Exploring chemical versatility within the tautomerase superfamily

Baas, Jan

Chapter 7

Summary and future perspectives

Catalytic promiscuity and the divergent evolution of enzyme function.

The textbook 'lock and key' paradigm that describes the general mechanism of enzyme catalysis has been shown to be too limited a view. In recent years, many examples of enzymes that catalyze more than one distinct reaction have been found; hence, not just one key fit's the lock. Enzyme promiscuity in terms of a broad substrate scope – i.e. substrate promiscuity – has been known for many years. More recently, however, it was found that the same enzyme can catalyze different chemical transformations, a feature which has been termed catalytic promiscuity. The study of catalytic promiscuity can provide valuable insights into the mechanism of enzyme evolution and resolve the contribution of active site residues to the mechanism of enzymatic activities. This, in turn, challenges researchers to predict new catalytic properties based on these mechanistic insights. These issues are discussed in **Chapter 1** using the tautomerase superfamily of enzymes as a model system.

Members of the tautomerase superfamily are characterized by a structural β-α-β-fold and a catalytic amino-terminal proline. The various chemical transformations catalyzed by members of this superfamily include the formation and/or cleavage of carbon-hydrogen (enol-keto tautomerizations), carbon-halogen (dehalogenase activities), carbon-oxygen (hydration and oxygenase activities) and carbon-carbon (aldol condensation and Michael-type addition activities) bonds. Catalytic promiscuity is a core feature of most, if not all, members of the tautomerase superfamily. Often the primary activity of one member is found to be a low-level promiscuous activity of another. These latent abilities may be vestiges from an evolutionary past, indicating that the activity of the specialist enzyme we see today has been nurtured during a course of evolution that started off with a generalist enzyme. Hence, an enzyme that features a promiscuous activity that fit's an evolutionary niche, may be selected as the starting template for the evolution of a new highly specialized enzyme. Thus, catalytic promiscuity may play a key role in the divergent evolution of function within enzyme families and superfamilies, which according to the studies reported in this thesis, might indeed be the case for the tautomerase superfamily.

Unraveling the mechanism of natural divergent evolution, and the analysis of the mechanistic roles which active site residues might play, challenges us to predict inherent catalytic abilities by mechanistic reasoning. This has allowed us to predict the ability of Pro-1 in 4-OT to act as a nucleophile in carbonyl-transformation reactions. As such, 4-OT was found to catalyze carbon-carbon bond-forming reactions, such as the aldol-condensation of acetaldehyde and benzaldehyde and the Michael-type addition of acetaldehyde to *trans*-nitrostyrene. In the laboratory, these activities can be further evolved by introducing strategic mutations in the active site that boost the rate at which these non-natural activities are catalyzed by 4-OT.

Two new examples of catalytic promiscuity: the dehalogenase activity of MsCCH2 and MIF.

The dehalogenation of the *trans*- and *cis*-isomer of 3-chloroacrylic acid (3-CAA) by a number of tautomerase superfamily members has been established in the past. These

chemically highly stable xenobiotics are rapidly converted into hydrochloric acid and malonate semialdehyde by two isomer-specific dehalogenases, CaaD and *cis*-CaaD, respectively. These organohalogens were first introduced into the environment in 1946, which raises the intriguing question of how nature has seemingly evolved two functional dehalogenases in a matter of a few decades.

Chapter 2 describes the identification, cloning and characterization of MsCCH2, a *cis*-CaaD homologue from *Mycobacterium smegmatis* stain MC2 155. The catalytic machinery of *cis*-CaaD was found to be only partially conserved in MsCCH2. MsCCH2 was found to act as a dehalogenase towards both isomers of 3-CAA, albeit with low catalytic efficiency. Instead, the enzyme was found to act as a robust phenylenolpyruvate tautomerase, by which the dehalogenase activity can be categorized as a case of catalytic promiscuity. Intriguingly, *cis*-CaaD functions as a promiscuous phenenolpyruvate tautomerase. The switched preferences of MsCCH2 and *cis*-CaaD indicate that MsCCH2 may be viewed as a look-a-like of an ancestral enzyme of *cis*-CaaD on the evolutionary route of becoming a full-fletched dehalogenase. This finding provides additional evidence for the crucial role of catalytic promiscuity in the evolution of new enzymatic activities in the tautomerase superfamily.

The question remained whether enhancing a low-level promiscuous activity is a facile evolutionary process or requires drastic – and hence more time consuming – changes to an enzyme. The work described in **Chapter 3** provides insight into this question by the discovery of low-level promiscuous dehalogenase activity in the mouse cytokine MIF and the 200-fold enhancement of this activity by the introduction of just two amino acid substitutions (I64V/V106L) in the active site pocket. This finding demonstrates that only a few changes to a cytokine are required to significantly boost its promiscuous dehalogenase activity. As MIF also functions as a phenylenolpyruvate tautomerase, and exhibits promiscuous dehalogenase activity, these findings again hint to an evolutionary route for the dehalogenase activity within the tautomerase superfamily where a tautomerase with promiscuous dehalogenase activity may have been recruited as the starting template for the evolution of fully active dehalogenase such as CaaD or *cis*-CaaD.

A unique member of the tautomerase superfamily: the cofactor-independent oxygenase RhCC.

The chemical versatility that is displayed by the various members of the tautomerase superfamily that have been characterized to date, is well illustrated by e.g. the remarkable rate-enhancement achieved by the dehalogenase activities of various members, and the recently discovered carbon-carbon bond-forming activities of 4-OT. The *cis*-CaaD homologue RhCC, from the soil-dwelling bacterium *Rhodococcus jostii* RHA1, is perhaps one of the most remarkable and unusual tautomerase superfamily members known; it's catalytic and structural characterization is described in **Chapter 4**. The enzyme was found to catalyze a cofactor-independent oxygenation reaction utilizing 4-hydroxyphenylenolpyruvate as substrate, which yields a complex mixture of

products. RhCC is part of a small group of cofactor-independent O_2-utilizing enzymes that is currently known to science. The complex mechanism which this enzyme employs to carry out the oxygenation reaction, which likely involves a feature termed 'substrate-assisted catalysis', presumably leads to the formation of a hydroperoxide intermediate. This highly reactive and unstable intermediate subsequently undergoes downstream reactions giving rise to various products. The mechanism of formation of three identified products is discussed as well as the presumed mechanism of hydroperoxide formation. The lack of control of RhCC to direct the chemistry of the reactive initial hydroperoxide to one (or two) oxygenation product(s), indicates that 4HPP is unlikely the natural substrate for RhCC. However, RhCC may very well be functional as an oxygenase towards a yet unknown substrate. Up to this date, RhCC stands alone in the tautomerase superfamily as a unique enzyme which has the remarkable capacity to activate molecular oxygen without the use of a redox-cofactor, finding a back-door to circumvent the thermodynamic as well as the chemical barrier that precludes the direct reaction between O_2 and organic substrates.

Discovery of the ability of 4-OT to catalyze carbon-carbon bond-forming reactions.
The analysis of the mechanistic details of promiscuous activities has provided a wealth of insight into the various roles which active site residues might play during catalysis. The discovery of promiscuous activities is mainly a chance event. However, the mechanistic knowledge gathered during past studies challenges us to predict catalytic abilities in a rational manner. This has led us to postulate that Pro-1 in 4-OT may act as a nucleophile, facilitated by the low pK_a value of ~6.4 of its secondary amine group. **Chapter 5** describes the screening of the reactivity of Pro-1 of 4-OT towards various carbonyl compounds and the discovery of the 4-OT catalyzed aldol-condensation of acetaldehyde and benzaldehyde yielding cinnamaldehyde as the final product.

Investigating Pro-1 mutants of 4-OT.
The role of Pro-1 in the catalytic mechanism of members of the tautomerase superfamily achieved much attention in the past. Generally, Pro-1 is mutated to an alanine in order to show the essential role which Pro-1 plays in the catalytic mechanism. The study of amino acid substitutions other than alanine or glycine is severely limited, as problems arise in the post-translational removal of the translation-initiating methionine. Recently, an engineered variant of *E. coli* methionine aminopeptidase (MetAP-*TG) was generated, which has the ability to remove the initiating methionine of peptide substrates with bulky or charged residues following Met-1. **Chapter 6** describes the production of the Pro-1 variants P1S, P1H and P1Q of 4-OT. P1S was found to be demethionylated by wild type MetAP; the 4-OT P1S enzyme shows low-level tautomerase activity towards the native substrate of 4-OT (2-hydroxymuconate, 2-HM) as well as decarboxylase activity towards oxaloacetate (a feature it shares with variants P1A and P1G). Both variants P1H and P1Q were obtained to a large extent in a demethionylated form by co-expression with MetAP-*TG; only a minor fraction of the total protein still harbors Met-1. The P1H/M1P2H mixture

was found to be almost catalytically inactive as a tautomerase towards 2-HM. An interesting feature observed for the P1Q variant is that glutamine exposed at the amino-terminus following demethionylation, undergoes an intramolecular cyclization reaction that yields pyroglutamate (pE). This is a unique way of introducing a proline analog at the amino-terminus of 4-OT by fermentation. The P1pE/M1P2Q mixture was found to still act as a robust tautomerase towards 2-HM.

General remarks and future perspectives.

The work reported in this thesis has shown that the identification and characterization of new members of the tautomerase superfamily can further extend our knowledge about the chemical versatility displayed by its members, as well as provide additional insight into the mechanistic diversity that is present. Catalytic promiscuity is a recurring feature among superfamily members and is likely to play a role in the divergent evolution of function within the superfamily. The promiscuous dehalogenase activity identified in MsCCH2 and MIF, the rather facile evolvability of which was shown for the latter, sets the stage to further investigate the past evolutionary routes that have resulted in the enzymes CaaD and *cis*-CaaD. The identification and characterization of new members of the *cis*-CaaD branch of the superfamily may fill the gaps between MsCCH2, Cg10062 and *cis*-CaaD and further shed light on the details of the evolutionary pathway that led to the functional *cis*-CaaD enzyme in a short time frame. An ultimate and highly informative goal would be to take MsCCH2 or the MIF I64V/V106L variant as a template for laboratory evolution towards a dehalogenase which displays similar catalytic parameters as the functional dehalogenases that evolved in nature. Do the catalytic properties of the intermediate enzymes generated during the course of laboratory evolution match with those of enzymes found in nature that are related to CaaD and *cis*-CaaD? Are these laboratory evolved intermediate enzymes characteristic of intermediates along the past evolutionary route from a tautomerase to dehalogenase? Such a study would provide key insights into the divergent evolution of function within an enzyme superfamily and the role which catalytic promiscuity might play in this process, and shed light on the requirements of the evolvability of this type of latent functions.

The work reported in this thesis on RhCC, the first member of the tautomerase superfamily which was found to catalyze an O_2-dependent reaction and to be cofactor-independent, sets the stage for further research into this unique phenomenon within the context of this superfamily. Homologues of RhCC present in *R. jostii* RHA1 and other organisms can be identified based on sequence similarity. 4HPP is an excellent mechanism-based probe to screen for new oxygenases within the superfamily, as this substrate supports the key mechanistic feature termed ‘substrate-assisted catalysis’, which is emerging as a common mechanistic requirement in the cofactor-less activation of molecular oxygen by different protein folds and families. Additionally, analogs of 4HPP can be used to probe the substrate scope of RhCC and related enzymes and may provide additional insights into the mechanistic details of the reaction. Mutagenesis and structural data will provide further insights into the mechanism of the cofactor-

independent activation of molecular oxygen within the β-α-β-fold architecture. Furthermore, it would be highly interesting to identify a homologue of RhCC of which its cofactor-independent oxygenase activity is of direct physiological relevance, and in which pathway such an activity may play an essential role. RhCC may represent the proverbial 'tip of the iceberg' in the exploration of the unusual cofactor-independent activation of molecular oxygen within the tautomerase superfamily of enzymes.