University of Groningen

Exploring chemical versatility within the tautomerase superfamily

Baas, Jan

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

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

Publication date:
2014

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Baas, J. (2014). *Exploring chemical versatility within the tautomerase superfamily: Catalytic promiscuity and the emergence of new enzymes*. [Thesis fully internal (DIV), University of Groningen]. [S.n.].

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

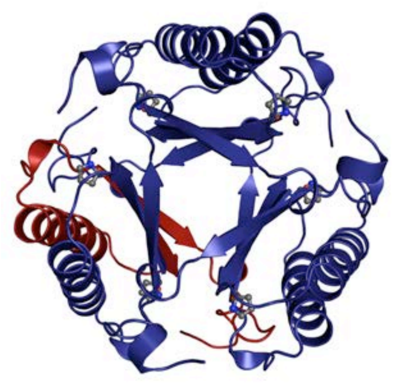
The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 12-12-2024

Chapter 1



- β-α-β building block
- Unique catalytic N-terminal proline
- Catalytically promiscuous

Recent advances in the study of enzyme promiscuity in the tautomerase superfamily.

Bert-Jan Baas, Ellen Zandvoort, Edzard M. Geertsema, and Gerrit J. Poelarends

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

Published in *ChemBioChem* (2013) 14, 917-926.

Abstract.
Exploiting enzyme promiscuity: Many enzymes exhibit catalytic promiscuity, which is the ability of an enzyme to catalyze an alternative reaction (or reactions) in addition to its biologically relevant one. These promiscuous activities may serve as starting points for both natural and laboratory evolution of new enzymatic functions. Recent advances in the study of enzyme promiscuity in the tautomerase superfamily, a group of homologous proteins with a characteristic β-α-β structural fold and a conserved catalytic amino-terminal proline, are discussed in this Minireview.

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 an enzyme to catalyze a chemically, and often mechanistically, distinct transformation (or transformations) in addition to its biologically relevant one.[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.[3c, 4, 5] 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, 3c, 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 lie formidable challenges and possibilities: the use of mechanistic reasoning to discover new promiscuous activities in existing enzymes, which could be exploited to generate novel biocatalysts.

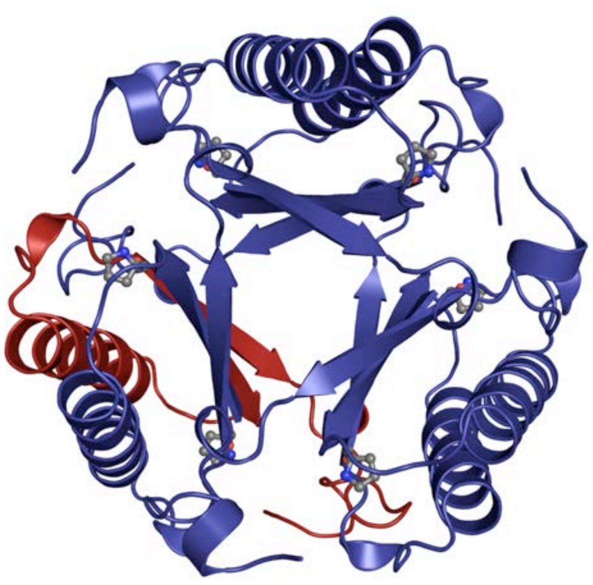


Figure 1. A ribbon diagram showing the β-α-β building block (highlighted in red) and the N-terminal proline residues (in ball/stick representation), that are characteristic features of members of the tautomerase superfamily, here represented by 4-oxalocrotonate tautomerase (4-OT).

These three topics will be addressed in this review 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).[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 the reactions catalyzed by these superfamily enzymes 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 review, recent advances in the study of catalytic promiscuity in the tautomerase superfamily and the implications derived thereof are reviewed.

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).[13,14] Both dehalogenases are found in soil-dwelling bacteria, where they are part of a catabolic pathway for 1,3-dichloropropene (Scheme 1), a xenobiotic nematocide introduced into the environment in the 1950s.[15] CaaD and *cis*-CaaD catalyze the hydrolytic dehalogenation of *trans*-3-CAA and *cis*-3-CAA, respectively, 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 2×10^{12}-fold rate enhancement when compared to the rate of spontaneous uncatalyzed hydrolysis of 3-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 products 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 tautomerization of phenylenolpyruvate to phenylpyruvate (Table 1, Scheme 2).[13] The presence of phenylpyruvate tautomerase (PPT) activity in these dehalogenases establishes a functional link with the known tautomerases in the superfamily. Although it is possible that new promiscuous activities might arise during divergence from an ancestor, the promiscuous PPT activity of CaaD and *cis*-CaaD is likely a vestige of the function of an evolutionary ancestor of these (presumably) recently evolved dehalogenases.

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

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

Enzyme abbreviations: CaaD, *trans*-3-chloroacrylic acid dehalogenase from *Pseudomonas pavonaceae* 170; *cis*-CaaD, *cis*-3-chloroacrylic acid dehalogenase from *Coryneform* bacterium strain FG41, 4-OT, 4-oxalocrotonate tautomerase from *Pseudomonas putida* mt-2, YwhB, a 4-OT homologue from *Bacillus subtilis*, Cg10062, a *cis*-CaaD homologue from *Corynebacterium glutamicum*, MsCCH2, a *cis*-CaaD family member from *Mycobacterium smegmatis*, MIF, macrophage migration inhibitory factor from *Mus musculus*.

Scheme 2. The tautomerization of (4-hydroxy)phenylenolpyruvate to (4-hydroxy)phenylpyruvate.

Scheme 3. The 4-OT-catalyzed tautomerization of 2-hydroxymuconate (i.e., 2-hydroxyhexa-2,4-dienedioate) to 2-oxo-3-hexenedioate.

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.

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 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, Scheme 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). Although the catalytic efficiency of 4-OT is about 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]

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). 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. The tautomerase YwhB, a 4-OT homologue from *Bacillus subtilis*, was found to be a promiscuous dehalogenase as well.[21,26] YwhB shares 36% pairwise sequence identity with 4-OT. Both Pro-1 and Arg-11 are conserved and are key catalytic residues for the observed tautomerase and dehalogenase activities of YwhB, 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). Although the direct progenitor of CaaD is unknown, it is thus likely that CaaD has evolved by divergent evolution from a 4-OT-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 promiscuous hydratase, converting 2-oxo-3-pentynoate (2-OP) to acetopyruvate

(Table 1, Scheme 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 ~120-fold lower compared to that of *cis*-CaaD (Table 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 is not the sole prerequisite for an enzyme being an efficient and specific dehalogenase. Discrimination between isomers and a high catalytic efficiency likely originate 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 from 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 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]

The study on Cg10062 shows that within the proteome of the bacterial kingdom *cis*-CaaD-like enzymes do exist, which may represent an intermediary stage in the evolution towards a fully functional dehalogenase. Without a primary activity that can be firmly linked to other members of the tautomerase superfamily, the low-level promiscuous hydratase and dehalogenase activities of Cg10062 offer few clues about the ancestral enzyme from which *cis*-CaaD may have evolved. To gain further insight into this matter, 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 are conserved in MsCCH2 as Pro-1, His-28, Arg-68 and Glu-112.

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.

Scheme 4. The hydration of 2-oxo-3-pentynoate (2-OP) to acetopyruvate.

The genomic context of the *mscch2* gene does not reveal any clear information about the physiological 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, Scheme 2 (R = H)]. Additionally, the enzyme also catalyzes the tautomerization of 4-hydroxyphenylenolpyruvate to 4-hydroxyphenylpyruvate [Table 1, Scheme 2 (R = OH)], albeit with lower efficiency.

Two promiscuous activities were identified in MsCCH2.[28] The enzyme acts as a hydratase converting 2-OP to acetopyruvate and as a dehalogenase using both *cis*- and *trans*-3-CAA (Table 1). The identification of MsCCH2 as a robust tautomerase with low-level promiscuous dehalogenase and hydratase activities established a functional link between the recently diverged *cis*-CaaD and a tautomerase of the same family. MsCCH2 could therefore resemble the progenitor of *cis*-CaaD, using its promiscuous dehalogenase activity as an essential starting point. Accumulation of beneficial mutations could then increase the promiscuous dehalogenase activity and yield an enzyme that is somewhere in between a specialist tautomerase and a specialist dehalogenase. An enzyme like Cg10062 might resemble such an intermediate on the evolutionary route from a specialist tautomerase toward a fully functional and isomer-specific dehalogenase.[28] Indeed, the low-level promiscuous tautomerase activity of *cis*-CaaD may be the vestige of the function of its progenitor.

The recurring ability of a tautomerase superfamily member to catalyze both the keto-enol tautomerization of phenylpyruvate and the hydrolytic dehalogenation of 3-CAA is an interesting observation in the study of the role of catalytic promiscuity in the evolution of CaaD and *cis*-CaaD (Table 1). This observation 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).[30] Intriguingly, MIF was found to act as a promiscuous dehalogenase towards *trans*-3-CAA (Table 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). Although the k_{cat} (6.0×10^{-3} s^{-1}) is rather low, this MIF mutant 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 of the PPT activity may indicate that these ancestral enzymes also featured this activity and that a tautomerase active site capable of catalyzing a keto-enol tautomerization of phenylpyruvate, 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 phenylenolpyruvate (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 (4-OT and YwhB as members of the same family as CaaD, and the *cis*-CaaD family members 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.

Scheme 5. Two proposed mechanistic routes for the CaaD-catalyzed dehalogenation of *trans*-3-CAA to malonate semialdehyde.

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 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 at 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 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 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. It was concluded that the active site of CaaD is equipped to ketonize enol compounds which makes the occurrence of an enol-intermediate in the dehalogenase mechanism more likely. Considering both mechanistic routes for the dehalogenation of *trans*-3-CAA (Scheme 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.

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.

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 pK_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 6).[32] The pK_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 by 2-OP but instead catalyzes a hydration reaction

Scheme 6. Modification of Pro-1 by 2-oxo-3-pentynoate yielding a covalent enamine-adduct. Subsequent reduction by $NaBH_4$ irreversibly traps the adduct.

converting 2-OP to acetopyruvate (Scheme 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 pK_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 pK_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 pK_a and that by altering the pH of the experimental conditions, one mechanism may be favored over the other.

Scheme 7. The proposed mechanism of the *cis*-CaaD-catalyzed conversion of 2,3-butadienoate to acetoacetate via water addition (A) or via a covalent enamine 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, Scheme 7). It was anticipated that the enzyme would add water to the electrophilic C-3 carbon atom of the substrate (Scheme 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 7B) that is favored over a mechanism involving the conjugate addition of water (Scheme 7A). The question was raised whether the dehalogenation of *cis*-3-CAA catalyzed by *cis*-CaaD could also proceed through a covalent intermediate in an addition-elimination type mechanism (Scheme 8).

Cl O O⁻ Enz NH ⇌ Cl O⁻ Enz NH + O⁻ $\xrightleftharpoons{-Cl^-}$ H O⁻ Enz NH + O $\xrightleftharpoons{+H_2O\ /\ -H^+}$ H O⁻ O O Enz NH

Scheme 8. A proposed alternative mechanism of the *cis*-CaaD-catalyzed dehalogenation of *cis*-3-CAA involving a covalent intermediate.

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, the formation of a covalent intermediate in the *cis*-CaaD-catalyzed dehalogenation of *cis*-3-CAA has not yet been demonstrated by labeling studies. 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.

Systematic screening for catalytic promiscuity

The majority of catalytically promiscuous activities known to this date has been discovered either by chance or by looking for a specific reaction catalyzed by a closely related enzyme.[7,9] This will undoubtedly continue in the future, and add to our understanding of evolutionary relationships between enzymes as well as yield insight into their mechanistic properties. The study of catalytic promiscuity has also shown that in many cases the same catalytic residues are involved in the primary and promiscuous activity of an enzyme but that these residues play different roles in the catalytic

mechanism of each activity. In other words, the various chemical properties of a set of active site residues are a determining feature in the chemical versatility of an enzyme's active site.[1a] This suggests another method of discovering catalytic promiscuity: catalytic promiscuity might be predicted based on the known chemical properties of active site residues. Recently, Poelarends and coworkers have predicted and identified a number of new promiscuous activities in 4-OT based on the chemical properties of the catalytic Pro-1 residue of this enzyme.

Pro1 N H O H $-H_2O$ / $+H^+$ + N Pro1 H H H H $-H^+$ N: Pro1 H

Scheme 9. Nucleophilic addition of the amino-terminal proline of 4-OT to acetaldehyde resulting in the formation of a covalent imine intermediate, which undergoes deprotonation to yield an enamine. Various other carbonyl compounds, e.g. various linear aliphatic aldehydes, can form a covalent enamine intermediate via the same mechanism.

The catalytic role of Pro-1 in the natural activity of 4-OT is that of a general acid/base.[11] In this tautomerase activity, Pro-1 abstracts the 2-hydroxyl proton of the substrate 2-hydroxymuconate and gives it back to the C-5 position (Scheme 3). This catalytic role of Pro-1 is possible because of its low pK_a of ~6.4 allowing it to function as a catalytic base at physiological pH. However, under these conditions, where it is deprotonated, Pro-1 might also be able to function as a nucleophile. This realization suggests possibilities in the area of nucleophilic catalysis and has led Poelarends and coworkers to hypothesize that 4-OT with its amino-terminal proline residue might have the ability to form enamines with carbonyl compounds (Scheme 9).[34] These enamine intermediates are nucleophilic species, which could subsequently react with a range of different electrophiles. As a final step, hydrolysis of the addition product from Pro-1 would regenerate the enzyme and yield the product. Hence, 4-OT might be able to act as a catalyst for carbon-carbon bond-forming reactions via this proposed mechanism. The feasibility of this concept is strengthened by the large amount of results in the field of aminocatalysis, in which the free amino acid proline and derivatives thereof are used as organocatalysts in carbon-carbon bond-forming reactions, which proceed via such enamine intermediates.[35] To investigate whether 4-OT indeed has this ability to react with carbonyl compounds, the enzyme was first screened with a number of carbonyl compounds for the formation of enamine intermediates.[34] Upon incubation (in separate reactions) with a set of small aldehydes and ketones as enamine donors, 4-OT was found to be inactivated in the presence of $NaCNBH_3$, which reduces the imine tautomer, leading to irreversible covalent modification of the enzyme. The covalently modified 4-OT was then analyzed by ESI-MS, which confirmed that the observed mass of the adduct was consistent with that of the reduced imine adduct of the starting carbonyl compound. Nano-LC-MS/MS analysis confirmed single-site modification of

Pro-1. These findings are consistent with the enamine-mechanism shown in Scheme 9. In order to test whether 4-OT is capable of catalyzing carbon-carbon bond-forming reactions, the chemical potential of the smallest enamine (that derived from acetaldehyde) was examined by screening 4-OT in the presence of acetaldehyde and various electrophiles. Intriguingly, 4-OT was found to catalyze two different types of carbon-carbon bond-forming reactions, which will be discussed below.

This systematic screening approach revealed that 4-OT can catalyze the aldol condensation of acetaldehyde with benzaldehyde yielding cinnamaldehyde (Scheme 10, Table 2).[34] By monitoring various control reactions by ^1H NMR spectroscopy and using synthetically prepared 4-OT, it was shown that the observed activity is indeed 4-OT-catalyzed. Further mechanistic insight into this promiscuous activity of 4-OT was gained by mutagenesis experiments. The native tautomerase activity of 4-OT is essentially based on three catalytic residues, being Pro-1, Arg-11 and Arg-39.[36] By mutating these residues to an alanine, Pro-1 and Arg-11 were shown to be crucial for the observed aldolase activity, whereas Arg-39 was found not to be essential. Analogous to the study of the dehalogenase activity of 4-OT, the 4-OT L8R variant was also tested to see if this active site mutation has an effect on the aldolase activity. Indeed, the 4-OT L8R variant has a 16-fold increased catalytic efficiency (Table 2). A possible explanation for this beneficial effect could be that the additional arginine residue exerts a stabilizing action on the negative charge that develops on the carbonyl oxygen of benzaldehyde during the course of the aldol coupling reaction. Furthermore, kinetic and labeling studies revealed that the reaction mechanism is sequential ordered. These results provide a basis for a mechanism in which Pro-1 reacts with acetaldehyde to form an enamine,

Pro1 -H_2O / +H^+ Pro1 -H^+ Pro1 +H^+ Pro1 +H_2O / -H^+ Pro1 -H_2O Pro1 OH

Scheme 10. Proposed mechanism of the 4-OT-catalyzed aldol condensation reaction of acetaldehyde with benzaldehyde yielding cinnamaldehyde as the final product.

which then attacks the electrophilic carbonyl carbon atom of benzaldehyde to form a carbon-carbon bond resulting in the aldol product as a covalent adduct to Pro-1 (Scheme 10). Subsequent dehydration of the aldol product and hydrolysis from Pro-1 yields cinnamaldehyde as the end product.

Due to the low catalytic rate of wild type 4-OT and the 4-OT L8R mutant, the next challenge was to further enhance the aldolase activity by mutagenesis. Zandvoort *et al*. postulated that the final hydrolysis step that releases the product from Pro-1 is likely to be the rate limiting step in the overall mechanism, based on the poor water accessibility of the active site of 4-OT.[37] Previously, Whitman and coworkers have analyzed the 4-OT F50A variant and have found that this mutation makes the active site more accessible to water, a feature that is also reflected by the slightly increased pK_a of Pro-1 (pK_a ~7.3) in rate limiting step in the overall mechanism, based on the poor water accessibility of the active site of 4-OT.[37] Previously, Whitman and coworkers have analyzed the 4-OT F50A

Table 2. Apparent kinetic parameters for the aldol, dehydration, retro-aldol and Michael-type addition reactions catalyzed by 4-OT, 4-OT L8R and 4-OT F50A.

Enzyme	Reaction	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
4-OT[34]		-	-	8.5×10^{-4}
4-OT L8R[34]		-	-	1.4×10^{-2}
4-OT F50A[37]		2.2×10^{-3}	4.3	0.5
4-OT F50A[37]		2.0×10^{-2}	2.8	7.2
4-OT F50A[37]		6.7×10^{-3}	1.7	3.9
4-OT[39]		1.7×10^{-2}	0.25	68
4-OT[39]		5.9×10^{-2}	1.6	37

Enzyme abbreviation: 4-OT, 4-oxalocrotonate tautomerase from *Pseudomonas putida* mt-2.

variant and have found that this mutation makes the active site more accessible to water, a feature that is also reflected by the slightly increased pK_a of Pro-1 (pK_a ~7.3) in this mutant.[38] Because of these observations, Zandvoort *et al*. tested whether the F50A mutation has a beneficial effect on the aldolase activity of 4-OT, reasoning that better water accessibility of the active site could enhance the rate of hydrolysis of the product from Pro-1 while retaining a pK_a of Pro-1 suitable for nucleophilic addition reactions. Indeed, the F50A mutant was found to have a profound effect on the aldolase activity, resulting in a 600-fold increase in k_{cat}/K_m relative to that of wild type 4-OT (Table 2).

During the course of analyzing the 4-OT-catalyzed aldol condensation reaction, both UV/VIS- and NMR spectroscopic analyses revealed cinnamaldehyde as the final product of the reaction. The presumed initial product of the aldol coupling reaction, i.e. 3-hydroxy-3-phenylpropanal, was not detected.[34,37] This raised the question whether the dehydration of this aldol product to form cinnamaldehyde is 4-OT catalyzed or not. Thus, 3-hydroxy-3-phenylpropanal was chemically synthesized.[37] When dissolved in the assay buffer in the absence of 4-OT, 3-hydroxy-3-phenylpropanal was found to slowly dehydrate and form cinnamaldehyde. In the presence of 4-OT, however, the rate of dehydration was significantly enhanced. Another interesting observation was the formation of acetaldehyde and benzaldehyde in addition to cinnamaldehyde when 3-hydroxy-3-phenylpropanal was incubated with 4-OT. This indicates that 4-OT catalyzes a retro-aldol reaction in addition to a dehydration reaction using 3-hydroxy-3-phenylpropanal as a substrate (Table 2). Analysis of the 4-OT P1A/F50A mutant showed greatly reduced activities for both reactions, indicating that both reactions are 4-OT-catalyzed and that Pro-1 is an important catalytic residue. This dual fate of 3-hydroxy-3-phenylpropanal in the presence of 4-OT is further proof that this product is a true intermediate in the 4-OT-catalyzed aldol condensation of acetaldehyde and benzaldehyde, even though it was not directly observed as a product of the aldol coupling reaction.

The second promiscuous 4-OT-catalyzed carbon-carbon bond-forming reaction which was discovered is the asymmetric Michael-type addition of acetaldehyde to *trans*-nitrostyrene or *p*-hydroxy-*trans*-nitrostyrene (Scheme 11).[39] The rationale behind selecting these particular substrates is based on the known reactivity of 4-OT towards acetaldehyde and the structural similarity of the Michael acceptors *trans*-nitrostyrene and *p*-hydroxy-*trans*-nitrostyrene to phenylpyruvate and *p*-hydroxyphenylpyruvate as known substrates of 4-OT. In addition, the product of the reaction between acetaldehyde and *trans*-nitrostyrene (4-nitro-3-phenylbutanal) is a precursor for the synthesis of the antidepressant phenibut. 4-OT was found to have pronounced activity towards these two Michael-type addition reactions (Table 2). The enzyme is highly stereoselective in the reaction of acetaldehyde and *trans*-nitrostyrene (*S*-enantiomer, 89% *ee*) and moderately stereoselective in the reaction of acetaldehyde and *p*-hydroxy-*trans*-nitrostyrene (*S*-enantiomer, 51% *ee*). In order to gain insight into the catalytic mechanism of this promiscuous Michael-type addition activity, the mutant enzymes 4-OT P1A, 4-OT R11A and 4-OT R39A were analyzed. These three variants showed complete (P1A) or nearly complete (R39A) loss of activity, or greatly diminished activity

R = H, OH

R = H; *ee* = 89%
R = OH; *ee* = 51%

Scheme 11. The 4-OT-catalyzed Michael-type addition of acetaldehyde to β-nitrostyrenes yielding the corresponding γ-nitroaldehydes.

(R11A), implicating all three residues as catalytically important in the overall mechanism. These results support a proposed mechanism in which Pro-1 reacts with acetaldehyde to form an enamine intermediate, which then undergoes a nucleophilic Michael-type addition to the C2-carbon atom of *trans*-nitrostyrene. Both arginine residues assist in catalysis either by positioning the substrate via its nitro functionality (Arg-11) or by delivering a proton to the C1-carbon atom of *trans*-nitrostyrene (Arg-39). Hydrolysis of the adduct from Pro-1 releases the nitroaldehyde product and regenerates the enzyme. This mechanism is analogous to that proposed for the 4-OT-catalyzed aldol condensation reaction, and supports the notion that enamine formation by the amino-terminal proline of 4-OT is the mechanistic basis of the promiscuous carbon-carbon bond-forming reactions catalyzed by 4-OT.

The discovery of these two types of carbon-carbon bond-forming activities of 4-OT is an interesting contribution to the field of catalytic promiscuity. Systematic screening for activities based on the chemical potential of active site residues has proven to be a powerful tool in the discovery of new promiscuous activities in enzymes.[34] In the specific case of 4-OT, the chemical properties of the unique amino-terminal proline can now be further exploited. It can be envisioned that the enzyme might catalyze yet other types of carbon-carbon bond-forming reactions (e.g. alkylation or Mannich reactions) in addition to those already discovered. The next step is to screen 4-OT towards a larger set of carbonyl compounds to explore the scope of enamine-intermediates that can be formed in the 4-OT active site. These intermediates can then be screened for reactivity towards a wide variety of electrophiles, which can lead to the discovery of new promiscuous reactions catalyzed by 4-OT. This approach is not limited to 4-OT alone, as the diversity of enzymes in the tautomerase superfamily could be a source of promiscuous catalysts for a variety of carbon-carbon bond-forming reactions. Initial discovery of useful promiscuous reactions will then be followed by mutational studies in order to improve the catalytic efficiency to a level that is required for application in industrial biocatalysis.

Concluding remarks

The tautomerase superfamily comprises a remarkable set of homologous proteins with diverse catalytic and biological functions in various physiological contexts. The most salient structural landmark for this superfamily of predominantly bacterial enzymes is a catalytically important amino-terminal proline embedded within the confines of a β-α-β structural fold. The essential Pro-1 residue can function as either a catalytic base or an acid, depending on the pK_a value, which is governed partly by the active site environment. Recent mechanistic and biochemical studies have revealed a diverse and partially overlapping pattern of promiscuous activities in the different superfamily members. The observation of shared promiscuous activities in these enzymes provides evidence for their divergent evolution from a common ancestor, gives hints regarding mechanistic relationships, and suggests that the emergence of new enzymes within the tautomerase superfamily was facilitated by catalytic promiscuity.

Most promiscuous activities of enzymes of the tautomerase superfamily have been discovered by looking for a specific reaction based on a family member. Identifying completely new promiscuous activities in these proline-based enzymes remains a challenge that requires a deep understanding of their reaction mechanisms as well as the chemical versatility of key catalytic residues. Recent systematic screening of 4-OT for promiscuous catalysis of carbonyl transformations, based on the mechanistic reasoning that its Pro-1 residue has the correct protonation state to function as a nucleophile at neutral pH, has led to the discovery of synthetically useful aldol and Michael-type addition activities in this enzyme. Further systematic screening of 4-OT and related superfamily members may prove to be a rewarding approach to discover new carbonyl transformation activities that could be exploited to develop efficient proline-based biocatalysts for carbon-carbon bond formation.

Acknowledgement

Work on enzyme promiscuity in our laboratory was supported in part by the Netherlands Organisation for Scientific Research (VIDI grant 700.56.421 to G. J. P.) and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement n° 242293 (to G. J. P.).

References.

1 a) A. Batbie, 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, *Biochemistry* **2009**, *48*, 1445-1453.

3 a) R. J. Kazlauskas, *Curr. Opin. Chem. Biol*. **2005**, *9*, 195-201; b) U. T. Bornscheuer, R. J. Kazlauskas, *Angew. Chem. Int. Ed.* **2004**, *43*, 6032-6040; c) O. Khersonsky, D. S. Tawfik, *Annu Rev Biochem.* **2010**, *79*, 471-505.

4 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.

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

35 a) S. Mukherjee, J. W. Yang, S. Hoffmann, B. List, *Chem. Rev.* **2007**, *107*, 5471-5569; b) P. Y. Bruice, *J. Am. Chem. Soc.* **1989**, *111*, 962–970; c) A. B. Northrup, D. W. C. MacMillan, *J. Am. Chem. Soc.* **2002**, *124*, 6798-6799.

36 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.

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

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

39 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.