

Reversal of Regioselectivity in Zinc-Dependent Medium-Chain Alcohol Dehydrogenase from *Rhodococcus erythropolis* toward Octanone Derivatives

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

The zinc-dependent medium-chain alcohol dehydrogenase from *Rhodococcus erythropolis* (ReADH) is one of the most versatile biocatalysts for the stereoselective reduction of ketones to chiral alcohols. Despite its known broad substrate scope, ReADH only accepts carbonyl substrates with either a methyl or an ethyl group adjacent to the carbonyl moiety; this limits its use in the synthesis of the chiral alcohols that 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-4-phenylbutyrate) opens up new routes in the synthesis of ethyl-2-hydroxy-4-phenylbutanoate, 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, sterically restricts

the access of ethyl 2-oxo-4-phenylbutyrate, octan-3-one or octan-4-one 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 WT ReADH (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 octan-4-one (ReADH W296A: 0.97 U/mg and ReADH WT: 0.06 U/mg) and a more than 30-fold decreased activity toward octan-2-one (ReADH W296A: 0.23 U/mg and ReADH WT: 7.69 U/mg). Computational and experimental results revealed the role of position W296 in controlling the substrate scope and regiopreference of ReADH for a variety of carbonyl substrates.

Introduction

Secondary alcohol dehydrogenases (ADHs) belong 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. Stereoselectivity is a unique characteristic

among secondary alcohol dehydrogenases^[1] that makes them of great interest for the biocatalytic production of chiral alcohols as a building blocks for the fine chemical industry.^[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 capable 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 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 *C. parapsilosis* ATCC 7330^[9] reduce 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 being able to reduce a broad variety of substrates, 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 emerged as a promising candidate for use in the industry.^[5,12] ReADH has a

[a] Dr. G. V. Dhoke,⁺ Dr. Y. Ensari,⁺ D. Y. Hacibaloglu, A. Gärtner, Dr. A. J. Ruff, Dr. M. Bocola, Dr. M. D. Davari
Lehrstuhl für Biotechnologie, RWTH Aachen University
Worringerweg 3, 52074 Aachen (Germany)
E-mail: m.davari@biotec.rwth-aachen.de

[b] Dr. Y. Ensari⁺
Kafkas University, Faculty of Engineering and Architecture
Department of Bioengineering
full address?, Kars (Turkey)

[*] These authors contributed equally to this work.



Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cbic.202000247> including 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 WT ReADH protein and comparison of activity of crude cell lysate and partially purified WT ReADH protein.



© 2020 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

homodimeric structure with a molecular weight of 36026 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 Prelog's rule,^[13] which leads to formation of *S* enantiomers with a high purity ($ee \geq 99\%$).^[12,14] ReADH has a high stereo-, chemo-, and regioselectivity towards variety of carbonyl substrates including acetophenone derivatives and aliphatic ketones, and can withstand temperatures of up to 65 °C.^[12,14] Recently, Kasprzak et al.^[15] have synthesized (*S*)-1-phenyl ethanol and ethyl (*R*)-4-chloro-3-hydroxybutanoate by employing a biphasic and a substrate-coupled cofactor regeneration system. ReADH reduces carbonyl substrates that have either a methyl or an ethyl group adjacent to carbonyl functionality.^[14] No crystal structure has 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 octan-4-one), we re-engineered the smaller binding pocket of ReADH. Thereby, we constructed a homology model of ReADH and used it for the molecular docking of carbonyl substrates to identify W296 as a key position in modulating substrate access to the catalytic zinc center. Identification was performed with our previously developed docking protocol.^[4b]

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 the synthesis of highly valuable chiral alcohols. First, we present the construction and evaluation of the ReADH homology model, then molecular docking studies of ethyl 2-oxo-4-phenylbutyrate, octan-2-one, octan-3-one, or octan-4-one by 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 WT ReADH and ReADH W296A toward octanone derivatives and a shift in regioselectivity of ReADH W296A are discussed.

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, a hybrid homology model was constructed with two different templates (ADH from *Rhodococcus ruber*, PDB ID: 3JV7,^[16] and ADH from *Aeropyrum pernix*, PDB ID: 1H2B^[17]) by using YASARA Structure Version 13.9.8.^[18] Details of the homology modeling are discussed in the Supporting Information.

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.

Structure-guided analysis of the binding pocket of ReADH

The catalytic residues that are involved in the carbonyl reduction mechanism and maintain the tetravalency of catalytic zinc ion were identified by comparing the constructed ReADH homology model with the crystal structure of cpADH5 (PDB ID: 4C40).^[19] A structural overlay of the homology model of ReADH and cpADH5 is shown in Figure S4 in the Supporting Information. The root mean square deviation (RMSD) was found to be 0.61 Å, thus indicating structural similarities between these two structures. It was found that the binding pocket of cpADH5 completely overlies the constructed model of ReADH, and cpADH5 shares overall 30.70% sequence identity (Figure S5).

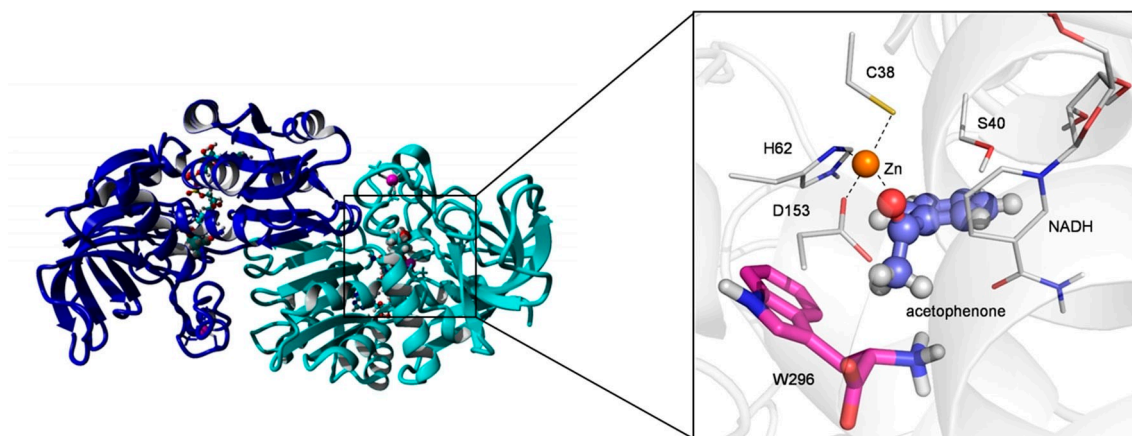


Figure 1. Homodimeric structure of ReADH modeled in YASARA structure version 13.9.8.^[18] with the docking pose of acetophenone. The catalytic active site along with the zinc ion (orange sphere) and its coordinating residues (C38, H62, and A153) along with NADH as a cofactor are shown. The substrate (acetophenone) is shown in ball and stick representation. The residue S40, which is involved in a proton relay mechanism, is also shown.

The active site of the optimized model was further examined by using YASARA.^[18] Previous literature has reported that the binding pocket of ADHs belonging to zinc-dependent medium-chain reductase family consists of small and large binding pockets.^[19b,20] These binding pockets largely determine the substrate scope and selectivity^[21] of ADHs. The large binding pocket is formed because of the shape of the entrance for the substrate from the solvent toward the catalytic zinc ion. The small binding pocket is defined by W296, which acts as a gatekeeper for this pocket. Conservation analysis showed that W296 is a 70% conserved residue in the family of zinc-dependent MDRs (Figure S6). A surface view of the small and large binding pockets of ReADH is given in Figure 2. The smaller

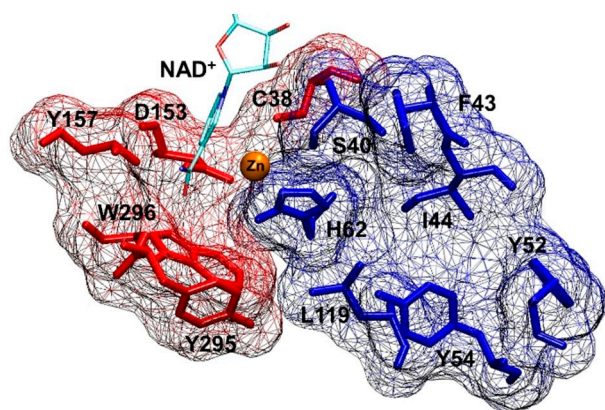


Figure 2. Binding pocket of ReADH consisting of small (red wireframe) and large binding pockets (blue wireframe). The NAD^+ cofactor is shown as a cyan stick, and the catalytic zinc ion as an orange sphere. Residues in the small (red) and large (blue) binding pocket are shown as sticks. The residues coordinating to the catalytic zinc ion (C38, H63, and D153) can be seen in the figure.

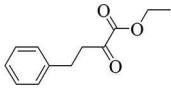
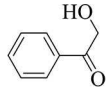
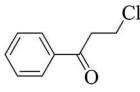
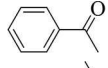
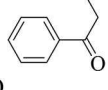
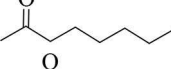
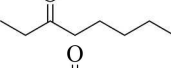
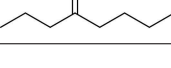
binding pocket is made up of residues C38, D153, T157, Y295, and W296; residues S40, F43, I44, Y52, Y54, H62, and L119 constituted the larger binding pocket.

The catalytic triad (Cys-His-Asp) that is coordinated to the catalytic zinc ion is common to both ReADH and cpADH5. The carbonyl reduction mechanism in cpADH5 was studied by Dhoke et al.^[22] First hydride transfer takes place in a carbonyl reduction mechanism, then the sequential events of proton transfer. Considering the similarities between the active sites in both the enzymes, we assumed a similar carbonyl reduction mechanism for ReADH (Figure S7). The serine residue (S46) that is involved in first proton transfer step (PT1) in cpADH5 is positioned at the same place (S40) in ReADH. In addition, the histidine residue (H39) occupies a similar position and orients toward the riboxyl sugar moiety of NADH in order to transfer a third proton (PT3) during the catalytic reduction mechanism. A hydrogen-bond network was observed in ReADH between S40, the riboxyl sugar of NADH, and H39 that is similar to the one in cpADH5.

Molecular docking of selected substrates

Based on the substrate-binding pockets in ReADH, we selected octan-2-one as a candidate substrate to carry out molecular docking studies. The rationale behind this selection was that it has one small (methyl) and one long (hexyl) alkyl chain along with carbonyl functionality. Along with this, substrates that have shown poor activity against MDRs sharing structural similarities in their active sites (cpADH5 and ReADH) were collected from literature^[23] (Table 1). The goal was to understand the structure-function relationship by using ReADH as a model enzyme and furthermore to design improved ReADH

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

Substrate	Structure	Reason for selection	Ref.
ethyl 2-oxo-4-phenylbutyrate (EOPB)		low activity for cpADH5	[23b]
2-hydroxyacetophenone (2HACP)		low activity for cpADH5	[23b]
3-chloropropiophenone (3CPP)		low activity for cpADH5	[23b]
acetophenone (ACP)		reference substrate	[4b,24]
propiophenone (PPP)		acetophenone derivative	this work
octan-2-one (2OCT)		based on ReADH binding pocket	this work
octan-3-one (3OCT)		studying ReADH structure-function relationship	this work
octan-4-one (4OCT)		studying ReADH structure-function relationship	this work

variants. In the literature, it was mentioned that the relative activity (% activity of acetophenone) of these substrates (EOPB, 2-hydroxyacetophenone, and 3-chloropropiophenone) are lower than 20% for cpADH5.^[23b] Similarly, (*S*)-carbonyl reductase from *C. parapsilosis* ATCC 7330 also shows low activity ($1.17 \pm 0.19 \text{ U mg}^{-1}$) towards EOPB.^[9]

Additionally, acetophenone was included as a reference substrate as reported for cpADH5;^[4b,24] propiophenone was included in order to determine the accessible space in the small binding pocket. Molecular docking was conducted by using our previously developed substrate-based docking protocol.^[20] The binding energy obtained from molecular docking of these substrates is listed in Table S2. The docking poses of 2OCT and 3OCT in the binding pocket of ReADH are depicted in Figure 3.

The docking pose of 2OCT revealed that the carbonyl oxygen of both substrates is at a distance of $< 2 \text{ \AA}$ from the zinc (Figure 3). This indicates strong binding of 2OCT to the catalytic zinc ion. As can be seen from Figure 3, when the good substrate 2OCT is in the binding pocket, it positions itself in such a way that its methyl group (next to the carbonyl) is in the small binding pocket. Thus, we assumed that substitutions at W296 could be beneficial for substrates with different alkyl chain lengths. Therefore, 3OCT (docking pose shown in Figure 3) and 4OCT were also included as substrates. With ReADH WT, noncatalytic binding modes of 4OCT and EOPB were obtained, as they have butyl and ethyl acetate next to their carbonyl functionality, respectively. In both cases, the carbonyl oxygen was far away ($D1 > 4 \text{ \AA}$) from the catalytic zinc ion.

Based on structure-based protein engineering of ADH from *C. parapsilosis* (cpADH5; small binding pocket; residue W286), we showed that the substrate scope can be controlled by modifying the smaller binding pocket in MDRs.^[25] This is because the side chain of W296 forming the small binding pocket introduces steric hindrance that prevents high activity of

ReADH toward bulky substrates with a larger substitution adjacent to their carbonyl functionality. A similar study on ADH from *Thermoanaerobacter brockii* (TbSADH) has revealed that mutating the residues (I86 A and W110T) of the binding pocket of ADHs not only changes its shape but also interferes with substrate–enzyme interactions.^[21] This can ultimately help in increasing the substrate scope toward bulkier substrates. Therefore, we designed ReADH variant W296A *in silico* on the 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 in the binding pocket of W296A ReADH (1300 \AA) compared to the WT (1050 \AA). With the increase in space, we achieved catalytic binding of EOPB and 4OCT (Figure S8). The W296A variant was further used for molecular docking of 4OCT and EOPB, from which it was observed that, for both substrates, the catalytic zinc ion is in close proximity ($D1 < 2 \text{ \AA}$) to the carbonyl oxygen, thus indicating the possibility of the zinc ion activating the carbonyl oxygen. The docked pose of EOPB with WT and W296A ReADH is shown in Figure 4. Afterwards, a NADH consumption assay was used to convert all these substrates.

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

Activity measurements of WT ReADH 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. WT ReADH was expressed and partially purified through heat treatment; its purity was checked by SDS-PAGE analysis (Figure S9). The activity toward various substrates was determined for cleared cell lysates, heat-treated

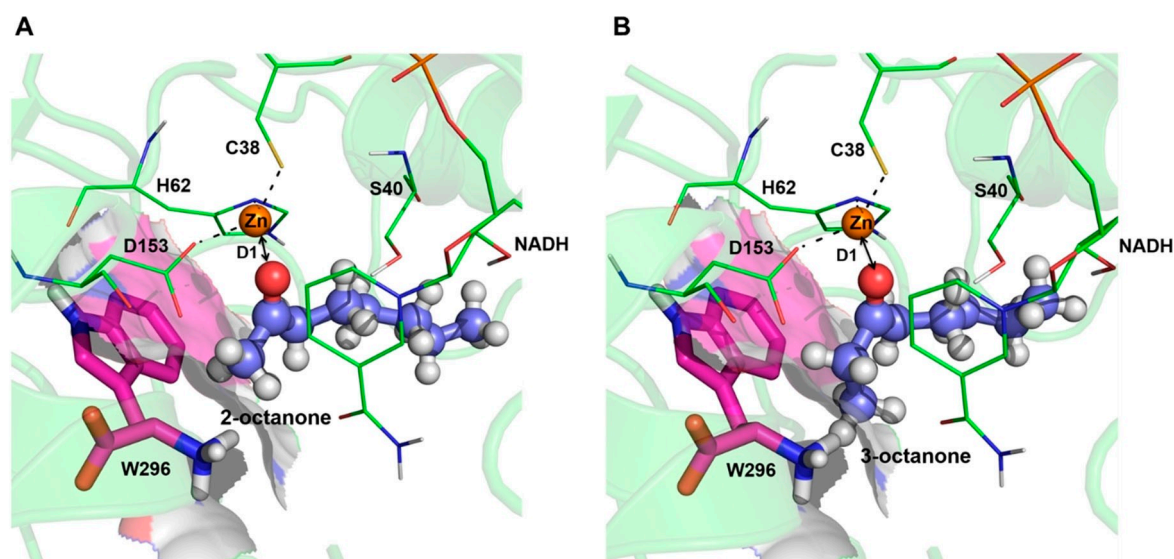


Figure 3. Molecular docking poses of A) octan-2-one and B) octan-3-one in the binding pocket of ReADH; substrates are shown in ball and stick representations; the active-site residues are shown as sticks. The zinc ion is shown as orange sphere, and residue W296 is shown as magenta sticks. Reciprocal arrows show the distance ($D1$) between the zinc ion and carbonyl oxygen of the substrate.

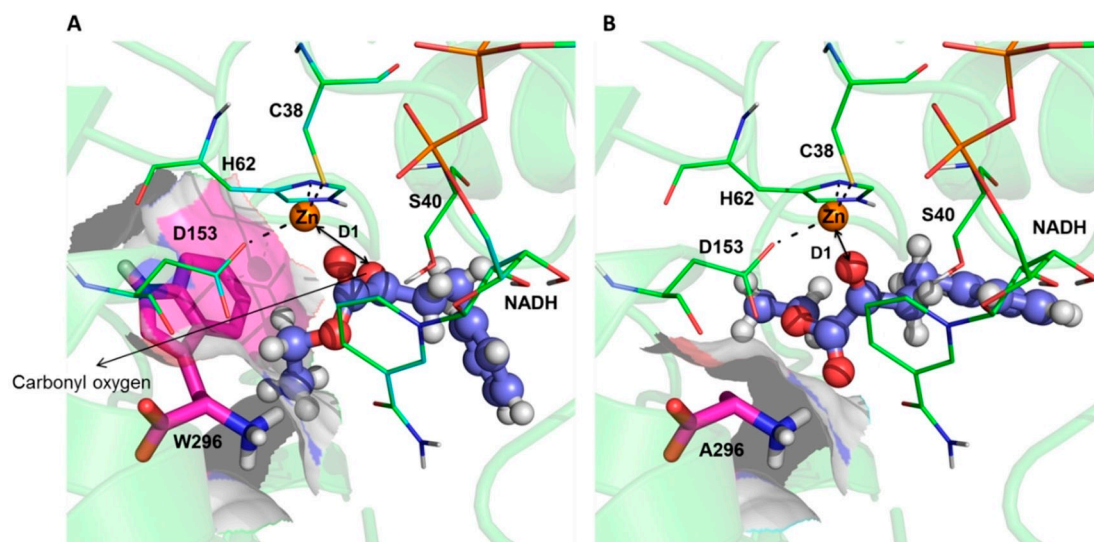


Figure 4. Molecular docking poses of EOPB in the binding pockets of A) WT and B) W296A ReADH. The substrates are shown in ball and stick representations, whereas the active-site residues are shown in stick representation. The zinc ion is shown as an orange sphere, and W296 and A296 are shown as magenta sticks. Reciprocal arrows show the distance ($D1$) between the catalytic zinc ion and the carbonyl oxygen of the substrate.

samples, and no significant activity change was observed for either sample (Figure S10). We could not determine the activity accurately toward 2HACP, 3CPP, and PPP due to the low solubility of the substrates during activity measurement, and thus they were not considered further.

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

One hundred eighty clones ($\geq 95\%$ coverage^[26]) were screened in a NADH-depletion assay for the 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. W296A ReADH showed higher activity toward EOPB than W296C or W296G. W296A (17.10 U mg^{-1}), W296C (14.4 U mg^{-1}), and W296G (9.1 U mg^{-1}) ReADH showed 3.6, 3.0, and 1.9-fold increased specific activity, respectively, compared to WT ReADH (4.7 U mg^{-1}). All three amino acids (alanine, glycine, and cysteine) have side chains with a low steric demand compared to the indole ring of tryptophan. A similar effect was also observed for cpADH5 toward methyl 3-hydroxyhexanoate and methyl 3-hydroxyoctanoate when alanine is substituted for Trp286 in the small binding pocket.^[25] The kinetic results for W296A, W296C, and W296G ReADH confirmed the computational analysis that tryptophan acts as a gatekeeper to the small binding pocket. EOPB could not be accommodated in the binding pocket of WT ReADH due to its size.

Furthermore, the specific activity toward acetophenone, which was used as reference substrate for cpADH5,^[4b] was determined for W296A, W296C, and W296G ReADH. These

variants showed drastic decreases in activity toward acetophenone (WT 4.3 U mg^{-1} , W296A 0.1 U mg^{-1}). Because W296A has been shown to have the highest activity among the improved variants, it was chosen for further analysis. The specific activities of the WT and W296A ReADH were determined toward 2OCT, 3OCT, and 4OCT in order to elucidate the observed binding of these substrates in the active site. WT ReADH shows a gradually decreased activity toward 3OCT and 4OCT (Figure 5). W296A ReADH has an opposite preference; W296A's activity is increased by changing the position of the carbonyl oxygen on the octanone chain. W296A ReADH has a >16 -fold increase activity toward 4OCT compared to the WT. A shift in the regioselectivity is observed for W296A ReADH. Whereas WT ReADH accommodates small chemical groups like methyl in acetophenones and octan-2-one, the W296A variant accepts larger chemical groups such as the propyl moiety of 4OCT in its small binding pocket.

Structural determinants of ReADH regioselectivity

Different carbonyl substrates were selected and investigated experimentally to explore the full substrate scope of ReADH. As discussed above, ReADH has a small and a large binding pocket; the smaller binding pocket restricts the entry of substrates that have a substituent larger than an ethyl group next to the carbonyl functionality. There is a significant similarity between the binding pockets of cpADH5 and ReADH, therefore it is highly likely that butyraldehydes^[4b] can also be considered as natural substrates of ReADH, which usually have one carbon on their smaller aliphatic chain. The selected substrate 2OCT orientates catalytically in the binding pocket of ReADH in such a way that its smaller aliphatic chain, a methyl group, is positioned in the small binding pocket (Figure 3A). Therefore, it

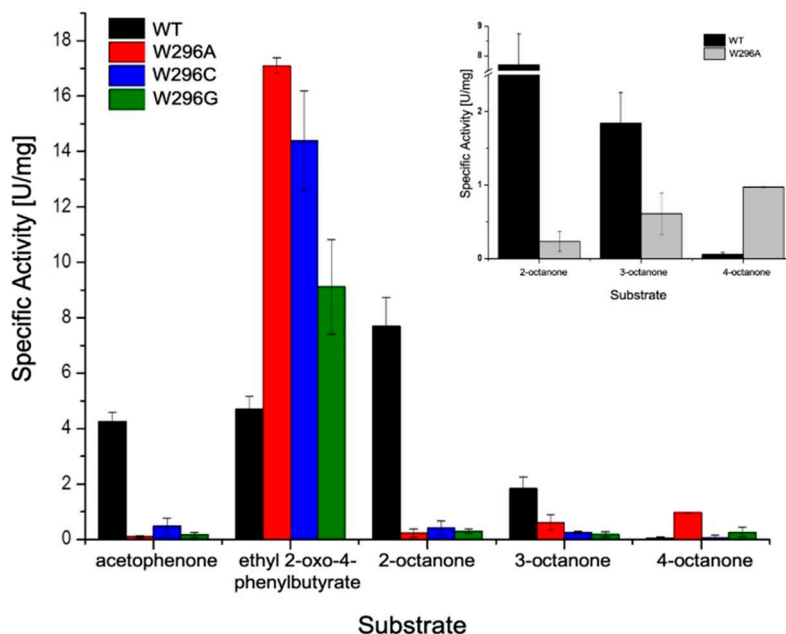


Figure 5. Specific activities of WT and improved variant ReADHs for the reduction of acetophenone, ethyl 2-oxo-4-phenylbutyrate, octan-2-one, octan-3-one, and octan-4-one. All activity measurements were performed with heat-purified lysate and 5 mM substrate; NADH depletion was monitored at 340 nm. The specific activities of WT and W296A ReADH for the reduction of octan-2-one, octan-3-one, and octan-4-one are shown in the inset.

has been proposed that substituting residue W296 with a smaller amino acid (e.g., alanine) increases the volume inside the small binding pocket, and this gives information on the definition and the boundaries of this pocket in WT ReADH toward substrates with different chain lengths. Thus, 3OCT and 4OCT, which are from the same octanone class with different positioning of carbonyl functionality, were tested to understand the structure-function relationship of ReADH. 3OCT and 4OCT along with 2OCT were tested experimentally, and the measured activities are depicted in Figures 5 and S11.

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

From our experimental data, it was found that the improved W296A variant showed completely the opposite trend in activity toward 2OCT, 3OCT, and 4OCT to WT ReADH. The activity increased with the change of carbonyl oxygen position; for 4OCT, the activity increases more than 16-fold. Inversion of

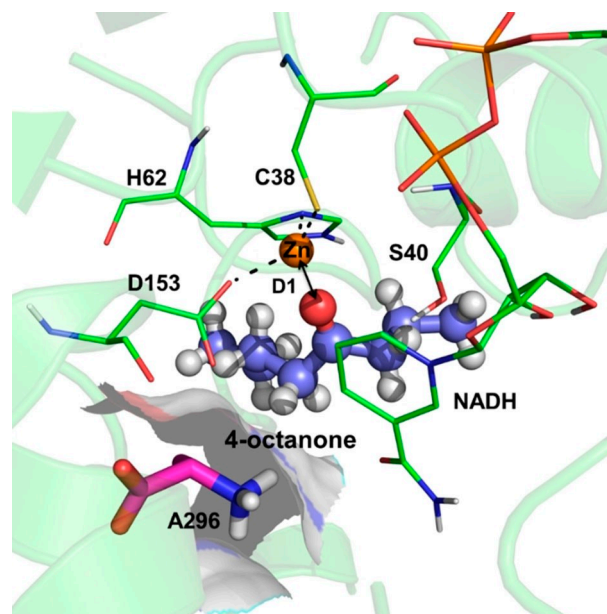


Figure 6. Molecular docking pose of octan-4-one in the binding pocket of W296A ReADH. The substrate is shown in ball and stick representation, the active-site residues are shown as sticks, the zinc ion is shown as orange sphere, and A296 is shown in magenta sticks. A reciprocal arrow shows the distance (D1) between the zinc ion and the carbonyl oxygen of the substrate.

the regioselectivity is observed for the W296A variant; instead of favoring small chemical groups like a methyl group in acetophenone and 2OCT, ReADH W296A favors larger chemical groups (i.e., propyl) in its small binding pocket. We demonstrate that the introduced substitutions in the smaller binding pocket

induce a systematic change in the catalytically competent substrate binding and the interactions of the carbonyl oxygen with the catalytic zinc ion, and thus determine the stereopreferences and regioselectivity of the ReADH variants. ReADH follows Prelog's rule and converts carbonyl compounds to their *S*-enantiomeric alcohols.^[12,14] However, the W296A variant could possibly show reversed enantioselectivity for the smaller substrates accepted by the WT, as the variant has a larger substrate binding pocket that allows the substrate to explore pro-*R* conformations with catalytically active hydride transfer distances. This phenomenon was observed in almost all previously engineered ketoreductases such as cpADH5^[25] and TbSADH.^[21,27]

Conclusions

Computer-assisted ReADH engineering revealed that enlarging the smaller binding pocket enables ReADH to convert bulkier substrates (ethyl 2-oxo-4-phenylbutyrate, octan-3-one, or octan-4-one) with larger alkyl chains adjacent to their carbonyl functionality. Molecular docking of substrates to WT ReADH showed that W296 is critical in controlling the substrate scope. SSM at position 296 yielded the improved ReADH variant W296A, which confirmed our computational prediction. Substitution of W296 with a smaller amino acid provided sufficient space for the catalytically competent binding of ethyl 2-oxo-4-phenylbutyrate and octan-4-one.

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

Experimental Section

Homology modeling: The amino acid sequence of ReADH (accession no: C0ZXL4, 348 aa) was retrieved from UniProt database^[28] in FASTA format. Homology modeling of the ReADH structure was performed by using YASARA Structure Version 13.9.8^[18] with the default settings (PSI-BLAST^[29] iterations: 6, *E* value cutoff: 0.5, templates: 5, OligoState: 4). Template structures were scored based on the position-specific scoring matrix (PSSM).^[30] Two X-ray templates were selected for hybrid modeling of ReADH structure (348 aa): Crystal structure of alcohol dehydrogenase from *R. ruber* (345 residues with quality score 0.487, PDB ID: 3JV7,^[16] resolution 2.0 Å) and crystal structure of the alcohol dehydrogenase from the hyperthermophilic archaeon *A. pernix* (343 residues with quality score 0.581, PDB ID: 1H2B,^[17] resolution 1.62 Å). The developed model is a hybrid homotetramer shown in Figure S1.

Mechanism-based substrate docking: The homodimeric structure of ReADH was used for substrate docking in AutoDock4.2^[31] within YASARA Structure Version 13.9.8.^[18a] In this mechanism-based docking protocol, reparameterization of the catalytic zinc environ-

ment including catalytic zinc ion and its coordinating residues (Cys38, His62, Asp153, and Water) was carried out. Additionally, reaction mechanism based distance filter criteria (distances *D*₁, *D*₂, and *D*₃ shown in Figure S2) was applied. *D*₁ is the distance between the carbonyl oxygen of the substrate and the catalytic zinc ion, *D*₂ is the distance between the hydrogen of the NADH cofactor and the carbonyl carbon of the substrate, and *D*₃ is the distance between the carbonyl oxygen and the OH group of Ser46. The details of the substrate-based docking protocol along with applied distance constraints can be found elsewhere.^[4b] For the substrate molecular-docking protocol, a bound model of the catalytic zinc ion with charge of +1.01 on the catalytic zinc ion^[12] and AM1-BCC charges^[32] for the substrate were used. PyMOL was used to prepare all the graphical images.

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

Construction of site-saturation libraries of ReADH: The expression vector pKA1, harboring the *S* alcohol dehydrogenase from *R. erythropolis* (ReADH)^[12] was used as the template for saturation mutagenesis of position W296. A modified two-step QuikChange mutagenesis (QCM) protocol was applied to generate mutant libraries. In the first step, two extension reactions were performed in separate tubes; one containing the forward primer (5'-ACAGTCCG-TATNNKGGTGCCCGC-3') and the other containing the reverse primer (5'-GCGGGCACCMNNATACGGAACTGT-3'). After 3 rounds of extension, the two reaction mixtures were mixed, and amplification was carried out for 15 cycles. The PCR products were digested with 20 U DpnI (New England Biolabs) in order to remove template DNA and purified with a PCR purification kit (NucleoSpin[®] Gel and PCR Clean-up kit, Macharey-Nagel, Düren, Germany). Libraries were transferred into *Escherichia coli* BL21 Gold (DE3) and plated on LB_{CM} agar plates.

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 μL LB medium supplemented with 34 μg mL⁻¹ 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 μL of 50 % (v/v) glycerol was added to each well, and the libraries were stored at -80 °C as a master microtiter plate.

Expression in microtiter plates was performed by duplicating the master plate with a 96-pin replicator to 150 μL of LB medium supplied with chloramphenicol (34 μg mL⁻¹) 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 μL of the pre-culture were transferred to V-bottom microtiter plates (transparent polystyrene plate, Corning GmbH, Kaiserslautern, Germany) containing 150 μL terrific broth (TB) supplemented with chloramphenicol (34 μg mL⁻¹), 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⁻¹ thiamine hydrochloride and 1 mM ZnCl₂. The 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) in an Eppendorf 5810R centrifuge (Eppendorf AG, Hamburg, Germany), and stored at -20 °C until further use.

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 pipetting in 150 μL lysis buffer (50 mM phosphate buffer pH 7.5, and 1 g L^{-1} lysozyme). Cells were lysed by incubating suspended cells in a microtiter plate shaker for 1 h (37 $^{\circ}\text{C}$, 900 rpm, 70% relative humidity), followed by centrifugation (4 $^{\circ}\text{C}$, 3220 g , 20 min). The SSM library was screened by monitoring NADH depletion, the result of ethyl 2-oxo-4-phenylbutyrate reduction to corresponding alcohol, spectrophotometrically (Tecan sunrise, Männedorf, Switzerland). The 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 adding 50 μL of 1 mM NADH, and absorption was monitored at 340 nm for 10 min ($\epsilon = 6200 \text{ M}^{-1} \text{ cm}^{-1}$).

Flask expression of ReADH protein: For flask expression of ReADH enzyme, 4 mL of overnight culture grown in LB medium with chloramphenicol (34 $\mu\text{g mL}^{-1}$; 37 $^{\circ}\text{C}$, 250 rpm) was inoculated into 250 mL TB medium supplemented with chloramphenicol (34 $\mu\text{g mL}^{-1}$), and trace element solution (0.25 mL containing; 3.4 mM CaCl_2 , 0.6 mM ZnSO_4 , 0.6 mM MnSO_4 , 54.0 mM $\text{Na}_2\text{-EDTA}$, 61.8 mM FeCl_3 , 0.6 mM CuSO_4 ; and 0.8 mM CoCl_2 in water), and the mixture was cultivated at 37 $^{\circ}\text{C}$, 250 rpm until the OD_{600} reached 0.6–0.8. Protein expression was induced by the addition of 0.1 mM IPTG and supplemented with thiamine hydrochloride (0.1 g L^{-1}) and ZnCl_2 (1 mM). Protein expression was carried out at 30 $^{\circ}\text{C}$, 250 rpm for 20–22 h. Cells were harvested by centrifugation (4 $^{\circ}\text{C}$, 3220 g , 15 min) and stored at -20°C until further use.

Activity determination of ReADH enzyme: ReADH 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 cells were disrupted by sonication for 5 min (30 s on + 30 s 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 $^{\circ}\text{C}$). Cleared cell lysate was further incubated at 65 $^{\circ}\text{C}$ for 15 min and centrifuged at 21300 g , 20 min, and 4 $^{\circ}\text{C}$. 80% of the *E. coli* proteins could be removed by heating the cell crude extract for 15 min at 65 $^{\circ}\text{C}$.^[12] The obtained supernatant was used for activity measurements. ReADH activity assays in cuvette-scale were performed on a Varian Cary 50 Bio UV/vis spectrophotometer (Agilent Technologies). 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 (pH 8.0). The reaction was started by adding 100 μL of partially purified ReADH. NADH absorption at 340 nm was monitored for 10 min ($\epsilon = 6200 \text{ M}^{-1} \text{ cm}^{-1}$). The protein concentration of the heat-treated lysate was determined by using the PierceTM BCA Protein Assay Kit according to the manufacturer's instructions (Thermo Fisher Scientific). Heat-purified ReADH was diluted accordingly in order to obtain linear absorbance change in the monitored timescale. The linear part of the measured activity curve was used to calculate enzymatic activity and stated in U per mg total protein concentration in heat-purified lysate. All measurements were performed in triplicate.

Acknowledgments

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.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: alcohol dehydrogenases · homology modeling · molecular docking · protein engineering · *Rhodococcus erythropolis*

- [1] a) F. Secundo, R. S. Phillips, *Enzyme Microb. Technol.* **1996**, *19*, 487–492; b) W. Hummel, M. R. Kula, *FEBS J.* **1989**, *184*, 1–13; c) C. L. Stone, T.-K. Li, 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, C. H. Senanayake, *Tetrahedron Lett.* **2001**, *42*, 8919–8921.
- [3] a) N. Itoh, N. Mizuguchi, M. Mabuchi, *J. Mol. Catal. B* **1999**, *6*, 41–50; b) J. C. Wang, M. Sakakibara, J. Q. Liu, T. Dairi, N. Itoh, *Appl. Microbiol. Biotechnol.* **1999**, *52*, 386–392.
- [4] a) H. Yamamoto, N. Kawada, A. Matsuyama, Y. Kobayashi, *Biosci. Biotechnol. Biochem.* **1999**, *63*, 1051–1055; b) G. V. Dhoke, Y. Ensari, M. D. Davari, A. J. Ruff, U. Schwaneberg, 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, F. Paradisi, *Org. Biomol. Chem.* **2017**, *15*, 9169–9175.
- [7] J. Cassidy, L. Bruen, E. Rosini, G. Molla, L. Pollegioni, F. Paradisi, *PLoS One* **2017**, *12*, e0187482.
- [8] a) Y. Xie, J.-H. Xu, Y. Xu, *Bioresour. Technol.* **2010**, *101*, 1054–1059; b) Y. Ni, C.-X. Li, H.-M. Ma, J. Zhang, J.-H. Xu, *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1111–1118; c) Y. Ni, C. X. Li, J. Zhang, N. D. Shen, U. T. Bornscheuer, J. H. Xu, *Adv. Synth. Catal.* **2011**, *353*, 1213–1217.
- [9] S. Sudhakara, A. Chadha, *Org. Biomol. Chem.* **2017**, *15*, 4165–4171.
- [10] a) S. Oda, Y. Inada, A. Kobayashi, H. Ohta, *Biosci. Biotechnol. Biochem.* **1998**, *62*, 1762–1767; b) W.-Q. Lin, Z. He, Y. Jing, X. Cui, H. Liu, A.-Q. Mi, *Tetrahedron: Asymmetry* **2001**, *12*, 1583–1587.
- [11] C. M. Nealon, M. M. Musa, J. M. Patel, R. S. Phillips, *ACS Catal.* **2015**, *5*, 2100–2114.
- [12] K. Abokitse, 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, 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. Vorbrod, G. Kunze, *Biochem. Eng. J.* **2016**, *106*, 107–117.
- [16] M. Karabec, A. Łyskowski, K. C. Tauber, G. Steinkellner, W. Kroutil, G. Grogan, K. Gruber, *Chem. Commun.* **2010**, *46*, 6314–6316.
- [17] L. Esposito, F. Sica, C. A. Raia, A. Giordano, M. Rossi, L. Mazzarella, A. Zagari, *J. Mol. Biol.* **2002**, *318*, 463–477.
- [18] a) E. Krieger, G. Vriend, C. Spronk, *YASARA. org* **993**; b) E. Krieger, G. Vriend, *Bioinformatics* **2014**, *30*, 2981–2982.
- [19] a) H. Man, C. Loderer, M. B. Ansorge-Schumacher, 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, R. Xiao, *Chem. Commun.* **2014**, *50*, 7770–7772.
- [20] G. V. Dhoke, C. Loderer, M. D. Davari, M. Ansorge-Schumacher, U. Schwaneberg, M. Bocola, *J. Comput.-Aided Mol. Des.* **2015**, *29*, 1057–1069.
- [21] M. A. María-Solano, A. Romero-Rivera, S. Osuna, *Org. Biomol. Chem.* **2017**, *15*, 4122–4129.
- [22] G. V. Dhoke, M. D. Davari, U. Schwaneberg, M. Bocola, *ACS Catal.* **2015**, *5*, 3207–3215.
- [23] a) A. Jakobinnert, R. Mladenov, A. Paul, F. Sibilla, U. Schwaneberg, M. B. Ansorge-Schumacher, P. D. de María, *Chem. Commun.* **2011**, *47*, 12230–12232; b) A. Jakobinnert, J. Wachtmeister, L. Schukur, A. V. Shivange, M. Bocola, M. B. Ansorge-Schumacher, U. Schwaneberg, *Protein Eng. Des. Sel.* **2013**, *26*, 291–298.
- [24] T. Schubert, W. Hummel, M.-R. Kula, M. Müller, *Eur. J. Org. Chem.* **2001**, *2001*, 4181–4187.
- [25] Y. Ensari, G. V. Dhoke, M. D. Davari, M. Bocola, A. J. Ruff, U. Schwaneberg, *Chem. Eur. J.* **2017**, *23*, 12636–12645.
- [26] A. E. Firth, W. M. Patrick, *Nucleic Acids Res.* **2008**, *36*, W281–W285.

- [27] B. Liu, G. Qu, J. Li, W. Fan, J. Ma, Y. Xu, Y. Nie, Z. Sun, *Adv. Synth. Catal.* **2019**, *361*, 3182–3190.
- [28] C. UniProt, *Nucleic Acids Res.* **2004**, *45*, D158–D169.
- [29] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
- [30] J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
- [31] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- [32] A. Jakalian, D. B. Jack, C. I. Bayly, *J. Comput. Chem.* **2002**, *23*, 1623–1641.
- [33] A. A. Canutescu, A. A. Shelenkov, R. L. Dunbrack, *Protein Sci.* **2003**, *12*, 2001–2014.

Manuscript received: April 22, 2020
Accepted manuscript online: May 16, 2020
Version of record online: June 30, 2020
