 min.; R$_{t-2}$ = 27.5 min. (Figure 4.8); $^1$H NMR (300 MHz, CDCl$_3$, 25 °C): δ 7.08 (d, $J$ = 8.5 Hz, 2H), 6.76 (d, $J$ = 8.5 Hz, 2H), 5.08 (br, 1H), 4.74 (dd, $J$ = 6.5, 2.9 Hz, 1H), 4.72

(dd, $J$ = 12.6, 6.5 Hz, 1H), 4.53 (dd, $J$ = 12.6, 9.1 Hz, 1H), 4.01-3.90 (m, 2H), 3.89-3.77 (m, 2H), 3.69 (dddd, $J$ = 9.1, 8.1, 7.3, 6.5 Hz, 1H), 2.09 (ddd, $J$ = 14.0, 8.1, 2.9 Hz, 1H), 1.94 (ddd, $J$ = 14.0, 7.3, 6.5 Hz, 1H) (Figure 4.10); $^{13}$C NMR (50 MHz, CDCl$_3$, 25 °C): δ 155.05, 131.13, 128.64, 115.81, 102.29, 80.64, 64.92, 64.84, 39.26, 37.31; HRMS (ESI): $m/z$ = 252.0880 [M-H]$^-$ (calcd. 252.0872 for C$_{12}$H$_{14}$NO$_5$).
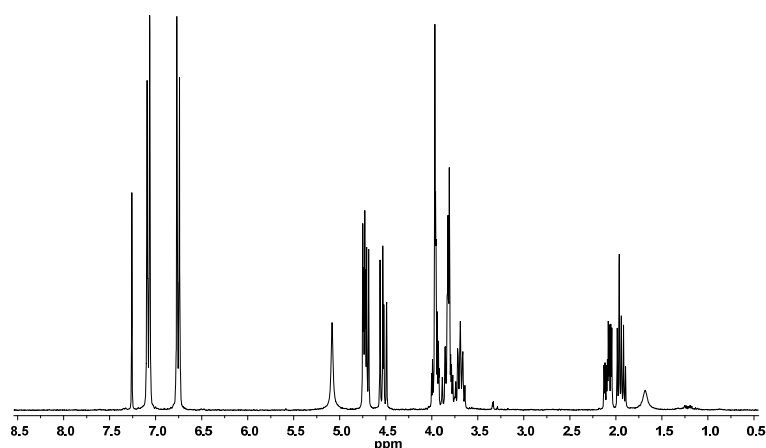


**Figure 4.10.** $^1$H NMR spectrum of 4-(1-(1,3-dioxolan-2-yl)-3-nitropropan-2-yl)phenol (**13**), the ethylene glycol acetal of **11**.

## Acknowledgements

# References

1. For reviews on organocatalysis, see: a) W. Notz, F. Tanaka, C. F. Barbas III, *Acc. Chem. Res.* **2004**, *37*, 580-591; b) A. Erkkilä, I. Majander, P. M. Pihko, *Chem. Rev.* **2007**, *107*, 5416-5470; c) S. Mukherjee, J. W. Yang, S. Hoffmann, B. List, *Chem. Rev.* **2007**, *107*, 5471-5569; d) L.-W. Xu, Y. Lu, *Org. Biomol. Chem.* **2008**, *6*, 2047-2053; e) L.-W. Xu, J. Luo, Y. Lu, *Chem. Commun.* **2009**, 1807-1821; f) B. List, *Angew. Chem.* **2010**, *122*, 1774-1779; *Angew. Chem. Int. Ed.* **2010**, *49*, 1730-1734.

2. a) B. List, R. A. Lerner, C. F. Barbas III, *J. Am. Chem. Soc.* **2000**, *122*, 2395-2396; b) W. Notz, B. List, *J. Am. Chem. Soc.*, **2000**, *122*, 7386-7387.

3. a) K. A. Ahrendt, C. J. Borths, D. W. C. MacMillan, *J. Am. Chem. Soc.* **2000**, *122*, 4243-4244; b) A. B. Northrup, D. W. C. MacMillan, *J. Am. Chem. Soc.* **2002**, *124*, 6798-6799.

4. a) M. Yamaguchi, N. Yokota, T. Minami, *J. Chem. Soc. Chem. Commun.* **1991**, 1088-1089; b) M. Yamaguchi, T. Shiraishi, M. Hirama, *Angew. Chem.* **1993**, *105*, 1243-1245; *Angew. Chem. Int. Ed.* **1993**, *32*, 1176-1178.

5. a) A. Kawara, T. Taguchi, *Tetrahedron Lett.* **1994**, *35*, 8805-8808; b) S. B. Tsogoeva, *Eur. J. Org. Chem.* **2007**, 1701-1716; c) P. García-García, A. Ladépêche, R. Halder, B. List, *Angew. Chem.* **2008**, *120*, 4797-4799; *Angew. Chem. Int. Ed.* **2008**, *47*, 4719-4721; d) O. G. Maltsev, A. S. Kucherenko, I. P. Beletskaya, V. A. Tartakovsky, S. G. Zlotin, *Eur. J. Org. Chem.* **2010**, 2927-2933.

6. a) B. List, *J. Am. Chem. Soc.* **2000**, *122*, 9336-9337; b) R. O. Duthaler, *Angew. Chem.* **2003**, *115*, 1005-1008; *Angew. Chem. Int. Ed.* **2003**, *42*, 975-978; c) J. W. Yang, C. Chandler, M. Stadler, D. Kampen, B. List, *Nature* **2008**, *452* (27), 453-455; d) C. Chandler, P. Galzerano, A. Michrowska, B. List, *Angew. Chem.* **2009**, *121*, 2012-2014; *Angew. Chem. Int. Ed.* **2009**, *48*, 1978-1980.

7. a) A. Erkkilä, P. M. Pihko, *J. Org. Chem.* **2006**, *71*, 2538-2541; b) A. Lee, A. Michrowska, S. Sulzer-Mosse, B. List, *Angew. Chem.* **2011**, *123*, 1745-1748; *Angew. Chem. Int. Ed.* **2011**, *50*, 1707-1710.

8. C. P. Whitman, *Arch. Biochem. Biophys.* **2002**, *402*, 1-13.

9. E. Zandvoort, B. J. Baas, W. J. Quax, G. J. Poelarends, *ChemBioChem* **2011**, *12*, 602-609.

10. a) M. Svedendahl, K. Hult, P. Berglund, *J. Am. Chem. Soc.* **2005**, *127*, 17988-17989; b) J. M. Xu, F. Zhang, B. K. Liu, Q. Wu, X. F. Lin, *Chem. Commun.* **2007**, 2078-2080; c) J. M. Xu, F. Zhang, Q. Wu, Q. Y. Zhang, X. F. Lin, *J. Mol. Catal. B: Enzym.* **2007**, *49*, 50-54; d) G. A. Strohmeier, T. Sović, G. Steinkellner, F. S. Hartner, A. Andryushkova, T. Purkarthofer, A. Glieder, K. Gruber, H. Griengl, *Tetrahedron* **2009**, 5663-5668, and references therein.

11. L. H. Chen, G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour, C. P. Whitman, *J. Biol. Chem.* **1992**, *267*, 17716-17721.

12. G. J. Poelarends, V. Puthan Veetil, C. P. Whitman, *Cell. Mol. Life Sci.* **2008**, *65*, 3606-3618.

13. a) S. Harayama, M. Rekik, K. L. Ngai, L. N. Ornston, *J. Bacteriol.* **1989**, *171*, 6251-6258; b) C. P. Whitman, B. A. Aird, W. R. Gillespie, N. J. Stolowich, *J. Am. Chem. Soc.* **1991**, *113*, 3154-3162.

14. S. C. Wang, W. H. Johnson, Jr., R. M. Czerwinski, S. L. Stamps, C. P. Whitman, *Biochemistry* **2007**, *46*, 11919-11929.

15. T. K. Harris, R. M. Czerwinski, W. H. Johnson, Jr., P. M. Legler, C. Abeygunawardana, M. A. Massiah, J. T. Stivers, C. P. Whitman, A. S. Mildvan, *Biochemistry* **1999**, *38*, 12343-12357.

16. E. A. Burks, C. D. Fleming, A. D. Mesecar, C. P. Whitman, S. D. Pegan, *Biochemistry* **2010**, *49*, 5016-5027.

17. 4-OT was synthesized by GenScript USA Inc. (Piscataway, NY).

18. J. T. Stivers, C. Abeygunawardana, A. S. Mildvan, G. Hajipour, C. P. Whitman, L. H. Chen, *Biochemistry* **1996**, *35*, 803-813.

19. In theory, the reaction could proceed via the enolate anion of **7** (a mechanism in which Pro-1 would function as a general base catalyst removing a proton from **7**) instead of via the derived enamine (a mechanism in which Pro-1 acts as nucleophile). Since Pro-1 rapidly reacts with **7** to form an imine/enamine, the alternative mechanism involving an enolate anion seems unlikely.

20. H. Raj, B. Weiner, V. Puthan Veetil, C. R. Reis, W. J. Quax, D. B. Janssen, B. L. Feringa, G. J. Poelarends, *ChemBioChem* **2009**, *10*, 2236-2245.

21. M. C. Fitzgerald, I. Chernushevich, K. G. Standing, S. B. H. Kent, C. P. Whitman, *J.Am. Chem. Soc.* **1995**, *117*, 11075-11080.

22. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, **1989**

23. W. J. Waddel, *J. Lab. Clin. Med.* **1956**, *48*, 311-314.

24. H. Gotoh, H. Ishikawa, Y. Hayashi, *Org. Lett.*, **2007**, *9*, 5307-5309.

25. I. Mager, K. Zeitler, *Org. Lett.*, **2010**, *12*, 1480-1483.

# 5

## An Unexpected Promiscuous Activity of 4-Oxalocrotonate Tautomerase: The *cis-trans* Isomerization of Nitrostyrene

*E. Zandvoort, E. M. Geertsema, B. J. Baas, W. J. Quax, G. J. Poelarends*

## Abstract

In addition to its natural tautomerization reaction, the enzyme 4-oxalocrotonate tautomerase (4-OT) can also promiscuously catalyze several carbon-carbon bond-forming reactions, including aldol condensation (Chapters 2 and 3) and Michael-type addition (Chapter 4) reactions. In the 4-OT-catalyzed Michael-type addition reaction, acetaldehyde is coupled to *trans*-nitrostyrene. While exploring *cis*-nitrostyrene as a potential electrophile in Michael-type addition reactions catalyzed by 4-OT, it was unexpectedly found that 4-OT catalyzes ($k_{cat}/K_m$ = 1.9 x $10^3$ $M^{-1}s^{-1}$) the isomerization of *cis*-nitrostyrene into *trans*-nitrostyrene. So far, this isomerase activity is the most efficient promiscuous activity of 4-OT. The amino-terminal proline residue (Pro-1) of 4-OT is shown to be essential for this *cis-trans* isomerization activity, and two possible catalytic mechanisms are proposed in which Pro-1 acts either as a nucleophile or as a general base.
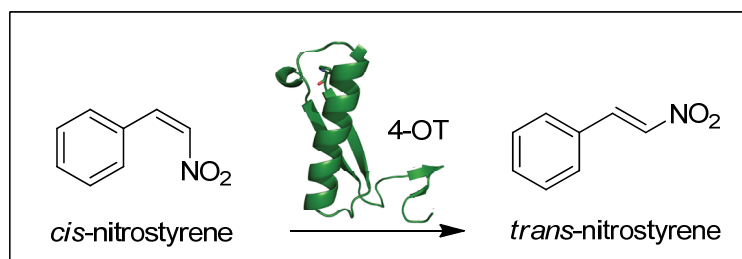
# Introduction

The tautomerase superfamily consists of homologous proteins with a characteristic β-α-β structural fold and a conserved catalytic amino-terminal proline.[1] 4-Oxalocrotonate tautomerase (4-OT) from *Pseudomonas putida* mt-2 is a member of this superfamily, and catalyzes the enol-keto tautomerization of 2-hydroxyhexa-2,4-dienedioate (**1**) into 2-oxohex-3-enedioate (**2**) (Scheme 5.1).[2] In this natural activity of 4-OT, the Pro-1 residue functions as the catalytic base and shuttles the 2-hydroxyl proton of **1** to the C-5 position to yield **2**.[3] Indeed, Pro-1 has the correct protonation state (p$K_a$ of Pro-1 in 4-OT is ~6.4) to be able to act as a catalytic base at cellular pH.[4] Intriguingly, 4-OT can also catalyze several C-C bond-forming reactions, including aldol condensation and Michael-type addition reactions.[5,6,7] In these unnatural activities of 4-OT, Pro-1 likely functions as a nucleophile, analogously to the role of the amino-acid L-proline and derivatives thereof in the field of organocatalysis.[8]
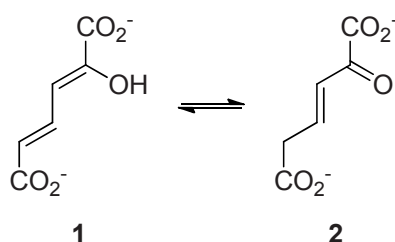


**Scheme 5.1**. The natural tautomerization reaction catalyzed by 4-OT.
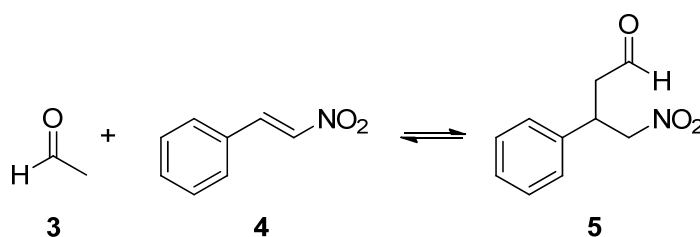


**Scheme 5.2.** The non-natural Michael-type addition catalyzed by 4-OT.

We have previously shown that 4-OT catalyzes the Michael-type addition of acetaldehyde (**3**) to *trans*-nitrostyrene (**4**) to yield 4-nitro-3-phenylbutanal (**5**) (Scheme 5.2).[7] While investigating the scope of this reaction, we assessed whether *cis*-nitrostyrene (**6**) (Scheme 5.3) can serve as an alternative substrate for 4-OT. Herein, we report the unexpected finding that 4-OT can promiscuously catalyze the *cis-trans* isomerization of nitrostyrene. Furthermore, we describe the characterization of this reaction, and propose

two possible mechanisms through which the conversion of **6** to **4** catalyzed by 4-OT might proceed.
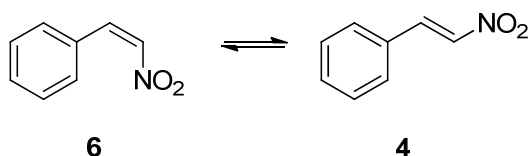


**Scheme 5.3.** The non-natural *cis-trans* isomerization of nitrostyrene catalyzed by 4-OT.

## Results and Discussion

The absorbance spectra of **4** and **6** largely overlap and both have an absorbance maximum around $\lambda=320$ nm corresponding to the vinylic C-C double bond (Figure 5.1A). However, the shape of the absorbance spectra differs slightly and the $\varepsilon_{320}$ values for **4** and **6** differ significantly, making it possible to distinguish between the two compounds when they are present in the same concentration ($\varepsilon_{cis,\mathbf{6}} = 0.31$ mM$^{-1}$ mm$^{-1}$ and $\varepsilon_{trans,\mathbf{4}} = 1.44$ mM$^{-1}$ mm$^{-1}$). To examine whether 4-OT catalyzes the Michael-type addition of **3** to **6**, 4-OT (73 µM) was incubated with **3** (50 mM) and **6** (2 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3). The reaction was monitored by UV spectroscopy and a decrease in A$_{320}$ was observed (Figure 5.1B). However, the starting absorbance was 2.8, which is ~4 times higher than the expected value of 0.7 for the concentration (2 mM) of **6** used (Figure 5.1A). Furthermore, the value of 2.8 accurately corresponds to the A$_{320}$ value expected for **4** (2 mM) (Figure 5.1A). When **6** and **3** were incubated without 4-OT, the A$_{320}$ value was stable at 0.7, which is in agreement with the expected A$_{320}$ value for **6** (2 mM) (Figure 5.1). Taken together, these observations suggest that 4-OT is able to rapidly catalyze (within the time of mixing 4-OT with **3** and **6**) the isomerization of **6** to **4** (Scheme 5.3).

To further investigate the ability of 4-OT to catalyze the isomerization of **6** to **4**, the enzyme (7.3 µM, monomer concentration) was incubated with **6** (2 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3), and the reaction was monitored by UV spectroscopy (Figure 5.2B). A fast increase in A$_{320}$ and a change in the shape of the absorbance spectrum were observed, indicating the formation of **4**. To affirm that the conversion of **6** to **4** is enzyme-catalyzed, a control reaction was performed without 4-OT, which did not result in an increase in A$_{320}$ (Figure 5.2A). Although the 4-OT sample used in this assay was of high purity (>95% as assessed by SDS-PAGE), the possibility existed that a contaminating enzyme from the expression host, rather than 4-OT, was catalyzing the isomerization reaction. Therefore, a 4-OT sample free of cellular proteins was prepared by total chemical synthesis.[9] Incubation of this synthetic 4-OT (7.3 µM, monomer concentration) with **6** (2 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) gave rise to a similar increase in A$_{320}$ as observed for the incubation of **6** with recombinant 4-OT (Figure 5.2C).
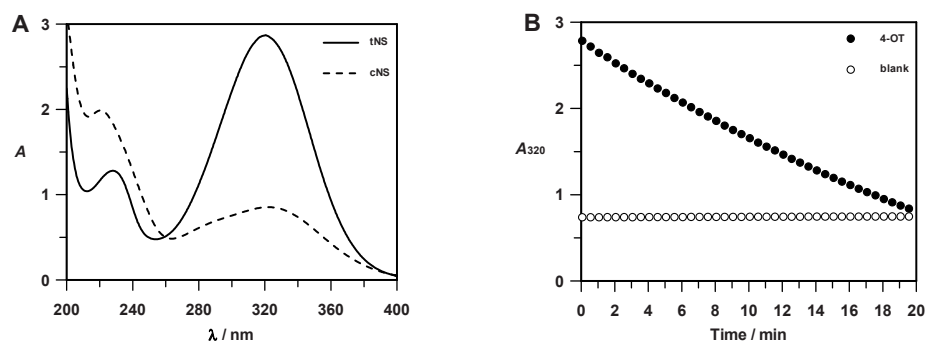
**Figure 5.1.** A) UV absorbance spectra of *cis*-nitrostyrene (**6**, cNS, 2 mM, $\varepsilon_{cis,\mathbf{6}} = 0.31$ mM$^{-1}$ mm$^{-1}$) and *trans*-nitrostyrene (**4**, tNS, 2 mM, $\varepsilon_{trans,\mathbf{4}} = 1.44$ mM$^{-1}$ mm$^{-1}$) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) containing 10% EtOH; B) UV traces monitoring the incubation of *cis*-nitrostyrene (**6**) in the presence of acetaldehyde (**3**) and with 4-OT, or without 4-OT (blank).



**Figure 5.2.** UV absorbance spectra monitoring the isomerization of *cis*-nitrostyrene (**6**) to *trans*-nitrostyrene (**4**) without enzyme (A), or catalyzed by recombinant 4-OT (B), synthetic 4-OT (C) or 4-OT P1A (D).

To verify that **4** is formed during incubation of 4-OT with **6**, the reaction was analyzed by $^1$H NMR spectroscopy. 4-OT ($4.4 \times 10^{-3}$ mM, 0.06 mol%) was incubated with **6** (7 mM) in Na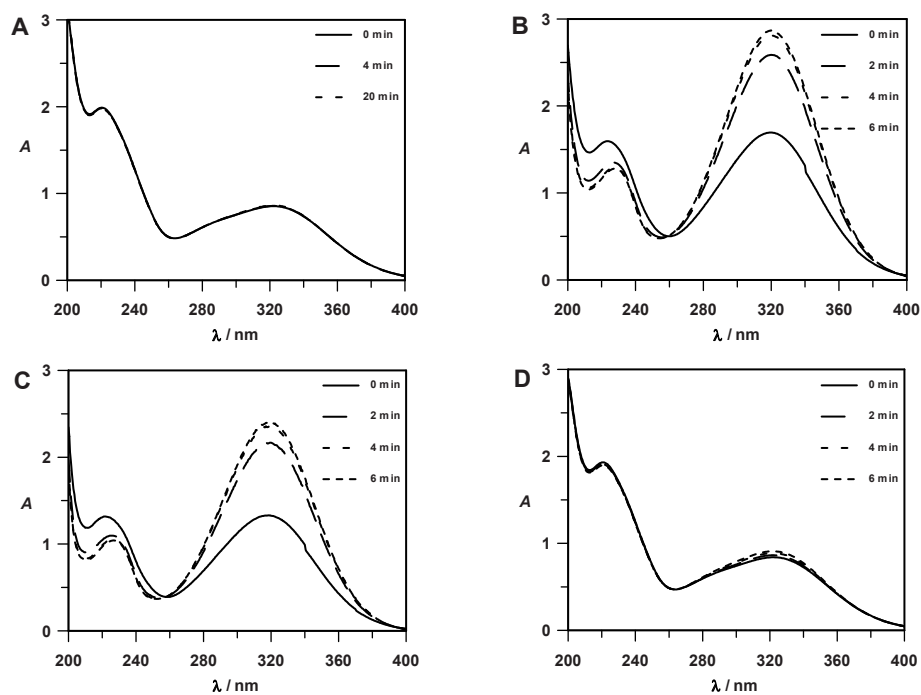D$_2$PO$_4$ buffer (pD 7.6) containing 30% (v/v) deuterated methanol (Figure 5.3). At the start of the reaction, characteristic signals (at 7.04 and 7.15 ppm) corresponding to the vinyl protons of **6** were detected (Figure 5.3A). These signals gradually diminished and signals assigned to the vinyl protons of **4** (at 7.87 and 8.11 ppm) appeared. Furthermore, the signals corresponding to the phenylic protons of **6** (7.40 ppm) were decreasing over time, while signals for the phenylic protons of **4** (at 7.52 and 7.66 ppm) emerged (Figure 5.3B-D). The final spectrum, recorded after completion of the enzymatic conversion of **6** (Figure 5.3D; t = 15 min), is identical to the spectrum of an authentic standard of **4** (Figure 5.3E), confirming that **4** is indeed the product of the 4-OT-catalyzed isomerization of **6**. Integration of the signals indicated that the enzymatic conversion of **6** to **4** is relatively fast, with 44 and 70% conversion after 2 and 5 minutes, respectively, and near full conversion after 15 minutes. A control reaction without 4-OT, but under otherwise identical conditions, showed no conversion of **6**, which indicates that nonenzymatic isomerization of **6** into **4** does not occur under the assay conditions used.
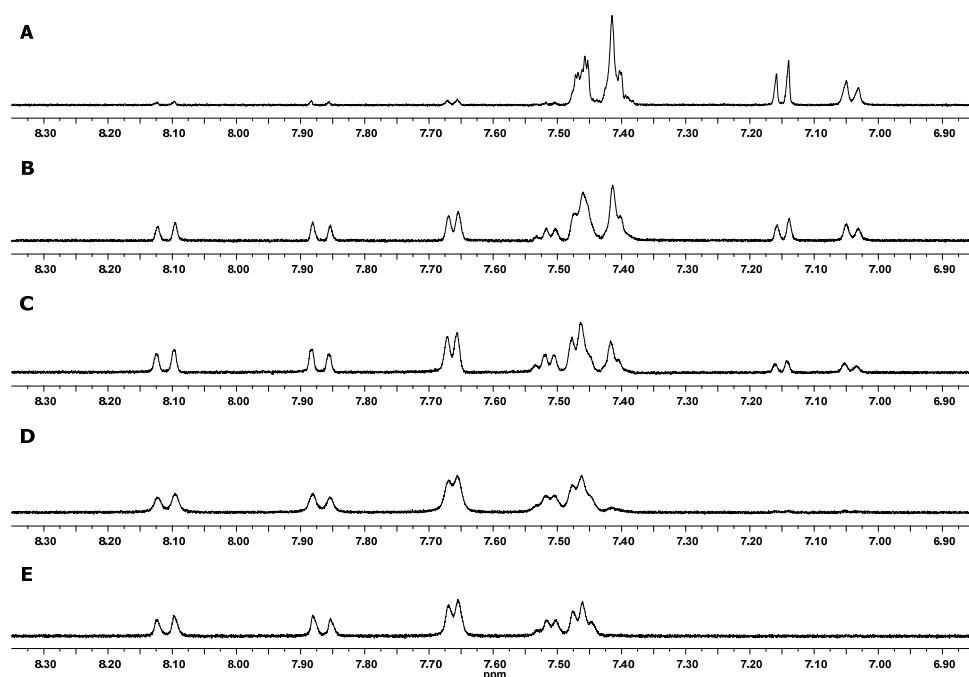


**Figure 5.3.** Stack plot of partial $^1$H NMR spectra monitoring the conversion of **6** to **4** catalyzed by 4-OT. A) t=0 min, B) t=2 min, C) t=5 min, D) t=15 min, E) reference compound **4**.

Having established that **4** is the product of the enzymatic isomerization of **6**, kinetic parameters were determined by using various concentrations of **6** (0.1-6 mM). 4-OT catalyzes the isomerization of **4** to **6** with a $k_{cat}$ of 4.8 s$^{-1}$ and a $K_m$ of 2.5 mM, yielding a $k_{cat}/K_m$ of $1.9 \times 10^3$ M$^{-1}$s$^{-1}$.
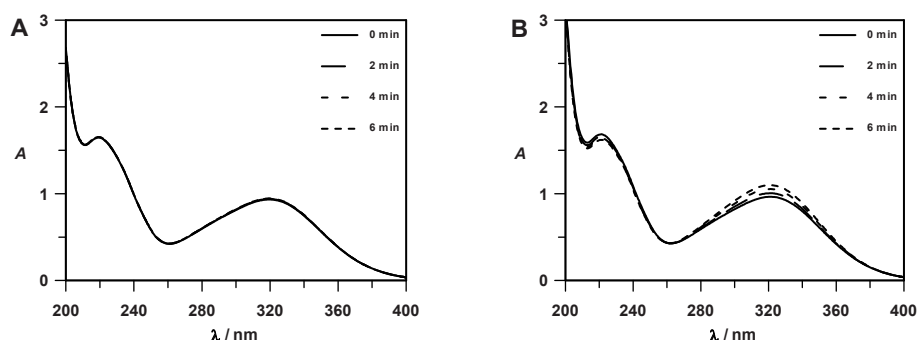


**Figure 5.4.** UV absorbance spectra monitoring the isomerization of *cis*-nitrostyrene (**6**) to *trans*-nitrostyrene (**4**) catalyzed by L-proline (A), or denatured 4-OT (B).

Next, we evaluated the importance of residue Pro-1 for the isomerization activity of 4-OT by mutation of the Pro-1 residue to an alanine. The corresponding 4-OT P1A mutant (7.3 μM, monomer concentration) was incubated with **6** (2 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) and the reaction was monitored by UV spectroscopy (Figure 5.2D). 4-OT P1A almost completely lost its activity, demonstrating that residue Pro-1 is essential for the 4-OT-catalyzed isomerization of **6** to **4**. Residues Arg-11 and Arg-39 are key catalytic residues in the natural tautomerase activity of 4-OT and interact with respectively the C-6 and C-1 carboxylate group of **1**.[10] To test whether one of these active site arginines might be important for binding the nitro-group (or nitronate) of **6**, these arginines were mutated into alanine residues, yielding mutants 4-OT R11A and 4-OT R39A. However, these mutations did not affect the ability of the enzyme to catalyze the isomerization of **6** to **4** (data not shown). This raised the question whether the active site of 4-OT is needed for this reaction, or whether the amino acid L-proline, or a peptide with an amino-terminal proline, would be sufficient for catalysis of this isomerization reaction. Therefore, two control experiments were performed. First, L-proline (7.3 μM) was incubated with **6** (2 mM) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3). UV spectroscopic analysis of this incubation indicated that no **4** was formed (Figure 5.4A). Even with a 100 times higher concentration of L-proline (730 μM) no isomerization of **6** into **4** was observed. Second, the same incubation experiment was performed, but now with a 4-OT sample that had been boiled for 20 minutes. The enzyme denatures upon boiling, which results in an unstructured peptide chain bearing a proline residue at the amino-terminus. The denatured enzyme almost completely lost its activity

(Figure 5.4B). Both control experiments lead to the conclusion that the active site of 4-OT is responsible for the isomerization of **6** to **4**. [11]

Considering the importance of the Pro-1 residue for the isomerization activity of 4-OT, two possible mechanisms can be proposed. In the first possible mechanism, 4-OT catalyzes the conversion of **6** to **4** through a covalent enzyme-substrate intermediate (Scheme 5.4). Residue Pro-1 acts as a nucleophile and attacks the C-2 carbon of **6**, leading to the formation of an *aci*-nitronate intermediate. In this covalent intermediate, formed between **6** and Pro-1, the original C-C double bond is transformed into a C-C single bond. Around this C-1-C-2 single bond, an internal rotation takes place to give a more favorable conformation with less repulsion between the nitro-group and the phenyl ring. Subsequent collapse of the *aci*-nitronate intermediate restores the C-C double bond and releases **4** from Pro-1. Such a covalent mechanism is plausible because Pro-1 has the correct protonation state to act as a nucleophile at pH 7.3, and because mechanisms involving nucleophilic attack on the unsaturated C-C bond have previously been proposed for natural isomerases like maleate isomerase, maleylacetoacetate isomerase, and maleylpyruvate isomerase.[12,13,14] In maleate isomerase the nucleophile is a cysteine residue, whereas in the case of maleylacetoacetate and maleylpyruvate isomerases the essential cofactor glutathione functions as the nucleophile. To find further support for this proposed covalent-intermediate-mechanism, we performed a labeling experiment to analyze whether a covalent intermediate between 4-OT and substrate **6** (or product **4**) can be formed. An aliquot of 4-OT (500 μg) was incubated with 10 mM **6** in 1 mM $NH_4HCO_2$ buffer (pH 7.3) containing 30% (v/v) MeOH and the reaction was stopped after 5 minutes by freezing the sample at -20 °C. ESI-MS analysis revealed two major peaks, one corresponding to unlabeled 4-OT (6811 Da; ~80%) and one consistent with the expected mass for 4-OT labeled with either **6** or **4** (6959 Da; ~20%). Unfortunately, the labeled amino acid residue could not be identified.[15]
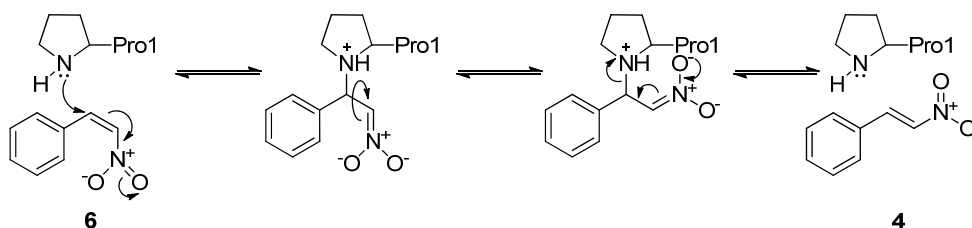


**Scheme 5.4.** A possible mechanism for the 4-OT-catalyzed isomerization of **6** to **4** in which Pro-1 acts as a nucleophile to form a covalent enzyme-substrate intermediate.

In another possible mechanism, residue Pro-1 acts as a general base that activates a water molecule (Scheme 5.5). By addition of the hydroxide ion to the C-C double bond and concomitant electron migration, an *aci*-nitronate intermediate is formed in which the original

unsaturated C-C bond is turned into a saturated C-C bond. Now, the molecule can adapt the more favorable conformation through internal rotation around this C-1-C-2 single bond. Collapse of the *aci*-nitronate intermediate results in restoration of the C-C double bond, and the release of product **4**. The possible role of Pro-1 as water-activating residue is suggested by a mechanism that has been proposed for the promiscuous dehalogenase activity of 4-OT, in which residue Pro-1 may activate a water molecule for attack at the C-3-position of *trans*-3-chloroacrylic acid.[16]



**6**                                                                **4**
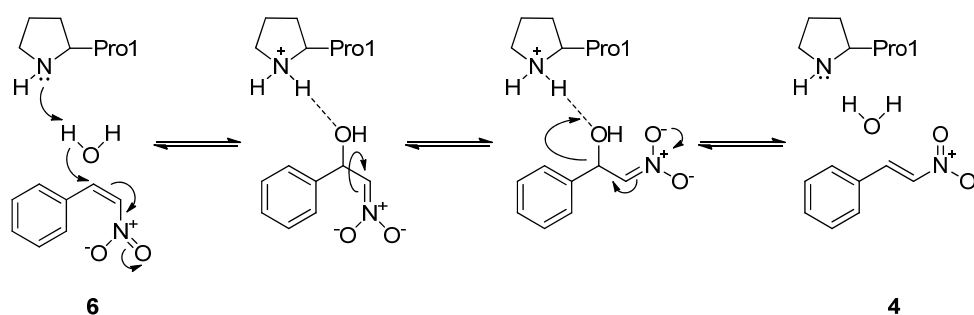
**Scheme 5.5.** A possible mechanism for the 4-OT-catalyzed isomerization of **6** to **4** in which Pro-1 acts as a general base to activate a water molecule.

## Conclusion

In conclusion, we set out to investigate whether 4-OT can catalyze the addition of **3** to **6**, similar to the previously reported promiscuous Michael-type addition of **3** to **4**.[7] However, we found that 4-OT does not catalyze the addition of **3** to **6**, but instead catalyzes the isomerization of **6** into **4**. This adds yet another chemically distinct reaction to the list of (promiscuous) reactions catalyzed by 4-OT. Apart from its natural tautomerase activity, 4-OT catalyzes the hydrolytic dehalogenation of *trans*-3-chloroacrylic acid, which is the natural activity of another member of the tautomerase superfamily, *trans*-3-chloroacrylic acid dehalogenase (CaaD).[1,16,17] 4-OT has also promiscuous activities that are not known to occur as natural activities within the tautomerase superfamily, including aldolase, dehydratase and Michael-type addition activities.[5,6,7] Furthermore, when Pro-1 in 4-OT is mutated to an alanine residue (4-OT P1A), the mutant enzyme is capable of the decarboxylation of oxaloacetate.[18] However, so far, the isomerization activity ($k_{cat} = 4.8$ s$^{-1}$; $k_{cat}/K_m$ of $1.9 \times 10^3$ M$^{-1}$s$^{-1}$) appears to be the highest promiscuous activity of 4-OT. We demonstrated that the Pro-1 residue of 4-OT is essential for this isomerization activity, but that the organocatalyst L-proline is not able to catalyze the *cis-trans* isomerization of nitrostyrene. This might be explained, at least in part, by the higher $pK_a$ value of the (secondary) amino group of L-proline as compared to the $pK_a$ of Pro-1 in 4-OT. Under the assay conditions (pH 7.3), the amino group of L-proline ($pK_a$ ~10.6) exists predominantly in the protonated form. This renders L-proline unable to act as a nucleophile or base, which might explain its inability to catalyze the isomerization reaction of **6** into **4**. In 4-OT, however, the $pK_a$ (~6.4) of Pro-1 is lowered as an effect of the surrounding active site environment of low dielectric constant.[4,19] This enables Pro-1 to function either as a nucleophile, or as a general base according to one of the proposed mechanisms. Future mechanistic and structural studies might shed light on the exact mechanism of this new promiscuous activity of 4-OT.

# Experimental Procedures

*Materials:* All chemicals were obtained from Sigma-Aldrich Chemical Co. (St. Louis, MO), unless stated otherwise. The sources for the biochemicals, buffers, solvents, components of Luria-Bertani (LB) media as well as the materials, enzymes, and reagents used in the molecular biology procedures are reported elsewhere.[20] High-purity synthetic 4-OT (lyophilized, with Met-45 replaced by Nle to prevent oxidation upon sample handling) was obtained from GenScript USA Inc. (Piscataway, NY) and folded into the active homohexamer as described before.[21]

*General methods:* Techniques for restriction enzyme digestions, ligation, transformation, and other standard molecular biology manipulations were based on methods described elsewhere.[22] The PCR was carried out in a DNA thermal cycler (model GS-1) obtained from Biolegio (Nijmegen, The Netherlands). DNA sequencing was performed by Macrogen (Seoul, Korea). Protein was analyzed by polyacrylamide gel electrophoresis (PAGE) using sodium dodecyl sulfate (SDS) gels containing polyacrylamide (10%). The gels were stained with Coomassie brilliant blue. Protein concentrations were determined by the Waddell method.[23] Kinetic data were obtained on a V-650 or V-660 spectrophotometer from Jasco (IJsselstein, The Netherlands) and were fitted by nonlinear regression data analysis using the GraFit program (Erithacus, Software Ltd., Horley, U.K.) obtained from Sigma Chemical Co. $^1$H NMR spectra were recorded on a Varian Inova 500 (500 MHz) spectrometer. Chemical shifts for protons are reported in parts per million scale ($\delta$ scale) and are referenced to CHCl$_3$ ($\delta$ = 7.26) or H$_2$O ($\delta$ = 4.80). The masses of 4-OT and 4-OT mutants were determined by ESI-MS using a Sciex API 3000 triple quadrupole mass spectrometer (AB Sciex, Concord, Ontario, Canada) which is housed in the Mass Spectrometry Facility Core in the Department of Pharmacy at the University of Groningen.

*Expression and purification of wild-type 4-OT and mutants:* 4-OT and 4-OT mutants were expressed in *E. coli* BL21(DE3) cells harboring the appropriate expression vector (based on the pET20b(+) plasmid). The construction of the vectors and the expression and purification of 4-OT and 4-OT mutants have been reported before.[5]

*Preparation of cis-nitrostyrene (6):* *Cis*-nitrostyrene (**6**) was prepared according to an adapted literature procedure.[24,25] *Trans*-nitrostyrene (**4**) (150 mg, 1.0 mmol) was dissolved in cyclohexane (200 mL). The solution was divided over 10 glass vials, which were placed between two UV lamps ($\lambda$=350 nm; Universal UV lamp TL-900 50/60 Hz, CAMAG, Muttenz, Switzerland). The set-up was covered to prevent the entrance of stray light. Upon irradiation, **4** isomerized to **6**. The progress of the reaction was followed by TLC (EtOAc: Hexane 1:5). After 48 hours, a stable 50/50 ratio was reached. The reaction mixtures were combined and cyclohexane was evaporated *in vacuo* at low temperature to prevent thermal isomerization. Compounds **6** and **4** were separated by column chromatography (silica gel, EtOAc/*n*-hexane 1/15) and **6** was obtained as a yellow oil (47 mg, 0.32 mmol, 32%). Spectroscopical data are in accordance with the literature.[25,26,27]

^{1}H NMR (500 MHz, CDCl$_3$, 25 °C); *trans*-nitrostyrene **4**: δ 8.02 (d, $J$ = 14.0 Hz, 1H), 7.59 (d, $J$ = 14.0 Hz, 1H), 7.57-7.44 (m, 5H); ^{1}H NMR (500 MHz, CDCl$_3$, 25 °C); *cis*-nitrostyrene **6**: δ 7.52-7.40 (m, 5H), 6.97 (d, $J$ = 9.5 Hz, 1H), 6.79 (d, $J$ = 9.5 Hz, 1H).

*General UV spectroscopic assay for the isomerization of cis-nitrostyrene (**6**) to trans-nitrostyrene (**4**):* The 4-OT-catalyzed isomerization of **6** to **4** was monitored by following the change in absorbance at 320 nm (ΔA$_{320}$). Since both **6** and **4** absorb at 320 nm ($\varepsilon_{cis}$ = 0.31 mM$^{-1}$ mm$^{-1}$ and $\varepsilon_{trans}$ = 1.44 mM$^{-1}$ mm$^{-1}$), ΔA$_{320}$ is the sum of the depletion of **6** and the simultaneous formation of **4**. In a 1 mm cuvette, 4-OT (WT or mutant) (15 μg, 7.3 μM, monomer concentration) was incubated with **6** (2 mM, 30 μL from a 20 mM stock solution in absolute ethanol) in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3; final volume 0.3 mL). Absorbance spectra were recorded from 200 to 400 nm.

*Kinetic assays:* The kinetic assays were performed at 22 °C by following the formation of **4** at 320 nm. A stock solution of 4-OT (5.5 μg/mL; 0.81 μM, monomer concentration) was made in 20 mM NaH$_2$PO$_4$ buffer (pH 7.3) and equilibrated at 22 °C. Stock solutions (1-60 mM) of **6** were prepared in absolute ethanol. To assay the enzyme activity, an aliquot of the enzyme solution (270 μL) was mixed with 30 μL of the appropriate stock solution of **6** in a 1 mm cuvette, yielding final concentrations of 0.73 μM 4-OT and 0.1 to 6 mM **6**. Compound **6** was contaminated with ~6% of **4**, and contained ~10% water; the concentration of **6** was corrected accordingly. The initial rates (mM/s) were plotted against the corrected substrate concentration (mM). The data were fitted to the Michaelis-Menten equation using GraFit and the kinetic parameters $k_{cat}$ and $K_m$ were determined.

*Preparation of NaD$_2$PO$_4$ buffer:* The preparation of NaD$_2$PO$_4$ buffer was described before.[6]

*^{1}H NMR spectroscopic analysis of the isomerization reaction:* Deuterated methanol (methanol-d4, 130 μL) was pre-mixed with NaD$_2$PO$_4$ buffer (pD 7.6, 437 μL) after which *cis*-nitrostyrene (**6**) was added (65 μL of a stock solution (70 mM) in methanol-d4). The isomerization reaction was started by the addition of 4-OT (30 ng/μL, 4.4 × 10$^{-3}$ mM, 0.06 mol%). The final reaction volume was 650 μL containing 30% (v/v) methanol-d4 and 7 mM **6**. ^{1}H NMR spectra were recorded after 2, 5 and 15 min. A control reaction without 4-OT, but under otherwise identical conditions, was performed as well.
^{1}H NMR (500 MHz, NaD$_2$PO$_4$/MeOD 70/30 v/v, 25 °C); *trans*-nitrostyrene **4**: δ 8.11 (d, $J$ = 13.5 Hz, 1H), 7.87 (d, $J$ = 13.5 Hz, 1H), 7.66 (d, $J$ = 7.5 Hz, 2H), 7.52 (t, $J$ = 7.5 Hz, 1H), 7.46 (t, $J$ = 7.5 Hz, 2H). ^{1}H NMR (500 MHz, NaD$_2$PO$_4$/MeOD 70/30 v/v, 25 °C); *cis*-nitrostyrene **6**: δ 7.47-7.40 (m, 5H), 7.15 (d, $J$ = 9.0 Hz, 1H), 7.04 (d, $J$ = 9.0 Hz, 1H).

*Labeling of 4-OT with cis-nitrostyrene (**6**):* At 22 °C, 4-OT (500 μg) was incubated with 10 mM **6** (25 μL from a 200 mM stock solution in methanol) in a final volume of 0.5 mL NH$_4$HCO$_2$ buffer (1 mM; pH 7.3) containing 30% (v/v) methanol. The reaction was stopped after 5 minutes by freezing the sample (-20 °C). The sample was analyzed by ESI-MS without further treatment.

## Acknowledgements

## References

1. a) C. P. Whitman, *Arch. Biochem. Biophys*. **2002**, *402*, 1-13; b) G. J. Poelarends, V. Puthan Veetil, C. P. Whitman, *Cell. Mol. Life Sci*. **2008**, *65*, 3606-3618.

2. a) S. Harayama, M. Rekik, K. L. Ngai, L. N. Ornston, *J. Bacteriol*. **1989**, *171*, 6251-6258; b) C. P. Whitman, B. A. Aird, W. R. Gillespie, N. J. Stolowich, *J. Am. Chem. Soc*. **1991**, *113*, 3154-3162; c) L. H. Chen, G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour, C. P. Whitman, *J. Biol. Chem*. **1992**, *267*, 17716-17721.

3. a) S. C. Wang, W. H. Johnson, Jr., R. M. Czerwinski, S. L. Stamps, C. P. Whitman, *Biochemistry* **2007**, *46*, 11919-11929; b) E. A. Burks, C. D. Fleming, A. D. Mesecar, C. P. Whitman, S. D. Pegan, *Biochemistry* **2010**, *49*, 5016-5027.

4. J. T. Stivers, C. Abeygunawardana, A. S. Mildvan, G. Hajipour, C. P. Whitman, L. H. Chen, *Biochemistry* **1996**, *35*, 814-823.

5. E. Zandvoort, B.-J. Baas, W. J. Quax, G. J. Poelarends, *ChemBioChem* **2011**, *12*, 602-609.

6. E. Zandvoort, E. M. Geertsema, W. J. Quax, G. J. Poelarends, *ChemBioChem* **2012**, *13*, 1274-1277.

7. E. Zandvoort, E. M. Geertsema, B.-J. Baas, W. J. Quax, G. J. Poelarends, *Angew. Chem*. **2012**, *124*, 1266-1269; *Angew. Chem. Int. Ed*. **2012**, *51*, 1240-1243.

8. For reviews on organocatalysis see a) W. Notz, F. Tanaka, C. F. Barbas III, *Acc. Chem. Res*. **2004**, *37*, 580-591; b) A. Erkkilä, I. Majander, P. M. Pihko, *Chem. Rev*. **2007**, *107*, 5416-5470; c) S. Mukherjee, J. W. Yang, S. Hoffmann, B. List, *Chem. Rev*. **2007**, *107*, 5471-5569; d) L.-W. Xu, Y. Lu, *Org. Biomol. Chem*. **2008**, *6*, 2047-2053; e) L.-W. Xu, J. Luo, Y. Lu, *Chem. Commun*. **2009**, 1807-1821; f) B. List, *Angew. Chem*. **2010**, *122*, 1774-1779; *Angew. Chem. Int. Ed*. **2010**, *49*, 1730-1734.

9. 4-OT was synthesized by GenScript USA Inc. (Piscataway, NY).

10. T. K. Harris, R. M. Czerwinski, W. H. Johnson, Jr., P. M. Legler, C. Abeygunawardana, M. A. Massiah, J. T. Stivers, C. P. Whitman, A. S. Mildvan, *Biochemistry* **1999**, *38*, 12343-12357.

11. It is important to note that similar enzymatic assays were performed in which 4-OT is replaced by either a proline-based tautomerase (YwhB or MIF) or by an aspartase (AspB or MAL) that lacks an amino-terminal proline. Both MIF and YwhB are catalyzing the isomerization of **6** to **4** at a similar rate as 4-OT WT, while AspB and MAL do not show any isomerase activity.

12. F. Fisch, C. Martinez Fleites, M. Delenne, N. Baudendistel, B. Hauer, J. P. Turkenburg, S. Hart, N. C. Bruce, G. Grogan, *J. Am. Chem. Soc.* **2010**, *132*, 11455-11457.

13. G. Polekhina, P. G. Board, A. C. Blackburn, M. W. Parker, *Biochemistry* **2001**, *40*, 1567-1576.

14. M. Marsh, D. K. Shoemark, A. Jacob, C. Robinson, B. Cahill, N.-Y. Zhou, P. A. Williams, A. T. Hadfield, *J. Mol. Biol.* **2008**, *384*, 165-177.

15. To determine the site of labeling, a sample of 4-OT partially labeled with **6** was digested by using trypsin and the peptide mixture was analyzed by MALDI-TOF mass spectrometry. No modified peptides carrying the extra mass corresponding to **6** were detected, indicating that the label was lost under the MALDI-TOF conditions. Attempts to trap the covalent intermediate by reduction with $NaCNBH_3$ were also unsuccessful.

16. S. C. Wang, W. H. Johnson, Jr., C. P. Whitman, *J. Am. Chem. Soc.* **2003**, *125*, 14282-14283.

17. G. J. Poelarends, J. L. Almrud, H. Serrano, J. E. Darty, W. H. Johnson, Jr., M. L. Hackert, C. P. Whitman, *Biochemistry* **2006**, *45*, 7700-7708.

18. A. Brik, L. J. D'Souza, E. Keinan, F. Grynszpan, P. E. Dawson, *ChemBioChem* **2002**, *3*, 845-851.

19. R. M. Czerwinski, T. K. Harris, W. H. Johnson, Jr., P. M. Legler, J. T. Stivers, A. S. Mildvan, C. P. Whitman, *Biochemistry* **1999**, *38*, 12358-12366.

20. H. Raj, B. Weiner, V. Puthan Veetil, C. R. Reis, W. J. Quax, D. B. Janssen, B. L. Feringa, G. J. Poelarends, *ChemBioChem* **2009**, *10*, 2236-2245.

21. M. C. Fitzgerald, I. Chernushevich, K. G. Standing, S. B. H. Kent, C. P. Whitman, *J. Am. Chem. Soc.* **1995**, *117*, 11075-11080.

22. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, **1989**.

23. W. J. Waddel, *J. Lab. Clin. Med.* **1956**, *48*, 311-314.

24. T. Ohwada, T. Ohta, K. Shudo, *J. Am. Chem. Soc.* **1986**, *108*, 3029-3032.

25. D. B. Miller, P. W. Flanagan, H. Shechter, *J. Org. Chem.* **1976**, *41*, 2112-2120.

26. M. Vuagnoux-D'Augustin, A. Alexakis, *Eur. J. Org. Chem.* **2007**, *35*, 5852-5860.

27. J. A. Burkhard, B. H. Tchitchanov, E. M. Carreira, *Angew. Chem.* **2011**, *123*, 5491-5494; *Angew. Chem. Int. Ed.* **2011**, *50*, 5379-5382.

# *6*    Summary and Future Perspectives

**Catalytic promiscuity as starting point for the development of new biocatalysts**

Enzymes are protein catalysts that catalyze a wide range of transformations with high chemical and stereospecificity. Furthermore, enzymes work under mild conditions, performing conversions in an aqueous environment and at moderate temperature. This specificity and the ability to work under mild reaction conditions make enzymes attractive catalysts for industrial applications. A drawback, however, is that enzymes in general only accept substrates and catalyze reactions for which they evolved, so application of enzymes for non-natural, man-designed chemical reactions is limited. It is a challenging task to broaden the industrial applicability of enzymes by re-designing enzymes to obtain new unnatural activities. How can this task be approached?

Nowadays, the notion that enzymes are highly specific catalysts performing only one chemical transformation is superseded. More and more examples are found of enzymes that catalyze more than one chemically distinct reaction. This phenomenon, called catalytic promiscuity, occurs also in the tautomerase superfamily of enzymes, a group of homologous proteins that share a characteristic β-α-β-fold and a unique catalytic amino-terminal proline. **Chapter 1** presents an overview of catalytic promiscuity within the tautomerase superfamily and the insights it provides concerning the divergent evolution of enzymes, enzyme mechanisms, and the chemical versatility of an enzyme's active site. In turn, this knowledge can be used to predict new unnatural activities for the same enzyme scaffold; in other words, catalytic promiscuity can be exploited as a springboard for the generation of new biocatalysts.

**Promiscuous carbon-carbon bond-forming activities in 4-oxalocrotonate tautomerase**

Carbon-carbon bond formation is of utmost importance for the synthesis of new compounds, allowing the expansion of the molecule's carbon skeleton. In organic synthesis a wide range of carbon-carbon bond-forming reactions including aldol coupling, Michael addition and Mannich reactions are applied, and organocatalysis has become one of the main areas in the asymmetric catalysis of these reactions. Of particular interest is aminocatalysis, in which secondary and primary amine catalysts are employed that react with carbonyl compounds to give enamine and iminium ion intermediates. For example, the amino acid proline and derivatives thereof are used as efficient (secondary amine) organocatalysts.

The enzyme 4-oxalocrotonate tautomerase (4-OT) is a member of the tautomerase superfamily and as such it possesses a characteristic amino-terminal proline. Pro-1 acts as a general base in 4-OT's natural activity, that is, the tautomerization of 2-hydroxyhexa-2,4-dienedioate into 2-oxohex-3-enedioate. However, Pro-1 provides a unique chemical

functionality in the active site that might also be suitable for carbon-carbon bond formation through an enamine intermediate, in reminiscence of organocatalysis. Therefore, the work described in this thesis focused on the screening of 4-OT for its ability to catalyze carbon-carbon bond-generating reactions, and on the characterization of these promiscuous non-natural activities.

We used a systematic screening strategy for the discovery of new promiscuous activities in 4-OT (**Chapter 2**). In this screening method reactive carbonyl compounds were identified that give rise to imine/enamine intermediates and it was demonstrated that these imines/enamines are formed with Pro-1 by proteinase treatment followed by LC-MS/MS. The chemical potential of the smallest enamine (generated from acetaldehyde) was further explored by combination with other carbonyl compounds, which resulted in the identification of an aldolase activity in 4-OT, in which acetaldehyde is coupled to benzaldehyde to yield cinnamaldehyde. The active site residues Pro-1 and Arg-11, one of the catalytic arginine residues that is important for the tautomerase activity of 4-OT, seem to be involved in the aldol reaction. In addition, it was shown that this low-level activity can be improved (16-fold in $k_{cat}/K_m$) by a single amino acid substitution (L8R). Based on these observations, we proposed a mechanism that might explain this aldolase activity.

To further improve the aldolase activity of 4-OT, we used a mechanism-based engineering approach (**Chapter 3**). It was hypothesized that hydrolysis of the product from the catalytic proline residue is the rate-limiting step in the overall aldol condensation reaction. In order to raise the reaction rate we employed a 4-OT variant (F50A) that is reported to have an active site more accessible to water, which might facilitate product release. Indeed, 4-OT F50A displayed a largely increased aldolase activity (600-fold in $k_{cat}/K_m$). Presumably, the aldol condensation reaction consists of two parts: the initial cross-coupling of acetaldehyde with benzaldehyde yields the aldol product 3-hydroxy-3-phenylpropanal, which then undergoes dehydration yielding the final product cinnamaldehyde. During the characterization of the 4-OT-catalyzed aldol condensation reaction by $^1$H NMR spectroscopy, aldol product 3-hydroxy-3-phenylpropanal was not observed. This raised the question whether the dehydration reaction is 4-OT-catalyzed, or whether the aldol product is highly unstable and undergoes rapid chemical dehydration. Incubation of both wild-type 4-OT and 4-OT F50A with 3-hydroxy-3-phenylpropanal demonstrated that the enzymes accept this aldol compound as substrate and catalyze its dehydration reaction (to yield cinnamaldehyde) and its retro-aldol reaction (to give acetaldehyde and benzaldehyde) at the same time.

Another fundamental reaction to expand the carbon skeleton of a molecule is the Michael addition reaction, in which a carbon nucleophile reacts with an α,β-unsaturated carbonyl compound. Again, the smallest enamine (generated from acetaldehyde) was

screened in combination with various Michael acceptors. Unfortunately, we did not observe any standard Michael addition activity with α,β-unsaturated carbonyl compounds, but we found that 4-OT catalyzes the Michael-type addition of acetaldehyde to *trans*-β-nitrostyrene (**Chapter 4**). The product 4-nitro-3-phenylbutanal is an important precursor for the antidepressant phenibut (4-amino-3-phenylbutanoic acid) and is obtained as the (*S*)-enantiomer in good enantiomeric excess (89% *ee*). For this reaction, 4-OT does not only need Pro-1 and Arg-11 residues, like in the aldolase activity, but also the other active site arginine, Arg-39, seems to be important. This led us to propose a mechanism in which Pro-1 acts as a nucleophile, Arg-11 ensures correct substrate binding by coordinating the nitro functionality of *trans*-β-nitrostyrene, and Arg-39 acts as a general acid catalyst by delivering a proton at C-α of *trans*-β-nitrostyrene.

### An outsider: a promiscuous non-carbon-carbon bond-forming activity in 4-OT

The 4-OT-catalyzed Michael-type addition of acetaldehyde to *trans*-β-nitrostyrene (**Chapter 4**) yields (*S*)-4-nitro-3-phenylbutanal as the product. In an attempt to enzymatically produce the (*R*)-enantiomer of 4-nitro-3-phenylbutanal, we examined whether *cis*-β-nitrostyrene is accepted as substrate by 4-OT for the addition of acetaldehyde. However, we found that 4-OT does not perform the addition of acetaldehyde to *cis*-β-nitrostyrene, but instead catalyzes the isomerization of *cis*-β-nitrostyrene to *trans*-β-nitrostyrene (**Chapter 5**). Pro-1 seems to be the most important catalytic residue in this transformation, and we demonstrated that a folded enzyme is important for activity. Two mechanisms were proposed for this 4-OT-catalyzed *cis-trans*-isomerization reaction, in which Pro-1 functions either as a nucleophile or as a general base, but so far we were unable to point out the exact function of Pro-1. Within this PhD project, the promiscuous isomerase activity of 4-OT is somewhat an outsider, since it does not involve the formation of a carbon-carbon bond.

### General remarks and future perspectives

We set out to discover new promiscuous carbon-carbon bond-forming activities in 4-OT. The first activity we found was an aldolase activity, in which acetaldehyde is coupled to benzaldehyde. Although the product cinnamaldehyde is not of synthetic interest, this finding provided 'proof of principle' that you can screen for promiscuous activities based on the predicted chemical potential of an enzyme's active site. Indeed, upon further screening, we did identify another promiscuous activity in 4-OT in which a carbon-carbon bond is generated: the asymmetric Michael-type addition of acetaldehyde to *trans*-β-nitrostyrene. Here, the product is of more synthetic relevance, since it is a precursor in the synthesis of the antidepressant phenibut. Furthermore, this enzymatic Michael-type addition complements the reported organocatalytic approach; the product is obtained with similar *ee*
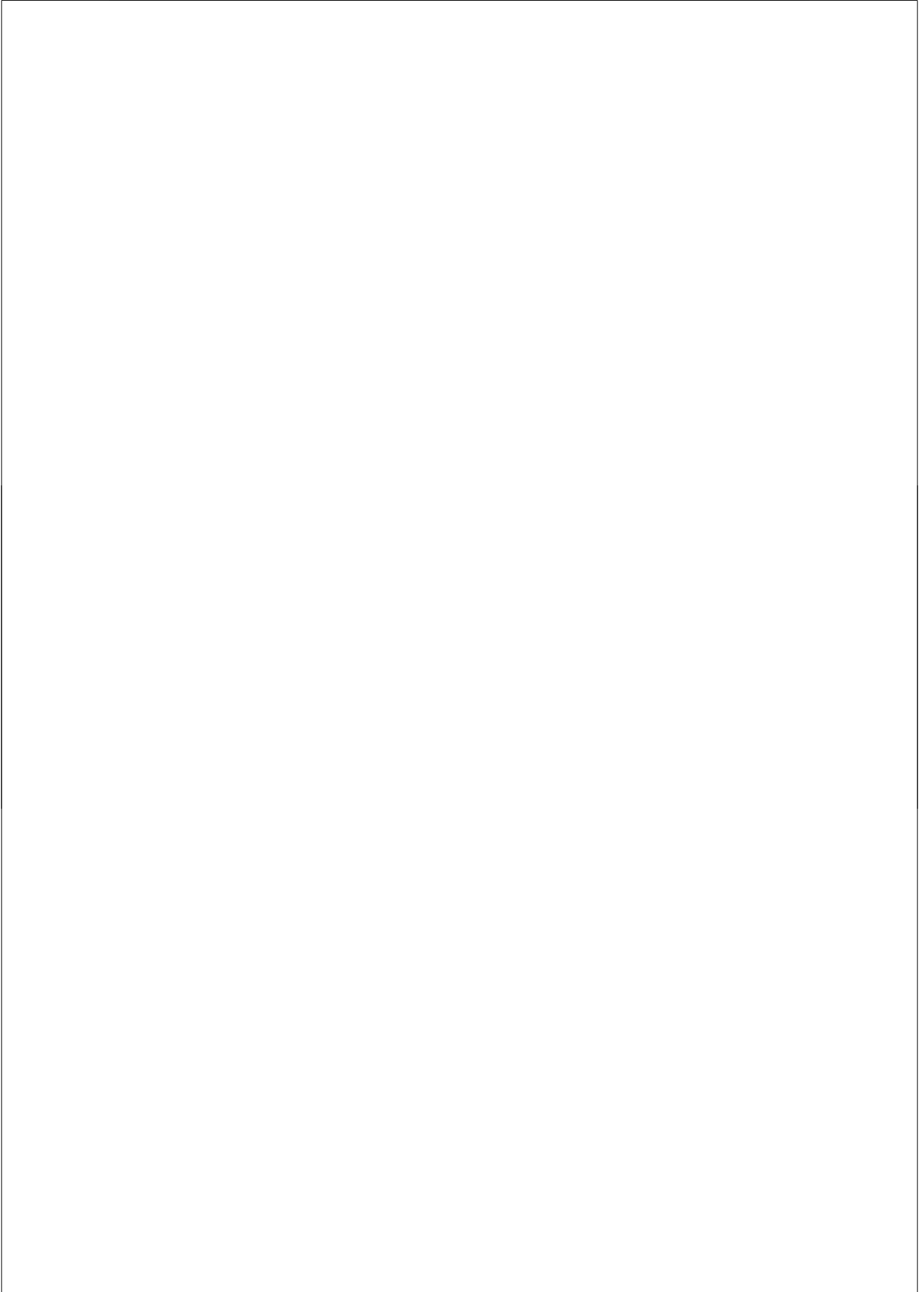
using less catalyst (0.7 versus 10-20 mol%) in a benign aqueous environment (versus organic solvents).

Overall, the results presented in this thesis serve as a good starting point for further development of 4-OT into a versatile and efficient biocatalyst, with improved activities and a broadened substrate and reaction scope. The fact that 4-OT's aldolase activity can be largely increased (600-fold) by only a single amino acid substitution (F50A) is an auspicious indication for the usefulness of directed evolution experiments with the aim to evolve variants with elevated activities. Since 4-OT also catalyzes the dehydration of the initial aldol product 3-hydroxy-3-phenylpropanal, the aldolase activity yields cinnamaldehyde, which is a non-interesting product from a synthetic point of view. We also found that 4-OT catalyzes the self-condensation of acetaldehyde and the self-condensation of propanal, resulting again in dehydrated products. To increase the synthetic usefulness of 4-OT in aldol reactions, it is interesting to design 4-OT mutants that retain high aldolase activity, but lack dehydration activity. It is worthwhile to note that preliminary results indicate that 4-OT catalyzes the aldol reaction of propanal and pyruvate, yielding 2-hydroxy-2,3-dimethyl-4-oxobutanoic acid as non-dehydrated product.

To broaden the substrate scope of the Michael-type reaction, it is intriguing to test whether 4-OT catalyzes the addition of acetaldehyde to the aliphatic Michael acceptor 4-methyl-1-nitropent-1-ene. The presumed product, 5-methyl-3-(nitromethyl)hexanal, is a precursor for pregabalin, an anticonvulsant drug. In addition, the use of longer aldehydes to replace acetaldehyde might further expand the product scope. The initial experiments that we performed along these lines are very promising.

To expand the reaction scope of 4-OT, it is fascinating to look into the possibilities of a 4-OT-catalyzed Mannich reaction. In a Mannich reaction a primary or secondary amine reacts with a carbonyl compound to form an imine. This imine is attacked by a carbon nucleophile, generating a new carbon-carbon bond. The Mannich reaction is a popular reaction in the synthesis of pharmaceuticals and natural products because of the ubiquitous presence of nitrogen in these compounds. The simplest set-up would be to use ethylamine or benzylamine in combination with an excess of acetaldehyde. Acetaldehyde would play a double role by not only reacting with the amine to yield an imine, but also forming a reactive enamine intermediate on 4-OT's Pro-1 residue, which can act as carbon nucleophile.

If successful, these efforts might lead to a set of biocatalysts, based on the 4-OT scaffold, that catalyze diverse asymmetric carbon-carbon bond-forming reactions and that might be industrially applicable in the future.

Nederlandse Samenvatting
voor de Geïnteresseerde Leek

**Enzymen als biokatalysatoren**

Eiwitten zijn opgebouwd uit aminozuren die via peptidebindingen aan elkaar gekoppeld zijn (Figuur 1). De gevormde peptideketen vouwt zich tot een driedimensionale structuur. Een enzym is een eiwit dat een bepaalde chemische reactie kan versnellen (katalyseren). Het katalyseert de omzetting van moleculen aan het begin van het proces (substraten) in andere moleculen (producten), maar wordt zelf niet verbruikt in de reactie (Schema 1). Katalyse vindt plaats op een specifieke plek in het enzym, de zogenaamde 'active site'. De aminozuren in de 'active site' binden de substraten en brengen ze in een oriëntatie die de chemische reactie vergemakkelijkt. De eigenschappen van de 'active site' bepalen of een substraat wel of niet gebonden wordt en maakt een enzym selectief, zowel voor specifieke substraten als voor het versnellen van specifieke chemische reacties.
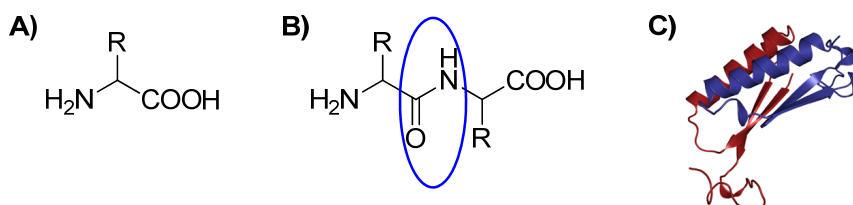


**Figuur 1.** A) Algemene structuur van een aminozuur. ___ = een binding; H=waterstof; N= stikstof; O= zuurstof; R= variabele groep; de rest van de moleculen bestaat uit koolstof; B) Twee aminozuren verbonden door een peptidebinding (omcirkeld); C) Een peptideketen vouwt in een driedimensionale structuur.
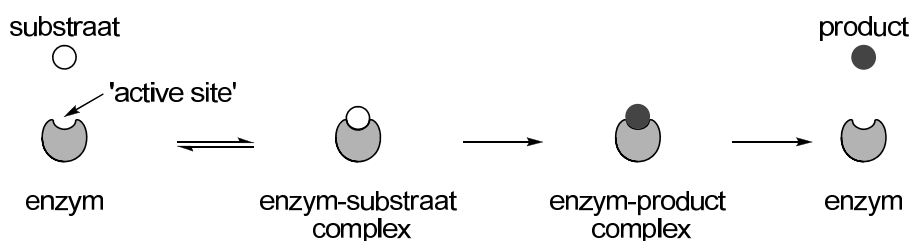


**Schema 1.** Een schematische representatie van een enzym-gekatalyseerde reactie.

Enzymen worden op grote schaal toegepast als biokatalysatoren in de industrie. De specificiteit van enzymen maakt het mogelijk producten te verkrijgen die met conventionele chemische technieken moeilijk te maken zijn. Bovendien werken enzymen onder milde, milieuvriendelijke condities: in water en bij kamertemperatuur. Er zijn heel veel verschillende enzymen die een breed scala aan reacties katalyseren, maar dit zijn reacties die in de natuur voorkomen. Het is moeilijk om enzymen te vinden die chemische reacties

katalyseren die niet in de natuur voorkomen, maar door de mens ontworpen zijn. Dit vormt een beperking voor de industriële toepassing van enzymen. Het is een uitdaging om de industriële toepasbaarheid van enzymen te verbreden door bestaande enzymen aan te passen.

## 'Katalytische promiscuïteit' als uitgangspunt voor de ontwikkeling van nieuwe biokatalysatoren

Van oudsher worden enzymen gezien als specifiek voor één chemische omzetting. Echter, meer en meer enzymen worden ontdekt die naast hun natuurlijke activiteit een andere chemische reactie katalyseren. Dit verschijnsel noemen we 'katalytische promiscuïteit' en het komt o.a. voor in de tautomerase-superfamilie van enzymen. Deze enzymen worden allemaal gevormd door een peptideketen die begint met het aminozuur proline (Pro-1) en die zich vouwt in een karakteristieke driedimensionale structuur ('β-α-β-fold'). Het proline-residu bevindt zich in de 'active site' en speelt een rol in de katalyse van reacties door de enzymen van de tautomerase-superfamilie.

In **Hoofdstuk 1** wordt een overzicht gegeven van katalytische promiscuïteit binnen de tautomerase-superfamilie. Het bestuderen van katalytische promiscuïteit geeft inzicht in de evolutie van enzymen en de mechanismes die een enzym gebruikt om een reactie te versnellen. Ook geeft katalytische promiscuïteit informatie over de chemische veelzijdigheid van de 'active site' van enzymen. Met andere woorden, eenzelfde aminozuur (bijvoorbeeld Pro-1) kan verschillende rollen vervullen in katalyse, afhankelijk van de reactieomstandigheden. Op basis van deze informatie kan men vervolgens proberen te voorspellen welke nieuwe niet-natuurlijke chemische reacties een enzym zou kunnen katalyseren: katalytische promiscuïteit vormt zo een springplank voor de ontwikkeling van nieuwe biokatalysatoren.

## Koolstof-koolstofbindingvormende reacties

Veel moleculen zijn grotendeels opgebouwd uit koolstof (Figuur 2). Daarom is het vormen van een koolstof-koolstofbinding (C-C) erg belangrijk voor de synthese van nieuwe stoffen. Er is een reeks aan chemische reacties waarin een C-C-binding gevormd wordt, waaronder aldolreacties, Michael-addities en Mannich-reacties (Schema 2). Voor het versnellen van deze reacties wordt o.a. gebruik gemaakt van organokatalyse. In organokatalyse wordt geen enzym gebruikt om chemische reacties te versnellen, maar maakt men gebruik van kleine organische moleculen. Vooral interessant is aminokatalyse, een veld binnen de organokatalyse waarin voornamelijk het aminozuur proline en afgeleiden hiervan gebruikt worden als katalysator.
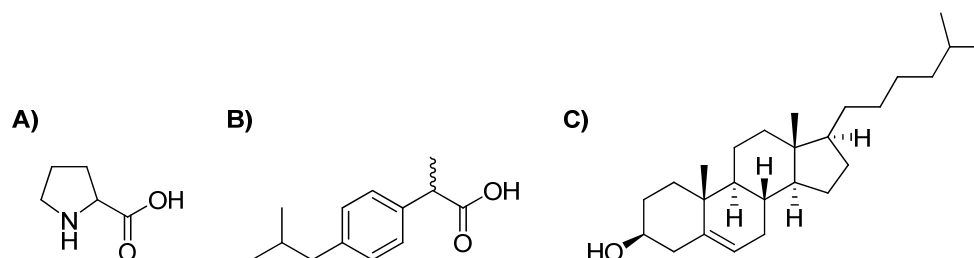
**Figuur 2.** Enkele voorbeelden van (organische) moleculen die grotendeels zijn opgebouwd uit koolstofatomen. A) het aminozuur proline; B) Ibuprofen; C) Cholesterol. ___ = een binding; H=waterstof; N= stikstof; O= zuurstof; de rest van de moleculen bestaat uit koolstof.
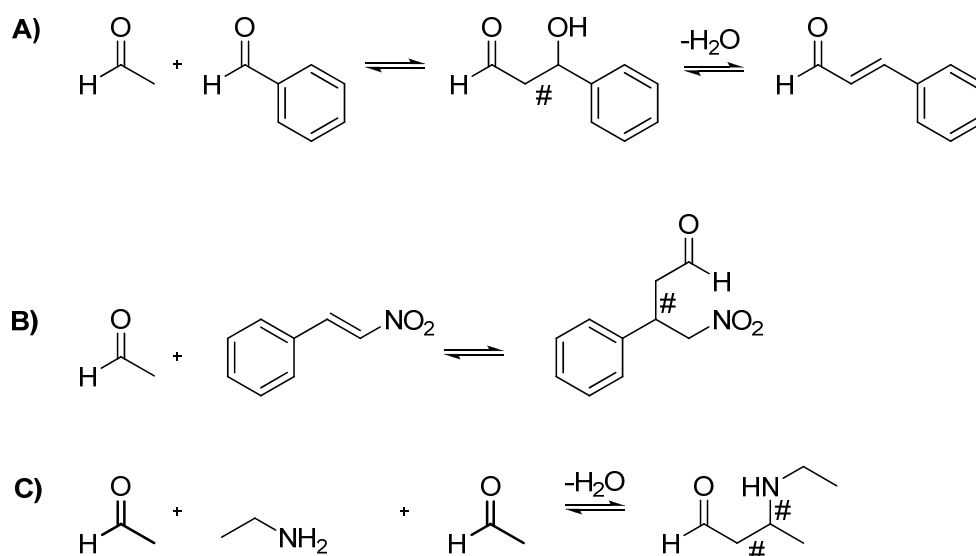


**Schema 2.** Enkele voorbeelden van C-C-bindingvormende reacties. A) de aldolcondensatie van acetaldehyde met benzaldehyde; B) de Michael-additie van acetaldehyde aan *trans*-nitrostyreen; C) Mannich-reactie van acetaldehyde en ethylamine. ___ = een binding; #= de nieuwe binding; H=waterstof; N= stikstof; O= zuurstof; de rest van de moleculen bestaat uit koolstof.

Het enzym 4-oxalocrotonaat tautomerase (4-OT) behoort tot de tautomerase-superfamilie en bezit dus het karakteristieke proline-residu (Pro-1) in de 'active site'. 4-OT katalyseert een tautomerisatiereactie, waarbij Pro-1 fungeert als base en een positief geladen waterstofatoom (proton) verplaatst op het substraat zodat het product gevormd wordt. Echter, uit **Hoofdstuk 1** is gebleken dat Pro-1 een chemisch veelzijdig aminozuur is en diverse rollen heeft binnen de tautomerase-superfamilie. De aanwezigheid van Pro-1 in de 'active site' van 4-OT maakt het wellicht mogelijk om C-C-bindingvormende reacties uit te voeren; immers, in aminokatalyse wordt het aminozuur proline gebruikt als katalysator van

C-C-bindingvormende reacties. Dit proefschrift beschrijft het onderzoek naar de mogelijkheid om C-C-bindingvormende reacties te katalyseren met 4-OT en de karakterisatie van deze promiscue, niet-natuurlijke reacties.

**4-OT katalyseert een aldolcondensatie**

De C-C-bindingvormende reacties waarvoor we wilden screenen, gaan allemaal uit van zogenaamde carbonylverbindingen. Deze verbindingen bevatten een zuurstofatoom dat via een dubbele binding gebonden is aan een koolstofatoom (bijvoorbeeld acetaldehyde in Schema 2). Om een C-C-binding te vormen moet een carbonylverbinding eerst reageren met Pro-1 tot een enamine-intermediair. Vervolgens kan dit enamine-intermediair verder reageren met andere moleculen wat resulteert in een nieuwe C-C-binding. Om nieuwe promiscue activiteiten te vinden in 4-OT hebben we een systematische screening strategie gebruikt (**Hoofdstuk 2**). Eerst is er getest met welke carbonylverbindingen Pro-1 een enamine-intermediair vormt. Het kleinste enamine-intermediair, gevormd door Pro-1 en acetaldehyde, is vervolgens getest in combinatie met verscheidene andere moleculen. Zo werd ontdekt dat 4-OT een niet-natuurlijke aldolase-activiteit heeft, waarin de aldolcondensatie van acetaldehyde met benzaldehyde wordt gekatalyseerd (Schema 2A). Naast Pro-1 is ook een arginine-residu (Arg-11) in de 'active site' belangrijk voor de aldolreactie. Hoewel de aldolcondensatie versneld wordt door 4-OT, is het een zeer lage activiteit. Om de activiteit van een enzym te verhogen kunnen mutaties worden aangebracht in de 'active site'. Dit betekent dat het ene aminozuur wordt vervangen door het andere aminozuur. De verschillende aminozuren beïnvloeden de enzym-gekatalyseerde reactie elk op een eigen manier. In 4-OT L8R, een 4-OT-variant waarin het aminozuur leucine (Leu-8) in de 'active site' is vervangen door een arginine-residu, is de promiscue aldolase-activiteit 16-maal verhoogd. Gebaseerd op deze waarnemingen hebben we een mechanisme voorgesteld dat de niet-natuurlijke aldolase-activiteit van 4-OT L8R kan verklaren (Schema 3).
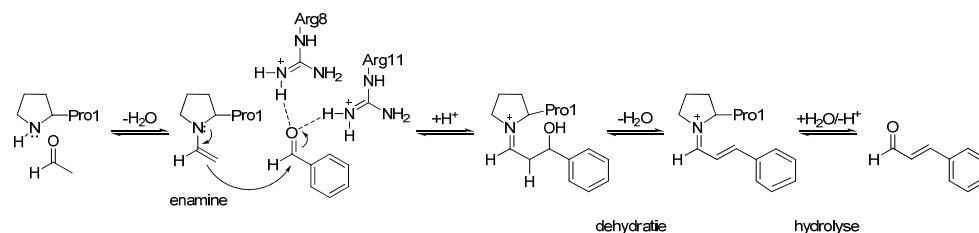


**Schema 3.** De aldolcondensatie van acetaldehyde met benzaldehyde die door 4-OT L8R gekatalyseerd wordt. Het verlies van water (dehydratie) is weergegeven terwijl het additieproduct nog op Pro-1 zit. Het kan echter ook zo zijn dat het additieproduct eerst los komt van Pro-1 door hydrolyse en vervolgens dehydrateert.

Om de aldolase-activiteit van 4-OT nog verder te verhogen hebben we goed gekeken naar de verschillende stappen in de reactie (Schema 3). Het vormen van een enamine-intermediair tussen Pro-1 en acetaldehyde gaat erg snel. Omdat het enamine-intermediair erg reactief is, verloopt de stap waarin een C-C-binding wordt gevormd tussen het enamine-intermediair en benzaldehyde waarschijnlijk ook snel. Onze hypothese was dat de laatste stap, waarin het product vrij komt van Pro-1 door hydrolyse, wellicht de langzaamste was. Door deze snelheidsbepalende stap te versnellen, zou de aldolase-activiteit van 4-OT verhoogd kunnen worden. We hebben de aldolase-activiteit getest van 4-OT F50A, een 4-OT-variant waarin een fenylalanine-residu (Phe-50) in de 'active site' is vervangen door het aminozuur alanine. Het is bekend dat deze 4-OT-variant een meer open 'active site' heeft, waardoor er makkelijker water binnen kan komen. Als er meer water aanwezig is, kan dit het 'losmaken' van het product door hydrolyse versnellen. Inderdaad, 4-OT F50A heeft een aldolase activiteit die 600 keer hoger is dan in gewoon 4-OT (**Hoofdstuk 3**).

De aldolcondensatie bestaat uit twee delen (Schema 2A): het koppelen van acetaldehyde en benzaldehyde tot het aldolproduct en de dehydratie van het aldolproduct tot het eindproduct. Tijdens de karakterisatie van de aldolcondensatie hebben we het aldolproduct niet waargenomen. Het zou kunnen dat het aldolproduct niet stabiel is en spontaan vervalt tot het eindproduct. Het is ook mogelijk dat de dehydratie van het aldolproduct door 4-OT gekatalyseerd wordt. In **Hoofdstuk 3** maken we het aldolproduct en tonen we aan dat 4-OT het aldolproduct accepteert als substraat. 4-OT katalyseert dus beide stappen van de aldolcondensatie.

**Intermezzo**

Voor het vervolg van deze samenvatting, is het noodzakelijk een korte uitleg te geven over isomeren. Isomeren zijn moleculen die opgebouwd zijn uit dezelfde atomen, maar die verschillen in de manier waarop deze atomen verbonden of georiënteerd zijn. Isomeren lijken op elkaar, maar zijn niet identiek en hebben vaak verschillende eigenschappen. Het ene isomeer kan een medicijn zijn, terwijl het andere isomeer niet actief is als medicijn, of zelfs giftig is. Er is een hoop te vertellen over isomeren, maar dit intermezzo is beperkt tot spiegelbeeldisomeren en *cis-trans*-isomeren, omdat die voorkomen in het vervolg van deze samenvatting.

Een klassiek voorbeeld om het begrip spiegelbeeldisomeren uit te leggen zijn onze handen (Figuur 3A). Beide handen zijn opgebouwd uit dezelfde onderdelen (handpalm en vingers), maar de onderdelen zijn anders georiënteerd: aan de linkerhand zit de duim rechts en de pink links, terwijl dat aan de rechterhand precies andersom is. Je handen zijn elkaars spiegelbeeld. Moleculen die isomeren van elkaar zijn, lijken gelijk maar zijn elkaars spiegelbeeldvorm (Figuur 3B). Om het verschil aan te geven wordt het ene isomeer (*S*)

genoemd en de spiegelbeeldvorm (*R*). De spiegelbeeldvormen zijn niet over elkaar heen te leggen.
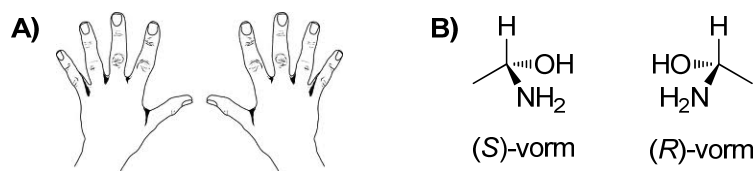


**Figuur 3**. A) De linker- en rechterhand lijken op elkaar maar zijn niet hetzelfde. (Bron: Florida Center for Instructional Technology Clipart ETC (Tampa, FL: University of South Florida, 2009), http://etc.usf.edu/clipart/37100/37133/us-10_37133.htm) B) Twee isomeren ((*S*) en (*R*)) zijn elkaars spiegelbeeld.

Behalve spiegelbeeldisomeren bestaan er ook andere soorten isomeren. Een soort die van belang is om te bespreken zijn de *cis*-*trans*-isomeren. Deze isomeren bevatten een dubbele C-C-binding. De groepen rondom de dubbele C-C-binding kunnen op verschillende manier georiënteerd zijn (Figuur 4). Als de groepen naar dezelfde kant wijzen spreken we van de *cis*-isomeer (*cis* is Latijn voor 'aan dezelfde kant'). In de *trans*-isomeer wijzen de groepen naar verschillende kanten (*trans* is Latijn voor 'aan de andere kant' of 'tegenover').
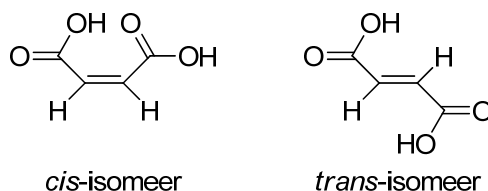


**Figuur 4.** Een voorbeeld van *cis*-*trans*-isomeren

### 4-OT katalyseert een Michael-additiereactie

Een ander voorbeeld van een C-C-bindingvormende reactie is de Michael-additiereactie (Schema 2B). Deze reactie gaat net als de aldolcondensatie uit van een carbonylverbinding. De carbonylverbinding vormt een enamine-intermediair met Pro-1, dat vervolgens verder kan reageren met moleculen die we Michael-acceptoren noemen. Een Michael-acceptor molecuul bevat altijd een dubbele C-C-binding (zie *trans*-nitrostyreen in Schema 2B) waaraan de carbonylverbinding gekoppeld wordt om zo een nieuwe C-C-binding te vormen. Om te testen of 4-OT een Michael-additiereactie kan katalyseren zijn we op dezelfde manier te werk gegaan als met de aldolcondensatie: het kleinste enamine-intermediair, gevormd door Pro-1 en acetaldehyde, is gecombineerd met verschillende moleculen (Michael-acceptoren). Dit resulteerde in de vinding dat 4-OT de Michael-additie van acetaldehyde aan

*trans*-nitrostyreen versnelt (**Hoofdstuk 4**). Het product van deze reactie is belangrijk, omdat het een bouwsteen is voor het antidepressivum phenibut. Echter, 4-OT maakt de verkeerde spiegelbeeldvorm van het product (de (*S*)-vorm), aangezien de actieve vorm van phenibut de (*R*)-vorm is.

## Een vreemde eend in de bijt: 4-OT katalyseert een isomerisatiereactie

In **Hoofdstuk 4** staat beschreven dat 4-OT de niet-natuurlijke Michael-additie van acetaldehyde aan trans-nitrostyreen versnelt, maar dat 4-OT de verkeerde vorm van het product vormt ((*S*) in plaats van (*R*)). Daarom hebben we getest of 4-OT ook de Michael-additiereactie van acetaldehyde aan *cis*-nitrostyreen katalyseert (**Hoofdstuk 5**). Het gebruik van *cis*-nitrostyreen in plaats van *trans*-nitrostyreen als substraat zou ervoor kunnen zorgen dat de juiste spiegelbeeldvorm (*R*) van het product wordt gemaakt. We vonden helaas dat 4-OT de Michael-additie van acetaldehyde aan *cis*-nitrostyreen niet versnelt. Verassend genoeg vonden we wel een andere reactie: 4-OT zet *cis*-nitrostyreen om in *trans*-nitrostyreen (Schema 4). Dit noemen we een isomerisatiereactie. Het is niet duidelijk hoe 4-OT deze reactie versnelt, maar het karakteristieke aminozuurresidu Pro-1 is belangrijk voor de reactie. Binnen dit promotieonderzoek is deze reactie een vreemde eend in de bijt, omdat er geen nieuwe C-C-binding wordt gevormd.
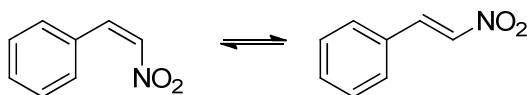


**Schema 4.** De isomerisatiereactie van *cis*-nitrostyreen naar *trans*-nitrostyreen.

## Algemene conclusies en toekomstperspectief

In dit promotieonderzoek hebben we gekeken naar de mogelijkheid om C-C-bindingvormende reacties te katalyseren met 4-OT. We hebben gevonden dat 4-OT meerdere reacties versnelt waarin een C-C-binding wordt gevormd, zoals een aldolcondensatie en een Michael-additiereactie. Per toeval ontdekten we ook dat 4-OT een *cis-trans*-isomerisatiereactie katalyseert, waarin geen nieuwe C-C-binding wordt gevormd. Al deze reacties zijn voorbeelden van niet-natuurlijke reacties van 4-OT.

Het ontdekken en karakteriseren van niet-natuurlijke activiteiten in 4-OT is een goed begin voor de verdere ontwikkeling van 4-OT tot een veelzijdige en efficiënte biokatalysator. Enzymen zijn vaak minder actief voor niet-natuurlijke reacties en versnellen deze reacties minder dan hun natuurlijke reacties. Dit geldt ook voor 4-OT. Om van 4-OT een efficiënte biokatalysator te maken zal de activiteit van 4-OT voor promiscue, niet-natuurlijke reacties verhoogd moeten worden. Dit kan door aminozuren in de 'active site' te veranderen, zoals we gezien hebben bij de aldolcondensatie. In de L8R- en F50A-varianten van 4-OT is slechts

één aminozuur vervangen, maar je kunt je voorstellen dat het vervangen van meerdere aminozuren een nog groter effect kan hebben op de enzymactiviteit.

Er zijn verschillende manieren om 4-OT (nog) veelzijdiger te maken in de toekomst. Het gebruik van andere substraten in een bekende reactie levert uiteindelijk andere producten op. In de Michael-additiereactie bijvoorbeeld kan in plaats van acetaldehyde een andere carbonylverbinding getest worden. Behalve het variëren van de substraten in al bekende reacties kan er onderzocht worden of 4-OT ook nieuwe C-C-bindingvormende reacties katalyseert. Een interessante reactie om te testen is een Mannich-reactie (Schema 2C). Deze reactie wordt veel toegepast bij het maken van medicijnen omdat daar vaak stikstof (N) in voorkomt. De veelzijdigheid van 4-OT kan, net als de efficiëntie, beïnvloed worden door aminozuren van het enzym te veranderen, met name in de 'active site'.

Wellicht leiden bovenstaande inspanningen ertoe dat 4-OT in de toekomst als efficiënte en veelzijdige biokatalysator kan worden toegepast in de industrie voor C-C-bindingvormende reacties.

# Dankwoord

Daar zijn ze dan…. De laatste paar bladzijden van dit proefschrift! Traditioneel enkele van de meest gelezen bladzijden, en om te schrijven echt niet de makkelijkste. Want natuurlijk heb ook ik het niet alleen gedaan, onderzoek is immers een teamsport. Niet alleen de expertise van de verschillende teamleden en de strategie van de coach is belangrijk, ook de verzorgers en waterdragers langs de kant zijn van essentieel belang voor het winnen van de wedstrijd!

Allereerst wil ik mijn promotores, prof. dr. G. J. Poelarends en prof. dr. W. J. Quax, bedanken. Gerrit, hartelijk dank voor je begeleiding in wat begon als een risicovol project. Wat waren we blij dat het werkte (ookal duurde het even voor we zelf overtuigd waren)! Ik heb veel geleerd over het doen van onderzoek, het uitvoeren van de juiste controles en het presenteren van resultaten in een publicatie. Het is mooi dat je, na jouw promotie tot 'associate professor', nu ook als officiële promotor kan optreden. Wim, omdat de dagelijkse begeleiding in handen was van Gerrit, stond je meer aan de zijlijn gedurende mijn promotietraject. Met prikkelende vragen bij werkbesprekingen en je kennis van toepassing en industrie heb je echter ook zeker een bijdrage geleverd aan dit onderzoek. Dank daarvoor!

Graag wil ik de leden van de leescommissie, prof. dr. R. P. H. Bischoff, prof. dr. ir. M. W. Fraaije en prof. dr. ir. A. J. Minnaard, bedanken voor hun vlotte en positieve beoordeling van dit proefschrift.

Ik heb tijdens dit onderzoek verschillende analysetechnieken toegepast. Sommige voerde ik zelf uit, zoals NMR spectroscopie. Pieter van der Meulen en dr. Renee Otten, zeer bedankt voor jullie hulp bij het oplossen van de problemen die ik tegenkwam!
Andere analyses werden uitgevoerd door anderen. Margot Jeronimus-Strathingh en Annie van Dam, hartelijk dank voor het uitvoeren van zoveel ESI-MS metingen. Marcel de Vries bedankt voor het verkrijgen van essentieel bewijs (dat Pro-1 gelabeld was) met behulp van nano-LC-MS/MS. Jullie worden alle drie, samen met dr. Hjalmar Permentier, bedankt voor jullie expertise in de interpretatie van resultaten.
Monique Smith wil ik bedanken voor het bepalen van de *ee* door HPLC analyse en dr. Wesley Browne voor zijn hulp bij het verkrijgen van *cis*-nitrostyreen.

Like in every other research group, apart from a few permanent employees, many people of various backgrounds come and go within the department of Pharmaceutical Biology. I would like to thank all colleagues I met in the past four years for the good working atmosphere and all the foreign snacks (from very sweet, to very spicy). I enjoyed working with you all!

Enkele collega's met wie ik veel direct heb samengewerkt wil ik graag bij naam noemen. Bert-Jan, bedankt voor je vriendschap en alle werkgerelateerde discussies die we gevoerd hebben. Fijn dat je één van mijn paranimfen bent! Edzard, de laatste twee jaar van mijn promotie heb ik met veel plezier met je samengewerkt. We praatten soms een wat andere 'taal', jij als organisch chemicus en ik als biochemica, maar we hebben wel een paar mooie 'werkjes' afgeleverd. Bedankt!

Zoals gezegd, de verzorgers en waterdragers langs de kant zijn ook van essentieel belang. Ze dragen niet altijd direct bij aan de wedstrijd, maar zijn er 'gewoon' en dat is erg fijn. Daarom wil ik mijn (schoon)familie en vrienden erg bedanken voor hun interesse in deze voor sommigen 'ver-van-hun-bed-show', voor hun steun wanneer dat nodig was en vooral voor alle gezelligheid. Karin, erg leuk dat je paranimf wilt zijn!

Rik, dankjewel…