Structural studies into ketosteroid dehydrogenases and S-selective ω-transaminases
van Oosterwijk, Cornelis Christiaan

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:
2016

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

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).

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.
Chapter 5

Summary and Outlook
In this thesis we have investigated the three-dimensional structures of two different classes of proteins, ketosteroid dehydrogenases and \( \omega \)-transaminases, which both are relevant for biotechnological applications. In addition, \( \Delta 4 \)-ketosteroid dehydrogenase may be a promising target for drug design against actinomycetal pathogens such as *Mycobacterium tuberculosis* and *Rhodococcus equi*.

### Ketosteroid dehydrogenases

3-Ketosteroid dehydrogenases are flavoproteins that play key roles in steroid ring degradation. The enzymes are abundantly present in actinobacteria, including the catabolic powerhouse *Rhodococcus jostii* and the pathogenic species *R. equi* and *M. tuberculosis*. These enzymes are of interest for the synthesis of steroid-based pharmaceuticals, for the biotechnological production of novel steroids from, for example, cheap plant sterols (phytosterols), and as potential targets for new antibacterials.

In chapters 2 and 3 of this thesis we have studied the 3-ketosteroid \( \Delta 4-(5\alpha) \)-dehydrogenase (\( \Delta 4-(5\alpha)-\text{KstD} \)) from *R. jostii*, which introduces a double bond between the C4 and C5 atoms of 3-keto-(5\( \alpha \))-steroids. While chapter 2 describes the successful overexpression, purification and crystallization of the enzyme, chapter 3 gives an account of the X-ray crystal structure elucidation of the enzyme at 1.6 Å resolution, and gives details of its structure and putative catalytic mechanism. As expected, the fold of the enzyme is very similar to that of its nearest homologues, the fumarate reductases. Both enzymes contain a FAD co-factor and catalyze a dehydrogenation reaction resulting in the formation of a carbon-carbon double bond. However, they act on very different substrates and their catalytic residues are not conserved. Indeed, at the start of our research the nature of the KstD catalytic residues was completely unclear. Our \( \Delta 4-(5\alpha)-\text{KstD} \) structure revealed for the first time the presence of three putative catalytic residues (Tyr-319, Tyr-466, and Ser-468) in a pocket near the isoalloxazine ring system of the FAD co-factor. These residues are conserved in \( \Delta 4-(5\alpha)-\text{KstDs} \), and are crucial for catalytic activity, as we could demonstrate through site-directed mutagenesis. A crystal structure of the enzyme with the bound reaction product 4-androstene-3,17-dione allowed deducing their putative roles in catalysis. Ser468 has a position suitable for serving as the catalytic base, abstracting a proton from the C4
atom of the substrate. It is assisted by Tyr319, which possibly facilitates shuttling of the abstracted proton to the solvent. Tyr466 can serve as the catalytic acid, stabilizing the enolate intermediate occurring during the reaction through a hydrogen bond with the C3 oxygen atom of the substrate. Finally, the FAD N5 atom is ideally positioned to abstract a hydride ion from the 5α-position of the substrate. These features fully explain the reaction catalyzed by Δ4-(5α)-KstDs.

The hydrophobic steroid product is predominantly bound through van der Waals interactions, including ring stacking with a tryptophan side chain as the most conspicuous interaction. Only one polar interaction, a hydrogen bond between the C3 carbonyl oxygen atom and the catalytic acid Tyr-466, is present. Interestingly, no specific substrate recognition interactions with the C17 carbonyl oxygen atom are observed.

Besides the structure of Δ4-(5α)-KstD, the structure of the Δ1-KstD from *Rhodococcus erythropolis* SQ1 was also solved in our lab (Rohman et al., 2012; 2013). This latter enzyme is also a flavoprotein and catalyses the formation of a double bond at the C1-C2 position of the A-ring. The two KstDs share ~26 % amino acid sequence identity and have similar folds, although with an overall rmsd of 2.9 Å. This high rmsd value can be explained by a slight rotation of the S-domain. A superposition of the separate F- and S-domains yields much lower rmsd values of 1.8 and 1.6 Å, respectively. Interestingly, compared to the *R. jostii* Δ4-(5α)-KstD, the Δ1-KstD protein has 22 additional residues, primarily present in the S-domain close to the active site.

Comparison of the Δ4-(5α)-KstD and Δ1-KstD structures with bound products shows that the steroids occupy the same general position, but are rotated by approximately 40 degrees (Figure 1). This rotation positions the steroid in Δ1-KstD such that instead of the C5 atom, the C1 atom is in an optimal position for hydride abstraction by the FAD cofactor. The catalytic acid residue (Tyr466 in Δ4-(5α)-KstD), responsible for stabilizing the enolate intermediate, is conserved in Δ1-KstD (Tyr487). However, the catalytic base (Ser468 in Δ4-(5α)-KstD) and the residue assisting in proton transfer (Tyr319 in Δ4-(5α)-KstD) are not conserved. In Δ1-KstD their functions are taken over by Tyr119 and Tyr318, respectively, which originate from completely different parts of the protein.
Figure 1 - Comparison of the binding of the reaction products 1,4-androstadiene-3,17-dione (ADD; green) and 4-androstene-3,17-dione (4-AD; blue) as bound to Δ1-KstD and Δ4-(5α)-KstD, respectively. The structures were superimposed based on the isoalloxazine ring of the FAD cofactor. An angle of approximately 40 degrees is observed between the bound ADD and 4-AD products. The cross-structure distances between Ser468 in Δ4-(5α)-KstD and the site of attack on ADD, and between Tyr318 in Δ1-KstD and the site of attack on 4-AD, respectively, are indicated.

It is an intriguing question how the Δ1- and Δ4-(5α)-KstD enzymes have diverged to catalyse the reaction at different sites on the substrate. Since both enzymes use the FAD cofactor for abstracting a hydride ion from the substrate, it seems likely that positioning the substrate’s C5 atom instead of the C1 atom near the reactive N5 atom of the FAD should have been a crucial step in their divergence. As shown by the crystal structures of the Δ1- and Δ4-(5α)-KstD enzymes, the change in binding position of the substrate is achieved via a rotation, whereby the substrate’s C3 carbonyl group retains its position, and the catalytic acid retains its function. Although in both KstDs the steroid A- and B-rings fit tightly in a hydrophobic pocket, a certain degree of flexibility must have been possible in the last common ancestor to allow such a rotation.
Summary and Outlook

**Table 1** - Distances between the catalytic base and the carbon atom at which the proton is abstraction takes place.

<table>
<thead>
<tr>
<th></th>
<th>Δ1-KstD, Tyr318</th>
<th>Δ4-KstD, Ser468</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADD, C2</td>
<td>3.2 Å</td>
<td>3.2 Å</td>
</tr>
<tr>
<td>4-AD, C4</td>
<td>3.6 Å</td>
<td>3.3 Å</td>
</tr>
</tbody>
</table>

Using only their isoalloxazine ring systems to superimpose the product-bound Δ1- and Δ4-(5α)-KstD structures revealed that the two catalytic base residues (Tyr318 in Δ1-KstD, Ser468 in Δ4-(5α)-KstD) are both at a good distance (3.2-3.6 Å) from the steroid’s carbon atom that is to be deprotonated, in both the Δ1- and Δ4- orientations of the substrates (Table 1). In other words, Tyr318 in Δ1-KstD would be able to deprotonate the substrate in both the Δ1- and Δ4- (5α)-KstD substrate orientations. Similarly, Ser468 in Δ4-(5α)-KstD would also have a position suitable for deprotonating the substrate in both orientations. Thus, from these observations we conclude that the necessity of a drastic change in the catalytic base and associated proton transfer residue is not obvious. Therefore, it seems possible that other Δ4-(5α)-KstD enzymes may exist that still use the Δ1-KstD catalytic tyrosine residues, but that so far have escaped identification by bioinformatics methods.

A still open question concerns the biological role of Δ4-(5α)-KstDs. Ketosteroid dehydrogenases are often part of large gene clusters, which code for enzymes involved in the degradation of cholesterol and related steroids. For several Δ1-KstDs it has been shown that they are involved in cholesterol breakdown, but the role of the Δ4-(5α)-KstDs is less clear. Cholesterol contains a double bond at the C5-C6 position, which, after the action of cholesterol oxidase, shifts to the C4-C5 position, making the action of a separate Δ4-(5α)-KstD superfluous for cholesterol degradation. Possibly, the Δ4-(5α)-KstD participates in the degradation of other steroids, in which a C4-C5 double bond in the A ring does need to be introduced, such as cholate (Mohn et al., 2012; Haußmann et al., 2013), allowing them to be channelled into the steroid degradation pathways. Further biochemical and genetic studies are necessary to establish the biological role of Δ4-(5α)-KstDs, perhaps assisted by our structural results.

Another question concerns the localization and organization of the steroid degradation proteins. For M. tuberculosis it has been shown that cholesterol...
accumulate in the cell envelope outside the cell, from where it may be internalized when needed (Brzostek et al, 2009). Indeed, the cholesterol degradation gene clusters of several actinobacteria contain putative cholesterol ABC transporter genes (Mohn et al, 2008). Once inside the cell, the immediate destination of cholesterol and its steroid degradation products, which are mainly hydrophobic and poorly soluble, is unclear. However, membrane proteomics experiments in R. jostii RHA1 have shown that many of the early enzymes in the steroid degradation pathways, including some of the KstDs, are membrane-associated (Haußmann et al, 2013). This suggests the presence of a localised, rapid degradation system for cholesterol and its hydrophobic intermediates. A further open question is whether the degradation enzymes perform their function on their own or in complexes with other enzymes. Such knowledge may ultimately also contribute to establish the physiological role of the Δ4-(5α)-KstDs. The final products of the steroid degradation route (pyruvate, acetyl-CoA and propionyl-CoA) are channelled directly into the primary metabolism, providing energy and central metabolites. However, the precise fate of the protons and electrons generated by the dehydrogenation of steroids by the KstDs is still elusive. Their natural acceptors have not yet been identified, nor is it known how they contribute to the generation of ATP and reducing equivalents. Nevertheless, the cholesterol catabolism plays an important role in the pathogenesis of a number of pathogenic actinomycetes, such as M. tuberculosis and R. equi. Since cholesterol degradation is essential for their survival inside the macrophages (Pandey & Sassetti, 2008), the enzymes involved are interesting targets for inhibitor and vaccine design (van der Geize et al, 2007; 2011). For example, knocking out the Δ1-KstD enzyme, which is essential for the pathogens, could sufficiently weaken the pathogen to allow its subsequent use in vaccine preparations. Furthermore, since there are no human homologs of the KstD enzymes (Knol et al, 2008), the KstDs may form attractive targets for inhibitor design. An obvious starting point for the design of inhibitors would be the bound steroid observed in the KstD active site. However, this compound has only two groups with hydrogen-bonding potential, i.e. its C3 and C17 carbonyl groups. In the KstD active site, the C3 carbonyl is hydrogen bonded to the catalytic acid residue, but the C17 carbonyl group does not interact with the protein. Further interactions are unspecific hydrophobic interactions, such as stacking of the steroid ring
structure with an aromatic residue. This lack of specific interactions may make it very difficult to develop specific inhibitors on the basis of the steroid core. Another complication for steroid-based inhibitor design is the wide presence of steroid-like molecules in humans, which risks interference of steroid-based KstD inhibitors with cholesterol and steroid hormone homeostasis and signalling pathways. To avoid such potential problems, a more promising approach could be to start from scratch, e.g. by using structure-guided fragment-based drug design approaches, which have proven successful in many cases (van Montfort & Workman, 2009). Nevertheless, before using our KstD4 structure for such an undertaking, the essentiality of Δ4-(5α)-KstD for the survival of the organism should be established beyond doubt.

Transaminases
Transaminases are PLP-dependent enzymes that reversibly exchange an amino group on one molecule for a carbonyl oxygen atom from another molecule. Transaminases are distinguished according to the position of the amino group that is transferred. α-Transaminases act on the α-carbon of an amino acid while ω-transaminases (ω-TAs) catalyse the transfer of an amino group from any of the other carbon atoms. In nature, ω-transaminases occur in two fold types (fold types I and IV), with the fold type I enzymes being (S)-selective, and the fold type IV enzymes being (R)-selective. In this thesis, we investigated two (S)-selective fold type I ω-transaminases from Arthrobacter sp. (Ars-ωTA) (Kariskoga, 2007) and Bacillus megaterium (BM-ωTA), respectively, which have 95% sequence identity. These enzymes are of interest for biotechnology for introducing an amino group into building blocks for chiral amine-containing bioactive compounds, which include anti-depressants, herbicides and antibacterials (Koszelewski et al, 2008; Höhne & Bornscheuer, 2009; Koszelewski et al, 2010b).

As detailed in Chapter 4 crystal structures were obtained of both transaminases, both in the native form and with bound (R)-α-methylbenzylamine or L-alanine, at resolutions of 2.3 Å or better. As expected from the 95% sequence identity of the enzymes their structures are highly similar, with an rmsd of less than 0.5 Å.

Four of the twenty amino acid differences between Ars-ωTA and BM-ωTA occur in the active site. Three of them (Tyr60Cys, Tyr164Phe, Val436Ala)
affect the size of the O pocket, which binds the aromatic substituent of their prototypic test substrate (S)-α-methylbenzylamine. As a result, the O pocket of Ars-ωTA is more spacious than that of BM-ωTA, suggesting that Ars-ωTA is able to accept larger substrates than BM-ωTA. However, activity assays with substrates containing a naphtyl group as the largest substituent did not reveal any differences. Possibly, differences in substrate size preference only occur with substrates possessing a substituent larger than a naphtyl group. The smaller P-pocket, which binds the methyl group of (S)-α-methyl-benzylamine, is not affected by the amino acid differences in the active site.

An essential step in the catalytic mechanism of ω-transaminases is the abstraction of a proton from the amino-group bearing carbon atom, to proceed from the external aldimine state (in which the substrate is covalently linked to the PLP cofactor via a Schiff base) to the ketimine intermediate. This proton abstraction requires that the hydrogen atom is pointing towards the active site base, the lysine residue, which is accomplished by binding the large aromatic substituent in the O pocket and the methyl group in the P pocket. The structure of the external aldimine state of BM-ωTA with the unfavoured (R)-α-methylbenzylamine enantiomer clearly showed why the enzymes are (S)-selective. (R)-α-methylbenzylamine is bound with its hydrogen atom pointing away from the catalytic lysine thus preventing its deprotonation. Yet, the enzymes do show some activity on the (R)-enantiomers, which may be explained by the presence of an alternative base such as e.g. a water molecule, or, for smaller substrates, by the flexibility of the binding interaction allowing access of the (R)-proton to the lysine residue. Therefore, tight binding of the substrate side chains in the O- and P-pockets would be important for ensuring the correct orientation of the hydrogen atom and thus could be an important factor for the enantioselectivity of the enzymes. Tight binding is achieved by the presence of two “gate” residues at the end of the substrate-binding tunnel close to the P-pocket, which form a tight-fitting slot for the substrate to bind in. Interestingly, Ars-ωTA and BM-ωTA differ in one of the gate residues: Ars-ωTA has a valine at position 242, while BM-ωTA has an Ala. The slightly larger valine may better restrain the position of the substrate’s phenyl group and thus reducing flexibility and preventing access of water. Although, in principle, the Val/Ala difference thus may affect the enantioselectivity of the enzymes, kinetic characterization has not yet borne this out.
The transaminase reaction consists of two half reactions on substrates with very different substituents (charged vs. (large) hydrophobic) that both bind in the O-pocket. To accommodate the binding of such different groups fold type I transaminases commonly use an arginine residue that switches between an in- (for binding the carboxylate) and out-position (for binding the hydrophobic side chain). In BM-ωTA and Ars-ωTA this arginine is Arg442, which is a different arginine residue from the switching arginine usually found in TAs (see e.g. Steffen-Munsberg et al. (2012a)) and it comes from a different region of the protein. Ars-ωTA and BM-ωTA have a less open active site compared to other TAs and an arginine at the usually observed position would obstruct the entrance to the active site, thus necessitating a different Arg to act as the switching arginine.

One of the key remaining questions on the catalytic mechanism of ω-transaminases is the role of water. Water provides the oxygen atom needed for formation of the keto-product from the ketimine intermediate. On the other hand, water may also function as an alternative base, which may result in the unwanted enantiomer being produced (see above). The (R)-α-MBA adduct structure contains two water molecules, at 4.1 and 5.4 Å from the nearby site of attack. It is not clear if any of them, and if so, which of them attacks the ketimine intermediate, nor whether any of them could act as alternative base to produce the unwanted enantiomer. A crystal structure of, for example, a substrate analogue trapped in the ketimine state could allow identification of the water molecule involved in the nucleophilic attack.

Recently, a considerable effort has been devoted to discover and characterize new ωTAs with unique substrate preferences and enhanced enantioselectivity. Such new enzymes add substantially to the growing toolbox of ωTA enzymes that can be used for easy selection of enzymes to be applied on new substrates or in new biocatalytic conversions. For (further) optimization the enzymes may subsequently be engineered for improved stability, enantio- and substrate-specificity and activity. Many of the currently applied enzymes use alanine as the amide donor substrate, which is a relatively expensive substrate for use in industry. Alanine usage may be reduced by the use of a co-substrate regeneration system or by engineering the enzymes to accept cheaper amine donors like, for example, isopropylamine. Finally, the stability of the enzymes is an important factor for industrial applications and adapting them to new temperatures, pH values and the presence of (organic) solvents is major goal for their application.
ray structures may assist in the rational mutagenesis of $\omega$TAs to enhance their applicability.