

ChemBioChem

Combining Chemistry and Biology



European Chemical Societies Publishing



Accepted Article

Title: Reversal of regioselectivity in zinc-dependent medium chain alcohol dehydrogenase from Rhodococcus erythropolis toward octanone derivatives

Authors: Mehdi D. Davari, Gaurao V. Dhoke, Yunus Ensari, Dinc Yasat Hacibaloglu, Anna Gärtner, Anna Joëlle Ruff, and Marco Bocola

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: ChemBioChem 10.1002/cbic.202000247

Link to VoR: https://doi.org/10.1002/cbic.202000247

WILEY-VCH

Reversal of regioselectivity in zinc-dependent medium chain alcohol dehydrogenase from *Rhodococcus erythropolis* toward octanone derivatives

Dr. Gaurao V. Dhoke[§],^[a] Dr. Yunus Ensari[§],^[a,b] Dinc Yasat Hacibaloglu,^[a] Anna Gärtner,^[a] Dr. Anna Joëlle Ruff,^[a] Dr. Marco Bocola,^[a] Dr. Mehdi D. Davari^{*},^[a]

^[a]Lehrstuhl für Biotechnologie, RWTH Aachen University, Worringerweg 3, 52074 Aachen, Germany

^[b]Kafkas University, Faculty of Engineering and Architecture, Department of Bioengineering, Kars, Turkey

§ G.V.D. and Y.E are joint first authors.

*Corresponding author: Dr. Mehdi D. Davari Lehrstuhl für Biotechnologie, RWTH Aachen University Worringerweg 3, 52074 Aachen (Germany) Fax:(+ 49) 241-80-22387 E-mail: <u>m.davari@biotec.rwth-aachen.de</u> Twitter: @mehdi_d_davari; @EnsariYunus; @DhokeGaurao; @RWTH.

Abstract

Zinc-dependent medium chain alcohol dehydrogenase from Rhodococcus erythropolis (ReADH) is one of the most versatile biocatalyst for stereoselective reduction of ketones to chiral alcohols. Despite the known broad substrate scope, ReADH accepts only carbonyl substrates with either methyl or ethyl group adjacent to carbonyl moiety which limits its use in synthesis of chiral alcohols which serve as a building blocks for pharmaceuticals. Protein engineering to expand the substrate scope of ReADH toward bulky substitutions next to carbonyl group (ethyl 2-oxo-4phenylbutyrate) opens up new routes in the synthesis of ethyl-2-hydroxy-4phenylbutanoate, an important intermediate for anti-hypertension drugs like enalaprilat and lisinopril. We have performed computer-aided engineering of ReADH toward ethyl 2-oxo-4-phenylbutyrate and octanone derivatives. W296 which is located in the small binding pocket of ReADH restricts sterically the access of ethyl 2-oxo-4phenylbutyrate, 3- or 4-octanone toward the catalytic zinc ion and thereby limits ReADH activity. Computational analysis was used to identify position W296 and Site Saturation Mutagenesis (SSM) yielded an improved variant W296A with a 3.6 fold improved activity toward ethyl 2-oxo-4-phenylbutyrate when compared to ReADH WT (ReADH W296A: 17.10 U/mg and ReADH WT: 4.7 U/mg). In addition, the regioselectivity of ReADH W296A is shifted toward octanone substrates. ReADH W296A has a more than 16 fold increased activity toward 4-octanone (ReADH W296A: 0.97 U/mg and ReADH WT: 0.06 U/mg) and a more than 30 fold decreased activity toward 2-octanone (ReADH W296A: 0.23 U/mg and ReADH WT: 7.69 U/mg). Computational and experimental results revealed the role of position W296 to control the substrate scope and regiopreference of ReADH toward a variety of carbonyl substrates.

Keywords: Alcohol dehydrogenase, *Rhodococcus erythropolis*, Protein engineering, Homology Modeling, Molecular docking

Running title: Reversal of regioselectivity of ADH from Rhodococcus erythropolis

1-Introduction

Secondary alcohol dehydrogenases (ADHs) are the enzymes belonging to the class of oxidoreductases (E.C.1.1.1.1, also called as keto-reductases) and represent an important class of biocatalysts because of their potential ability to stereospecifically reduce prochiral carbonyl compounds. Stereoselectiveness is a unique characterstic among secondary alcohol dehydrogenases,^[1] which makes them of great interest for the biocatalytic production of chiral alcohols as a building blocks for fine chemical industries.^[2] There are few described ADHs namely phenylacetaldehyde reductase from Cornynebacterium sp. Strain,^[3] Candida parapsilosis aldehyde dehydrogenase 5 (cpADH5),^[4] and ADH from Lactobacillus kefir,^[5] which are cabable of reducing ketones with bulky side chains. Recently, an ADH from the halophilic archaeon Haloferax volcanii (HvADH2) has been identified^[6] and engineered by rational design approach to broaden its substrate scope towards the conversion of a variety of aromatic substrates.^[7] Furthermore, a bacterial β-ketoacyl-ACP reductase (FabG) from Bacillus sp. ECU0013^[8] and NADPH-dependent (S)-carbonyl reductase from Candida parapsilosis ATCC 7330^[9] reduces ethyl 2-oxo-4-phenylbutanoate (EOPB) to (S)-ethyl 2-hydroxy-4-phenylbutanoate (S-EHPB). Optically pure EHPB is used in the synthesis of an important intermediate for anti-hypertension drugs such as enalaprilat and lisinopril^[10]. In spite of having capability of reducing broad variety of substrate, these enzymes differ in their substrate specificity and stereoselectivity. In addition, they have insufficient operational long-term stability or limited substrate acceptance which restricts their industrial applications. The ability to control the substrate specificity and stereochemistry of ADH reactions is of increasing interest in biocatalysis.^[11] In recent years, a novel NADH-dependent ADH from *Rhodococcus* erythropolis DSM 43297 (ReADH) has been emersed as a promising candidate for use in industry.^[5, 12] ReADH has a homodimeric structure and a molecular weight of 36,026 kDa per subunit.^[12] It consists of 348 amino acids per subunit and belongs to the zinc-dependent medium chain alcohol dehydrogenase (MDR) family. ReADH follows the Prelog's rule^[13] which leads to formation of S-enantiomers with a high purity $(e.e. \ge 99 \%)$.^[12, 14] ReADH has a high stereo-, chemo- and regioselectivity towards variety of carbonyl substrates including acetophenone derivatives, aliphatic ketones and can withstand temperatures up to 65 °C.^[12, 14] Recently, Kasprzak et. al^[15] has reported the synthesis of 1-(S)-phenyl ethanol and ethyl (R)-4-chloro-3-

10.1002/cbic.202000247

hydroxybutanoate employing a biphasic and a substrate-coupled cofactor regeneration system. ReADH reduces carbonyl substrates that have either a methyl or a ethyl group adjacent to carbonyl functionality^[14] and a crystal structure has not yet been reported.

In this work, we aimed to understand the structural determinants controlling the substrate scope of ReADH. In order to expand the substrate scope of ReADH toward bulkier substrates (ethyl 2-oxo-4-phenylbutyrate, 3- or 4-octanone), we reengineered the smaller binding pocket of ReADH. Thereby we constructed a homology model of ReADH and used for molecular docking of carbonyl substrates to identify positions W296 as a key position in modulating the substrate access to the catalytic zinc center. Identifcation was performed with our previously developed docking protocol.^[4b]

2. Methods

2.1 Homology modeling

The amino acid sequence of ReADH (Accession no: C0ZXL4, 348 aa) was retrieved from UniProt database^[16] in FASTA format. Homology modeling of ReADH structure was performed by using YASARA Structure Version 13.9.8^[17] with the default settings (PSI-BLAST^[18] iterations: 6, E value cutoff: 0.5, templates: 5, OligoState: 4). Template structures was scored based on position specific scoring matrix (PSSM).^[19] Two X-ray templates were selected for hybrid modeling of ReADH structure (348 aa): Crystal structure of alcohol dehydrogenase from *Rhodococcus ruber* (345 residues with quality score 0.487, PDB ID: 3JV7,^[20] resolution 2.0 Å) and crystal structure of the alcohol dehydrogenase from the hyperthermophilic archaeon *Aeropyrum pernix* (343 residues with quality score 0.581, PDB ID: 1H2B,^[21] resolution 1.62 Å). The developed model is a hybrid homotetramer shown in **Figure S1** in SI.

2.2 Mechanism based substrate docking

The homodimeric structure of ReADH was used for substrate docking using AutoDock4.2^[22] within YASARA Structure Version 13.9.8.^[17a] In this mechanism-based docking protocol, reparameterization of catalytic zinc environment including

catalytic zinc ion and its coordinating residues (Cys38, His62, Asp153, and Water) was carried out. Additionally, reaction mechanism based distance filter criteria (distances D1, D2, and D3 shown in **Figure S2** in **SI**) was applied. D1 signifies the distance between carbonyl oxygen of the substrate to the catalytic zinc ion, D2 signifies the distance between the hydrogen (H) of NADH cofactor to the carbonyl carbon (C) of substrate and D3 signifies the distance between carbonyl oxygen to the (HO) group of Ser46. The detailed of substrate based docking protocol along with applied distance constraint can be found elsewhere.^[4b] For substrate molecular docking protocol, bonded model of catalytic zinc ion with charge of +1.01 on catalytic zinc ion¹² and AM1-BCC charges^[23] for the substrate were used. Pymol was used to prepare all the graphical images.

2.3 Chemicals

All chemicals were purchased from Sigma (Steinheim, Germany), Aldrich (Steinheim, Germany), Fluka (Steinheim, Germany), Merck (Darmstadt, Germany) and Roth (Karlsruhe, Germany) if not stated otherwise. NAD⁺ cofactor was purchased from Sigma-Aldrich (Steinheim, Germany). Primers were ordered from Eurofins MWG (Ebersberg, Germany).

2.4 Construction of Site-Saturation Libraries (SSM) of ReADH

The expression vector pKA1, harboring the (*S*)-alcohol dehydrogenase from *Rhodococcus erythropolis* (ReADH),^[12] was used as template for saturation mutagenesis of position W296. Modified two step QuikChange Mutagenesis (QCM) protocol was applied to generate mutant libraries. In the first step, two extension reactions are performed in separate tubes; one containing the forward primer (5`-ACAGTTCCGTAT**NNK**GGTGCCCGC-3') and the other containing the reverse primer (5'- GCGGGCACC**MNN**ATACGGAACTGT -3'). After 3 rounds of extension, the two reactions are mixed and amplification carried out for 15 cycles. Obtained PCR products were digested with 20U *DpnI* (New England Biolabs, Frankfurt am Main, Germany) in order to remove template DNA and purified with PCR purification kit (NucleoSpin[®] Gel and PCR Clean-up kit, Macharey-Nagel, Düren, Germany).

Libraries were transformed into *E.coli* BL21 Gold (DE3) and plated on LB_{CM} agar plates.

2.5 Protein Expression in 96-Well Plates and Preparation of Crude Cell Extracts

Libraries were constructed by transferring single colonies from agar plate to 96-well microtiter plates (Greiner Bio-One GmbH, Frickenhausen, Germany) filled with 150 μ L LB medium supplemented with 34 μ g/ml chloramphenicol. Plates were tightly sealed with insulation tape and incubated in a microtiter plate shaker (Multitron II, Infors GmbH, Einsbach, Germany; 16 h, 37 °C, 900 rpm, 70 % humidity). For long-term storage, 100 μ l of 50 % (v/v) glycerol was added to each well and the libraries were stored at -80 °C as master microtiter plate.

Expression in microtiter plates was performed by duplicating master plate with 96-pin replicator to 150 μ L of LB medium supplied with chloramphenicol (34 μ g/ml) in 96-well flat bottom microtiter plates (pre-culture). Plates were closed with lids, tightly sealed with insulation tape and incubated for 16 h (37 °C, 900 rpm, 70 % relative humidity). Then, 10 μ L of the pre-culture were transferred to V-bottom microtiter plates (transparent polystyrene plate, Corning GmbH, Kaiserslautern, Germany) containing 150 μ I terrific broth (TB) medium supplemented with chloramphenicol (34 μ g/ml), trace element solution (0,25 mL containing; 3.4 mM CaCl₂, 0.6 mM ZnSO₄, 0.6 mM MnSO₄, 54.0 mM Na₂–EDTA, 61.8 mM FeCl₃, 0.6 mM CuSO₄ and 0.8 mM CoCl₂ in water), 0.1 mM IPTG, 0.1 g/L thiamine hydrochloride and 1 mM ZnCl₂. Plates were closed with lids, tightly sealed and incubated in a microtiter plate shaker for 20 h (30 °C, 900 rpm, 70 % relative humidity). Expression cultures were harvested by centrifugation (4 °C, 3220 g, 15 min) using an Eppendorf 5810R centrifuge (Eppendorf AG, Hamburg, Germany) and stored at -20 °C until further use.

2.6 96-well MTP activity assay for screening of SSM libraries.

The expressed variants in 96-well microtiter plates were taken out of the freezer and kept at room temperature for 10 min. Cell lysates were prepared by resuspending each cell pellet by pipeting in 150 μ L lysis buffer (50 mM phosphate buffer pH 7.5, and 1 g/L

lysozyme). Cell lysis was performed by incubating suspended cells in a microtiter plate shaker for 1 h (37 °C, 900 rpm, 70 % relative humidity) and followed by centrifugation (4 °C, 3220 g, 20 min). Screening of SSM library was performed through monitoring NADH depletion, result of ethyl 2-oxo-4-phenylbutyrate reduction to corresponding alcohol, spectrophotometrically (Tecan sunrise, Männedorf, Switzerland). Reaction setup contains 145 µl phosphate buffer pH 7.5, 50 µl of crude cell lysate, and 5 µl of 0.5 M ethyl 2-oxo-4-phenylbutyrate in ethanol. The reaction was initiated by addition of 50 µl of 1 mM NADH and absorption was monitored at 340 nm for 10 min (ϵ = 6200 1/(M cm)).

2.7 Flask Expression of ReADH Protein

For flask expression of ReADH enzyme, 4 ml of overnight culture grown in LB medium with chloramphenicol (34 μ g/ml; 37 °C, 250 rpm) was inoculated into 250 ml TB medium supplemented with chloramphenicol (34 μ g/ml) and trace element solution (0,25 ml containing; 3.4 mM CaCl₂, 0.6 mM ZnSO₄, 0.6 mM MnSO₄, 54.0 mM Na₂– EDTA, 61.8 mM FeCl₃, 0.6 mM CuSO₄ and 0.8 mM CoCl₂ in water) and cultivated at 37 °C, 250 rpm until OD600 reaches 0.6-0.8. Protein expression was induced by addition of 0.1 mM IPTG and supplemented with thiamine hydrochloride (0.1 g/L) and ZnCl₂ (1 mM). Protein expression was carried out at 30 °C, 250 rpm for 20 – 22 h. Cells were harvested by centrifugation (4 °C, 3220 g, 15 min) and stored at -20 °C until further use.

2.8 Activity determination of ReADH enzyme

ReADH enzyme was partially purified by heat treatment. For this procedure, 1 g cell pellet was dissolved in 4 mL 50 mM phosphate buffer pH 7.5 and cell disruption was performed by sonication for 5 min (30 sec. on + 30 sec. off, 40 % amplitude, SONICS, Vibra cell, VCX-130 Frankfurt-Germany). Cell lysate containing soluble enzyme was clarified by centrifugation (21300 g, 20 min, and 4°C). Cleared cell lysate was further incubated at 65 °C for 15 min and centrifuged at 21300 g, 20 min, and 4 °C. 80 % of the *E. coli* proteins could be removed by treating the cell crude extract for 15 min at 65 °C.^[12] Obtained supernatant was used for activity measurements. ReADH activity

assays in cuvette-scale were performed in the Varian Cary 50 Bio UV/vis spectrophotometer (Agilent Technologies, Darmstadt, Germany). 20 µL of substrates in ethanol (10 mM final concentration) were mixed with 50 µL 5 mM NADH (0.25 mM final concentration) and 830 µL 50 mM phosphate buffer at pH 8.0. The reaction was started by addition of 100 µL of partially purified ReADH enzyme. NADH absorption at 340 nm was monitored for 10 min (ϵ = 6200 1/(M cm)). Protein concentration of the heat treated lysate was determined using the PierceTM BCA Protein Assay Kit according to the manufactures instructions (Thermo Fisher Scientific, Darmstadt, Germany). Heat-purified ReADH was diluted accordingly in order to obtain linear absorbance change in the monitored time scale. The linear part of the measured activity curve was used for calculation of enzymatic activity and stated in U/mg total protein concentration in heat-purified lysate. All measurements were performed in triplicate.

3. Results and discussion

The main aim of the study was to understand the structural determinants for the substrate scope of ReADH and its expansion toward bulky substrates for synthesis of highly valuable chiral alcohols. First, we present the construction and evaluation of ReADH homology model. Followed by molecular docking studies of ethyl 2-oxo-4-phenylbutyrate, 2-, 3- or 4-octanone employing our previously developed mechanism based docking protocol. In the next paragraph, *in silico* generation of the ReADH W296A variant and results on docking studies are shown. Finally, the catalytic performance of ReADH WT and ReADH W296A toward octanone derivatives and a shift in regioselectivity of ReADH W296A are discussed.

3.1 Structural modeling of ReADH

The main aim behind the construction of a homology model for ReADH is to get insights into the substrate binding pocket of ReADH. Based on the sequence of ReADH, hybrid homology model was constructed with two different templates (PDB ID: 3JV7;^[20] ADH from *Rhodococcus ruber*, and PDB ID: 1H2B;^[21] ADH from

10.1002/cbic.202000247

Aeropyrum pernix) using YASARA Structure Version 13.9.8.^[17] Details of homology modeling is discussed in SI.

On closer inspection of the active site of ReADH, it was observed that the catalytic zinc ion is coordinated to C38, H62, and D153; the fourth coordination site is occupied by acetic acid (adapted from template structure PDB ID: 3JV7). However, the structural zinc ion is coordinated to four cysteine residues (C92, C95, C98, and C106). The constructed homology model shown in **Figure 1**.



Figure 1. Homodimeric structure of ReADH modeled using YASARA Structure Version 13.9.8.^[17] with the docking pose of acetophenone. The catalytic active site along with the zinc ion (orange sphere) and its coordinating residues (C38, H62, and A153) along with NADH as a cofactor are shown. Substrate (acetophenone) is shown in ball and stick representation. The residue S40 which is involved in proton relay mechanism is also shown.

3.2 Structure-guided analysis of the binding pocket of ReADH

The catalytic residues that are involved in the carbonyl reduction mechanism and maintains the tetravalency of catalytic zinc ion were identified by comparing the constructed ReADH homology model with the crystal structure of cpADH5 (PDB ID: 4C4O).^[25] The structural overlay of homology model of ReADH and cpADH5 is shown in **Figure S4** in SI. Root Mean Square Deviation (RMSD) was found to be 0.61 Å indicating structural similarities between these two structures. It was found that the binding pocket of cpADH5 is completely overlay to the constructed model of ReADH

and cpADH5 shares overall 30.70 % sequence identity (depicted in **Figure S5** in SI). The active site of the optimized model was further examined using YASARA Structure Version 13.9.8.^[17] Previous literatures have reported that the binding pocket of ADHs belonging to zinc-dependent medium chain reductase family consists of small and large binding pockets.^[25b, 26] These binding pocket mainly determines substrate scope and selectivity^[27] of ADHs. The large binding pocket is formed because of the shape of the entrance of substrate from the solvent media toward the catalytic zinc ion. However, the small binding pocket is defined by W296 which acts as a gatekeeper for smaller binding pocket. Conservation analysis showed that W296 is a 70 % conserved residue in the family of zinc-dependent MDRs (depicted in **Figure S6** in SI). The surface view of the identified small and large binding pocket of ReADH is depicted in **Figure 2**. The smaller binding pocket is constituted by the residues C38, D153, T157, Y295, W296 and the larger binding pocket is constituted by the residues S40, F43, I44, Y52, Y54, H62, and L119.



Figure 2. Binding pocket of ReADH consisting of small (red wireframe) and large binding pocket (blue wireframe). NAD⁺ cofactor is shown as stick and catalytic zinc ion as a orange sphere. Residues in the small (red color) and large (blue color) binding

pocket are shown as sticks. The coordinating residues (C38, H63, and D153) to catalytic zinc ion can be seen in the figure.

The catalytic triad (Cys-His-Asp) which is coordinated to catalytic zinc ion is common in both (ReADH and cpADH5). The carbonyl reduction mechanism in cpADH5 was studied by Dhoke *et. al*^[28] and firstly the hydride transfer takes place in carbonyl reduction mechanism followed by the sequential events of proton transfers. Considering the active site similarities in both the enzymes, we assumed a similar carbonyl reduction mechanism for ReADH (depicted in **Figure S7** in SI). The serine residue (S46) which is involved in first proton transfer step (PT1) in cpADH5 is positioned at the same place (S40) in ReADH. In addition to this, histidine residue (H39) is also occupies the similar position and orients toward the riboxyl sugar moiety of NADH in order to transfer third proton transfer (PT3) during the catalytic reduction mechanism. A hydrogen bond network was observed in ReADH between the S40riboxyl sugar of NADH-H39 which is similar to the one in cpADH5.

3.3 Molecular docking of selected substrates

Based on the substrate binding pockets in ReADH, we selected 2-octanone as a candidate substrate for carrying out molecular docking studies. The rationale behind selection of 2-octanone as a candidate substrate was that it has one small (methyl) and one long (hexyl) alkyl chain along with carbonyl functionality. Along with this, substrates which have shown poor activity against MDRs sharing structural similarities in the active sites (cpADH5 and ReADH) were collected from literature^[29] (listed in **Table 1**). The goal is to understand the structure-function relationship using ReADH as a model enzyme and furthermore to design ReADH improved variants. In the literature, it was mentioned that the relative activity (% activity of acetophenone) of these substrates (ethyl 2-oxo-4-phenylbutyrate (EOPB), 2-hyroxyacetopheone, and 3-chloropropiophenone) are lower than 20 % for cpADH5.^[29b] Similarly, *(S)*-carbonyl reductase from *Candida parapsilosis* ATCC 7330 also shows low activity (1.17 \pm 0.19 U/mg) towards EOPB.^[9]

Accepted Manuscrig

Substrate	Structure	Reason of	Reference
		selection	
ethyl 2-oxo-4-		Low activity for	Jakoblinnert
phenylbutyrate (EOPB)		cpADH5	2013 ^[29b]
2-hydroxyacetophenone	HO	Low activity for	Jakoblinnert
(2HACP)		cpADH5	2013 ^[29b]
3-chloropropiophenone	Cl	Low activity for	Jakoblinnert
(3CPP)		cpADH5	2013 ^[29b]
acetophenone (ACP)	< → O O O O O O O O O O O O O	Reference substrate	Schubert 2001 ^[30]
			and Dhoke
			2015 ^[4b]
propiophenone (PPP)		acetophenone	t. w.
		derivative	
2-octanonone (2OCT)		Based on ReADH	t.w.
		binding pocket	
3-octanone (3OCT)		Studying ReADH	t.w.
	\sim \sim \sim \sim	structure-function	
		relatioship	
4-octanone (4OCT)	$\uparrow 0$	Studying ReADH	t.w.
	$\sim \sim \sim \sim$	structure-function	
		relationship	

Table 1. Substrates used in this study along with their structures.

t.w. this work

Additionally, acetophenone was included as a reference substrate as reported for cpADH5.^[4b, 30] Furthermore, propiophenone was included in order to determine the accessable space in the small binding pocket. All the substrates used in this study are listed in **Table 1**. The molecular docking was conducted using our previously developed substrate based docking protocol.^[26] The binding energy obtained from molecular docking of these substrates is listed in **Table S2** in SI. The docking pose of 2OCT and 3OCT in the binding pocket of ReADH is depicted in **Figure 3**.

From the docking pose of 2OCT, it was revealed that the carbonyl oxygen group of both substrates is at a distance of <2 Å (shown in **Figure 3**). This indicates stronger binding of 2OCT to the catalytic zinc ion. As can be seen from the **Figure 3**, when the good substrate 2OCT is in the binding pocket, it positions itself in such way that its methyl group (next to carbonyl functionality) is settled in the small binding pocket. Therefore, we assumed that substitutions at the W296 position could be beneficial for substrates with varied alkyl chain lengths. Therefore, 3OCT (docking pose shown in **Figure 3**) and 4OCT were additionally included as substrates. With ReADH WT, non-catalytic binding modes of 4OCT and EOPB were obtained as 4OCT has butyl and EOPB has ethyl acetate substitution next to their carbonyl functionality, respectively. In both cases, the carbonyl oxygen was far away (D1>4 Å) from the catalytic zinc ion.



Figure 3. Molecular docking pose of **A**) 2-octanone and **B**) 3-octanone in the binding pocket of ReADH; substrates are shown in ball and sticks representations, whereas, all the active site residues are shown in sticks representation. Catalytic zinc ion is shown as orange sphere and residue W296 is shown as magenta sticks. Reciprocal arrows show the distance (D1) between the catalytic zinc ion and carbonyl oxygen of the octanone substrates.

Based on structured based protein engineering of ADH from *Candida parapsilosis* (cpADH5; small binding pocket; residue W286), we showed that by modifying the smaller binding pocket in MDRs, the substrate scope can be controlled.^[31] This is

because the side chain of tryptophan residue (W296) forming the small binding pocket introduces steric hindrance, which prevents high activity of ReADH toward bulky substrates with larger substitution adjacent to their carbonyl functionality. A similar study on ADH from Thermoanaerobacter brockii (TbSADH) has revealed that by mutating the residues (I86A, and W110T) of binding pocket of ADHs not only changes the shape of the binding pocket but also interfere with the substrate-enzyme interactions.^[27] This ultimately can help in increasing the substrate scope toward bulkier substrates. Therefore, we designed in silico the ReADH W296A variant with an assumption that by substituting W296 with a smaller amino acid, the catalytically competent binding orientation of 4OCT and EOPB could be achieved. We increased the space (1300 Å) in the binding pocket of ReADH W296A when compared to ReADH WT (1050 Å). With the increase in space, we achieved the catalytic binding EOPB and 4OCT as shown in Figure S8 in SI. Designed ReADH W296A variant was further used for molecular docking of 4OCT and EOPB. From the docking experiment with ReADH W296A variant, it was observed that for both the substrates (EOPB and 4OCT), the catalytic zinc ion is in close proximity (distance D1<2 Å) to that of the carbonyl oxygen indicating the possibility of activation of carbonyl oxygen by the catalytic zinc ion. The docked pose of EOPB with ReADH WT and ReADH W296A variant is shown in Figure 4. Afterwards, NADH consumption assay was used to convert all these substrates.



14

Figure 4. Molecular docking pose of ethyl 2-oxo-4-phenylbutyrate (EOPB) in the binding pocket of ReADH; (**A**) ReADH WT and (**B**) ReADH W296A variant; substrate is shown in ball and sticks representations whereas all the active site residues are shown in stick representation. Catalytic zinc ion is shown as orange sphere and residue W296 and A296 are shown as magenta sticks. Reciprocal arrow shows distance (D1) between catalytic zinc ion and carbonyl oxygen of the substrate.

3.4 Screening of SSM W296 library toward ethyl 2-oxo-4-phenylbutyrate and activity measurement of ReADH variants

Activity measurements of ReADH WT toward selected substrates EOPB, 2HACP, 3CPP, ACP, PPP, 2OCT, 3OCT, and 4OCT were performed to gain molecular insights and to evaluate the molecular docking simulations. ReADH WT was expressed and partially purified through heat treatment; purity was checked by SDS-PAGE analysis (**Figure S9** in SI). Activity toward various substrates was determined for cleared cell lysates, heat treated samples and no significant activity change was observed for both samples (**Figure S10** in SI). We could not determine activity accurately toward 2HACP, 3CPP and PPP due to low solubility of substrates during activity measurement, and thus they were not further considered.

As computational prediction shows that W296 is a key player for the catalytic binding of EOPB, we generated a SSM library at position W296. The rationale behind generation of SSM library at W296 position was to explore the full diversity and to gain deeper knowledge and to understand structure-function relationship of ReADH.

One hundred eighty clones (≥ 95 % coverage^[32] were screened through the NADH depletion assay for reduction of EOPB. Ten beneficial clones were sequenced and three different substitutions were obtained (alanine, glycine, and cysteine). Shake flask expression was performed for all three variants and specific activities were determined after partial purification. ReADH W296A variant showed a high activity toward EOPB when compared to ReADH W296C and W296G. ReADH W296A, W296C, and W296G variants showed 3.6, 3.0, and 1.9 fold increased specific activity when compared to ReADH WT: 4.7 U/mg; W296A:17.10 U/mg; W296C: 14.4; and W296G: 9.1 U/mg). All three amino acids (alanine, glycine, and cysteine) have side chains with a low steric demand when compared to the indole ring of tryptophan.

A similar effect was also observed for cpADH5 toward methyl 3-hydroxyhexanoate and methyl 3-hydroxyoctanote when tryptophan residue at position 286 in small binding pocket was substituted to alanine.[31] Kinetic results of ReADH W296A, W296C, and W296G confirmed the computational analysis that Trp acts as a gatekeeper to the small binding pocket. EOPB could not be accommodate in the binding pocket of ReADH WT due to its size. Furthermore, specific activity toward acetophenone which was used as reference substrate for cpADH5.^[4b] was determined for ReADH W296A, W296C, and W296G. Obtained variants showed a drastic activity decreases toward acetophenone (ReADH WT 4.3 U/mg, W296A 0.1 U/mg activity). Because W296A has shown to have the highest activity among the improved variants, it was chosen for further analysis. Specific activities of the ReADH WT and W296A were determined toward 2OCT, 3OCT, and 4OCT in order to elucidate the observed binding of these substrates in the active site. ReADH WT shows a gradually decreased activity toward 3OCT and 4OCT (See Figure 5). ReADH W296A variant has an opposite preference; W296A's activity gets increased by change of carbonyl oxygen position on octanone chain. The ReADH W296A variant has more than 16 fold activity increase toward 4OCT when compared to the ReADH WT. A shift in the regioselectivity is observed for the ReADH W296A variant. Whereas ReADH WT accommodates small chemical groups like a methyl group in acetophenones and 2octanone, the ReADH W296A variant accepts in its small binding pocket larger chemical groups such as propyl moiety of 4OCT.



Figure 5. Specific activities of ReADH WT and improved variants for reduction of acetophenone, ethyl 2-oxo-4-phenylbutyrate, 2-octanone, 3-octanone, and 4-octanone. All activity measurements were performed with heat-purified lysate at 5 mM substrate concentration and NADH depletion was monitored at 340 nm.; WT, black; W296A, red; W296C, blue and W296G, green. Specific activities of ReADH WT and W296A variant for reduction of 2-octanone, 3-octanone, and 4-octanone are shown on embedded figure. WT, black; W296A, gray.

3.5 Structural determinants of ReADH regioselectivity

Different carbonyl substrates were selected and investigated experimentally to explore full substrate scope of ReADH. As discussed earlier, ReADH has a small as well as large binding pocket and smaller binding pocket restricts the entry of substrates which has larger substitutuent than ethyl group next to the carbonyl functionality. There is a significant similarity between binding pocket of cpADH5 and ReADH, therefore it is highly likely that butyraldehydes^[4b] can also be considered as the natural substrate of ReADH, which has one carbon on smaller aliphatic chain. The selected substrate 20CT orientates catalytically in the binding pocket of ReADH in such a way that its smaller aliphatic chain having methyl group is positioned in the small binding pocket as shown in **Figure 3A**. Therefore, it has been proposed that the substituting residue W296 with smaller amino acid (e.g. alanine) increases the volume inside the small binding pocket which gives information on the definition and the boundaries of this pocket of ReADH WT toward the substrates with different chain lengths. Thus, 30CT and 40CT from the same octanone class with different positioning of carbonyl functionality were tested to understand the structure-function relationship of ReADH. 30CT and 40CT along with 20CT were tested experimentally and the measured activity were depicted in **Figure 5 and Figure S11** in SI.

ReADH WT shows higher activity toward 2OCT compared to ReADH W296A variant as there is space only for methyl or ethyl group in the smaller binding pocket. In case of 3OCT and 4OCT, ReADH WT showed less activity compared to 2OCT (depicted in **Figure 5**) because of the increase in the side chain length. It was found that on increasing the length of aliphatic side chain next to carbonyl moiety on octanone chain, the activity of ReADH WT decreases. ReADH WT shows gradually decreased activity toward 3OCT and 4OCT. Therefore, the smaller binding pocket of ReADH WT enzyme was engineered to understand the substrate scope. The improved ReADH W296A variant obtained after SSM at W296 position was used further for this purpose. Molecular docking simulations also revealed that by substituting W296 position with smaller amino acids, more space can be accessible for substrates having larger than ethyl group next to the carbonyl functionality. This can be clearly seen from docking pose of 4OCT depicted in **Figure 6**.



Figure 6. Molecular docking pose of 4-octanone in the binding pocket of ReADH; ReADH W296A variant; substrate is shown in ball and sticks representation, whereas all the active site residues are shown in sticks representation. Catalytic zinc ion is shown as orange sphere and residue A296 are shown in magenta sticks. Reciprocal arrow shows distance (D1) between catalytic zinc ion and carbonyl oxygen of the substrate.

From our experimental data, it was found that the improved ReADH W296A variant showed completely opposite trend in activity toward 2OCT, 3OCT, and 4OCT compared to ReADH WT. The activity has increased by change of carbonyl oxygen position. For 4OCT, the activity increases more than a 16 fold for ReADH W296A variant compared to the ReADH WT. Inversion of the regioselectivity is observed for the ReADH W296A variant; instead of favoring small chemical groups like a methyl group in acetophenone and 2OCT, ReADH W296A variant favors larger chemical groups (i.e. propyl) in its small binding pocket. We demonstrate that the introduced substitutions in the smaller binding pocket induces systematic change in the catalytically competent substrate binding and the interactions of the carbonyl oxygen with the catalytic zinc ion and thus determining the stereopreferences and regioselectivity of the ReADH variants. ReADH follows Prelog's rule and convert

carbonyl compounds to their S-enantiomeric alcohols. ^[12, 14] However, generated ReADH W296A variant possibly could show reversal in enantioselectivity for smaller substrates accepted by ReADH WT, since ReADH W296A variant has larger substrate binding pocket which allows the substrate to explore *pro*-(R) conformations with catalytically active hydride transfer distances. This phenomenon was observed almost in all previously engineered ketoreductases such as cpADH5 ^[31] and TbSADH ^[27, 33].

4. Conclusions

Computational assisted ReADH engineering revealed that enlarging the smaller binding pocket enables ReADH to convert bulkier substrates (ethyl 2-oxo-4-phenylbutyrate, 3- or 4-octanone) with larger alkyl chains adjacent to their carbonyl functionality. Molecular docking of substrates to ReADH WT showed that W296 is critical in controlling the substrate scope. SSM study at W296 position yielded impoved ReADH W296A variant which confirms our computational prediction. Substitution of W296 with smaller amino acid provided sufficient space for the catalytically competent binding of ethyl 2-oxo-4-phenylbutyrate and 4-octanone substrates.

From our point of view comprehensive study on ReADH provides a molecular basis to expand substrate scope and regioselectivity of ReADH. This study would open novel synthetic routes for synthesis of an important intermediate for anti-hypertension drugs like enalaprilat and lisinopril. Based on the structural determinants, it is very likely that the obtained reengineering knowledge can be transferred to other zinc-dependent ADHs with small and large binding pocket.

Abbreviations

- ADH Alcohol Dehydrogenase
- ReADH Alcohol Dehydrogenase from Rhodococcus erythropolis
- cpADH5 ADH5 from Candida parapsilosis
- HT Hydride transfer
- PT Proton transfer
- EOPB ethyl 2-oxo-4-phenylbutyrate
- 2HACP 2-hydroxyacetophenone)
- 3CPP 3-chloropropiophenone
- ACP acetophenone
- PPP propiophenone
- 20CT 2-octanonone
- 3OCT 3-octanone
- 4OCT 4-octanone

ASSOCIATED CONTENT

Supporting Information

Supporting information includes homotetrameric structure of ReADH, details of structural modeling of ReADH, Catalytically applied distance filter criteria, table with molecular docking energy of the substrates, sequence as well as structural alignment of cpADH5 and ReADH, conservation analysis of ReADH followed by proposed catalytic mechanism of ReADH, calculated binding pockets, SDS-PAGE analysis of ReADH WT protein and Comparison of activity of crude cell lysate and partially purified ReADH WT protein.

AUTHOR INFORMATION

Corresponding Author

m.davari@biotec.rwth-aachen.de (M. D. Davari)

Notes:

The authors declare no competing financial interest.

ACKNOWLEDGMENT

M.D.D and G.D. would like to acknowledge DFG SPP 1166 BioNoCo for financial support. Y.E. thanks Kafkas University, Kars, Turkey for funding. Simulations were performed with computing resources granted by JARA-HPC from RWTH Aachen University under project RWTH0116.

References

- [1] a) F. Secundo and R. S. Phillips, *Enzyme Microb. Technol.* **1996**, *19*, 487-492; b)
 W. Hummel and M. R. KULA, *FEBS J.* **1989**, *184*, 1-13; c) C. L. Stone, T.-K. Li and
 W. Bosron, *J. Biol. Chem.* **1989**, *264*, 11112-11116.
- [2] a) R. N. Patel, *Adv. Appl. Microbiol.* **1997**, *43*, 91-140; b) J. W. Hilborn, Z.-H. Lu, A.
 R. Jurgens, Q. K. Fang, P. Byers, S. A. Wald and C. H. Senanayake, *Tetrahedron Lett.* **2001**, *42*, 8919-8921.
- [3] a) N. Itoh, N. Mizuguchi and M. Mabuchi, *J. Mol. Catal. B: Enzym.* **1999**, *6*, 41-50;
 b) J. C. Wang, M. Sakakibara, J. Q. Liu, T. Dairi and N. Itoh, *Appl. Microbiol. Biotechnol.* **1999**, *52*, 386-392.
- [4] a) H. Yamamoto, N. Kawada, A. Matsuyama and Y. Kobayashi, *Biosci. Biotechnol. Biochem* 1999, 63, 1051-1055; b) G. V. Dhoke, Y. Ensari, M. D. Davari, A. J. Ruff, U. Schwaneberg and M. Bocola, *J. Chem. Inf. Model.* 2016, *56*, 1313-1323.
- [5] W. Hummel in New alcohol dehydrogenases for the synthesis of chiral compounds, Vol. 58 Springer, 1997, pp. 145-184.
- [6] D. Alsafadi, S. Alsalman and F. Paradisi, Org. Biomol. Chem. 2017, 15, 9169-9175.
- [7] J. Cassidy, L. Bruen, E. Rosini, G. Molla, L. Pollegioni and F. Paradisi, *PloS one* 2017, 12, e0187482.
- [8] a) Y. Xie, J.-H. Xu and Y. Xu, *Bioresour. Technol.* 2010, 101, 1054-1059; b) Y. Ni,
 C.-X. Li, H.-M. Ma, J. Zhang and J.-H. Xu, *Appl. Microbiol. Biotechnol.* 2011, 89, 1111-1118; c) Y. Ni, C. X. Li, J. Zhang, N. D. Shen, U. T. Bornscheuer and J. H. Xu, *Adv. Synth. Catal.* 2011, 353, 1213-1217.
- [9] S. Sudhakara and A. Chadha, Org. Biomol. Chem. 2017, 15, 4165-4171.

10.1002/cbic.202000247

- [10] a) S. Oda, Y. Inada, A. Kobayashi and H. Ohta, *Biosci. Biotechnol. Biochem.* 1998,
 62, 1762-1767; b) W.-Q. Lin, Z. He, Y. Jing, X. Cui, H. Liu and A.-Q. Mi, *Tetrahedron: Asymmetry* 2001, *12*, 1583-1587.
- [11] C. M. Nealon, M. M. Musa, J. M. Patel and R. S. Phillips, ACS Catalysis 2015, 5, 2100-2114.
- [12] K. Abokitse and W. Hummel, Appl. Microbiol. Biotechnol. 2003, 62, 380-386.
- [13] V. Prelog, *Pure Appl. Chem.* **1964**, *9*, 119-130.
- [14] H. Gröger, W. Hummel, C. Rollmann, F. Chamouleau, H. Hüsken, H. Werner, C.
 Wunderlich, K. Abokitse, K. Drauz and S. Buchholz, *Tetrahedron* 2004, 60, 633-640.
- [15] J. Kasprzak, F. Bischoff, M. Rauter, K. Becker, K. Baronian, R. d. Bode, F. Schauer, H. M. Vorbrodt and G. Kunze, *Biochem. Eng. J.* 2016, 106, 107-117.
- [16] C. UniProt, Nucleic Acids Res. 2004, 45, D158-D169.
- [17] a) E. Krieger, G. Vriend and C. Spronk, YASARA. org 993; b) E. Krieger and G. Vriend, *Bioinformatics* 2014, 30, 2981-2982.
- [18] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Res.* **1997**, 25, 3389-3402.
- [19] J. D. Thompson, D. G. Higgins and T. J. Gibson, *Nucleic Acids Res.* **1994**, 22, 4673-4680.
- [20] M. Karabec, A. Łyskowski, K. C. Tauber, G. Steinkellner, W. Kroutil, G. Grogan and K. Gruber, *Chem. Commun.* **2010**, *46*, 6314-6316.
- [21] L. Esposito, F. Sica, C. A. Raia, A. Giordano, M. Rossi, L. Mazzarella and A. Zagari, J. Mol. Biol. 2002, 318, 463-477.
- [22] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785-2791.

24

[23] A. Jakalian, D. B. Jack and C. I. Bayly, J. Comput. Chem. 2002, 23, 1623-1641.

- [24] A. A. Canutescu, A. A. Shelenkov and R. L. Dunbrack, *Protein Sci.* 2003, 12, 2001-2014.
- [25] a) H. Man, C. Loderer, M. B. Ansorge-Schumacher and G. Grogan, *ChemCatChem* 2014, *6*, 1103-1111; b) S. Wang, Y. Nie, Y. Xu, R. Zhang, T.-P.
 Ko, C.-H. Huang, H.-C. Chan, R.-T. Guo and R. Xiao, *Chem. Commun.* 2014, *50*, 7770-7772.
- [26] G. V. Dhoke, C. Loderer, M. D. Davari, M. Ansorge-Schumacher, U. Schwaneberg and M. Bocola, J. Comput. Aided Mol. Des. 2015, 29, 1057-1069.
- [27] M. A. Maria-Solano, A. Romero-Rivera and S. Osuna, Org. Biomol. Chem. 2017, 15, 4122-4129.
- [28] G. V. Dhoke, M. D. Davari, U. Schwaneberg and M. Bocola, ACS Catalysis 2015, 5, 3207-3215.
- [29] a) A. Jakoblinnert, R. Mladenov, A. Paul, F. Sibilla, U. Schwaneberg, M. B. Ansorge-Schumacher and P. D. de María, *Chem. Commun.* 2011, 47, 12230-12232; b) A. Jakoblinnert, J. Wachtmeister, L. Schukur, A. V. Shivange, M. Bocola, M. B. Ansorge-Schumacher and U. Schwaneberg, *Protein Eng. Des. Sel.* 2013, 26, 291–298.
- [30] T. Schubert, W. Hummel, M.-R. Kula and M. Müller, *Eur. J. Org. Chem.* 2001, 2001, 4181-4187.
- [31] Y. Ensari, G. V. Dhoke, M. D. Davari, M. Bocola, A. J. Ruff and U. Schwaneberg, *Chem. Eur. J.* 2017, 23, 12636-12645.
- [32] A. E. Firth and W. M. Patrick, Nucleic Acids Res. 2008, 36, W281-W285.
- [33] B. Liu, G. Qu, J. Li, W. Fan, J. Ma, Y. Xu, Y. Nie, and Z. Sun, *Adv. Synth. Catal.***2019**, *361*, 3182-3190.

Graphical Abstract



A versatile ADH from *Rhodococcus erythropolis* (ReADH) accepts only carbonyl substrates with either methyl or ethyl group adjacent to carbonyl moiety. Modified ReADH converts bulky substrates (ethyl 2-oxo-4-phenylbutyrate) and opens up a new route for synthesis of ethyl-2-hydroxy-4-phenylbutanoate, which is an important intermediate for anti-hypertension drugs like enalaprilat and lisinopril.