<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *ACS Catalysis*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Maurer, D., Enugala, T R., Hamnevik, E., Bauer, P., Lüking, M. et al. (2018)
Stereo- and Regioselectivity in Catalyzed Transformation of a 1,2-Disubstituted Vicinal
Diol and the Corresponding Diketone by Wild Type and Laboratory Evolved Alcohol
Dehydrogenases
ACS Catalysis, 8(8): 7526-7538
<https://doi.org/10.1021/acscatal.8b01762>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-355854>

Stereo- and Regioselectivity in Catalyzed Transformation of a 1,2-Disubstituted Vicinal Diol and the Corresponding Diketone by Wild Type and Laboratory Evolved Alcohol Dehydrogenases

Dirk Maurer,^{†,‡} Thilak Reddy Enugala,^{†,‡} Emil Hamnevik,[‡] Paul Bauer,[§] Malin Lüking,[§] Dušan Petrović,[§] Heidi Hillier,[‡] Shina C. L. Kamerlin,^{*,§} Doreen Dobritzsch,^{*,‡} and Mikael Widersten^{*,‡}

[‡] Department of Chemistry – BMC, Uppsala University, Box 576, SE-751 23 Uppsala, Sweden

[§] Department of Cell and Molecular Biology, Uppsala University, Box 596, SE-751 24 Uppsala, Sweden

* Corresponding authors:

S. C. L. Kamerlin (kamerlin@icm.uu.se), D. Dobritzsch (doreen.dobritzsch@kemi.uu.se) and M. Widersten (mikael.widersten@kemi.uu.se)

Abstract

ADH-A from *Rhodococcus ruber* DSM 44541 catalyzes the oxidation of (*S*)-1-phenylethanol 3,000-fold more efficiently as compared to the 2-hydroxylated derivative (*R*)-phenylethane-1,2-diol. The enzyme is also highly selective for *sec*-alcohols with comparably low activities with the corresponding primary alcohols. When challenged with a substrate containing two secondary alcohols, such as 1-phenylpropane-(1*R*,2*S*)-diol, ADH-A favors the oxidation of the benzylic carbon of this alcohol. The catalytic efficiency, however, is modest in comparison to the activity with (*S*)-1-phenylethanol. To investigate the structural requirements for improved oxidation of vicinal diols we conducted iterative saturation mutagenesis combined with activity screening. A first-generation variant, B1 (Y54G, L119Y) displays a two-fold higher k_{cat} value with 1-phenylpropane-(1*R*,2*S*)-diol and a shift in the cooperative behavior in alcohol binding, from negative in the wild type, to positive in B1, suggesting a shift from a less active enzyme form (T) in the wild type to a more active form (R) in the B1 variant. Also, the regiopreference changed to favor oxidation of C-2. A second-generation variant, B1F4 (F43T, Y54G, L119Y, F282W), shows further improvement in the turnover and regioselectivity in oxidation of 1-phenylpropane-(1*R*,2*S*)-diol. The crystal structures of the B1 and B1F4 variants describe the structural alterations to the active site the most significant of which is a re-positioning of a Tyr side-chain located distal to the coenzyme and the catalytic zinc ion. The links between the changes in

structures and stereoselectivities are rationalized by molecular dynamics simulations of substrate binding at the respective active sites.

Keywords Alcohol dehydrogenase; alcohol oxidation; directed evolution; enzyme engineering; biocatalysis; stereoselectivity; crystal structure; molecular dynamics simulations

1. INTRODUCTION

The significant contribution of biocatalytic transformations to the sustainable manufacturing of chemicals is undisputed. Hence, a fundamental understanding of structure-activity relationships of enzymes that catalyze relevant chemical transformations is necessary for the future development of effective biocatalysts. Of particular importance is understanding of the structural requirements of the active site in order to steer stereoselectivity in substrate acceptance or the resulting stereoconfiguration of reaction products.

A cornerstone in synthetic chemistry is the oxidation/reduction of alcohols and their respective carbonyl derivatives. These transformations are generally catalyzed by alcohol dehydrogenases in biological systems. A class of particularly attractive synthetic building blocks are α -hydroxy ketones (acyloins) which are precursors in the synthesis of *e.g.* bio-active natural products.^{1,2} The biocatalytic productions of acyloins has been approached using different enzyme systems including monooxygenases introducing alcohol functions into phenyl-substituted aliphatic ketones,³ or by alcohol dehydrogenase-catalyzed oxidation of 1,2-diols.^{4,5} Successful biocatalytic transformations of reduction of the corresponding ketones, producing asymmetric alcohols from acyloins, diketones or cyclic ketones, have also been reported making use of whole-cell systems,⁶⁻⁸ naturally occurring dehydrogenases^{9,10} or *in vitro* engineered variants.^{11,12}

Alcohol dehydrogenase A (ADH-A) from the bacterium *Rhodococcus ruber* (strain DSM 44541) is a potential biocatalyst candidate for redox transformations. ADH-A has been shown to retain its catalytic function in high concentrations of acetone and isopropanol,¹³⁻¹⁵ a feature that is beneficial from different perspectives. In particular, it allows for the regeneration of the redox state of the coenzyme NAD⁺ or NADH, depending on the reaction direction, simply by adding either acetone or isopropanol as a co-substrate in the reaction mixture. Hence, only catalytic amounts of the costly coenzyme are required with maintained good conversion into product. The possibility of adding higher concentrations of organic solvents also facilitates solvation of non-polar substrates that may be of synthetic interest.

ADH-A displays strong regiopreference for *sec*-alcohols and ketones. Primary alcohols are very poor substrates in comparison.¹⁶ For example, (1*S*)-1-phenylethanol ((*S*)-**1** in **Figure 1A**) is oxidized 5,000-fold more efficiently than 2-phenylethanol. Furthermore, vicinal diols are comparably poor substrates; (1*R*)-1-phenylethanediol ((*R*)-**2**) is transformed with a 3,000-fold lower efficiency as compared to (*S*)-**1**. As expected from the strong preference for *sec*-alcohols, given the choice of two alcohol functions, as in (*R*)-**2**, ADH-A exclusively oxidizes the secondary alcohol to 2-hydroxyacetophenone (**5**). This capacity to regioselectively oxidize 1,2-diols, producing the acyloin products, makes ADH-A a suitable model for studies that may clarify the underlying reasons for the observed regioselectivity in this enzyme. The dramatic differences in activity towards (*S*)-**1** and (*R*)-**2** further qualifies this enzyme for studies of substrate selectivity, again with special focus on stereoselectivity. Furthermore, the established stereoselectivity of ADH-A for *sec*-alcohols motivated us to study the regiopreference with a vicinal diol substrate containing two secondary alcohol functions, (1*R*,2*S*)-**3**. We conducted ISM-driven laboratory evolution of the enzyme with the aim to isolate more efficient catalysts of acyloins from (1*R*,2*S*)-**3**. Here, we demonstrate the successful isolation of ADH-A variants that exhibit increased turnover numbers for the oxidation of (1*R*,2*S*)-**3** and in the reduction of the corresponding diketone (**8**). We are also able to rationalize the stereopreference from molecular dynamics simulations using crystal structures of these variants as the structural basis.

2. MATERIALS AND METHODS

2.1. Chemicals and reagents. (*R*)-**1** (≥ 99 %, #77848), (*S*)-**1** (≥ 99 %, #77849), (*R*)-**2** (99 %, #302163), (*S*)-**2** (99 %, #302155), **4** (≥ 99 %, #00790), **5** (98 %, #445134) and **8** (99 %, #223034) were purchased from Merck (Darmstadt, Germany). (1*R*,2*S*)-**3** was produced by hydrolyzing 1-phenyl-(1*S*,2*S*)-propaneoxide (98 %, #287792, Merck) using epoxide hydrolase as described before¹⁷ (**Figure 1B**). All other reagents, solvents, substrates, coenzymes, microbial growth media, and chromatographic resins were purchased from commercial sources at highest available purity.

2.2. Gene library construction. The design of libraries for ISM-driven¹⁸⁻²⁰ evolution of ADH-A targeted amino acid residues constituting the active-site cavity as deduced from the crystal structure (PDB code: 3jv7²¹). The construction of library B has been described before.²² Library B1F was constructed similarly with some modifications and was created from the first-generation ‘B1’ variant as parent. The ‘F’ site contained two residues, F43 and F282*, where the asterisk indicates that this residue is located in the neighboring subunit of the dimer. Codon degeneracy was minimized using

oligonucleotides containing mixtures of NDT, VMA, ATG and TGG codons,²³ thus, the theoretical number of unique sequences in the constructed library was 400. Detailed descriptions of the library construction and sequences of used oligonucleotides are provided in the Supporting Information (see also **Figure S1 and Table S1**). The gene constructs were subcloned into the *Xho*I and *Spe*I sites of pGT7ADHA-5H.¹⁶ Plasmids were transformed into *Escherichia coli* BL21-AI (Invitrogen) already carrying a plasmid for over-expression of chaperonins GroEL/ES,²⁴ through electroporation and grown on LB plates containing 100 mg/ml ampicillin and 30 mg/ml kanamycin.

2.3. Library screening. Screening was performed essentially as described earlier.²² Clones were incubated in round-bottom Nunc 96-well plates over 2 days. Day one, 350 μ l 2TY (1.6 % (w/v) tryptone, 1.0 % (w/v) yeast extract, 0.5 % (w/v) NaCl) containing 100 μ g/ml ampicillin and 30 μ g/ml kanamycin was added to each well and single colonies were inoculated into designated wells. Each plate contained four parent enzyme clones as baseline controls and four negative controls with cells expressing an unrelated protein from the same expression construct (yeast YNR064c²⁵) and four blank controls containing only growth medium. Gas permeable seals were placed on the plates followed by incubation at 30 °C for 16 hours under rotation at 150 rpm. On the following day, overnight cultures were diluted 14-fold into new plates containing 325 μ l 2TY, 50 μ g/ml ampicillin and 30 μ g/ml kanamycin. Gas permeable seal was placed on the plates and the plates were incubated as described for 3 hours. After this incubation, β -D-thiogalactopyranoside (IPTG) and L-arabinose were added to each well to final concentrations of 1 mM and 0.04 % (w/v), respectively, and the plates were incubated as before for 16 hours. Following protein expression, cells were harvested by centrifugation at 3,000 $\times$ g, at 4 °C for 20 min. Supernatants were removed and plates were dried up-side down on paper tissues for 2 min before being covered with alumina foil. Sealed plates were stored at -80 °C until analyzed. Cells were lysed by addition of 100 μ l lysis solution (10 ml B-PER (Thermo Fisher Scientific, Waltham, MA, USA; 4 mg lysozyme, 2 mg DNase, 1 EDTA-free Complete mini-tablet, Merck; 10 μ M ZnSO₄), then incubated for 1 h on a table-top mixer (Vortex-Genie 2, Scientific Industries, Bohemia, NY, USA) at lowest speed. Lysed cells were diluted with 100 μ l assay buffer (0.1 M sodium-phosphate, pH 8.0, 10 μ M ZnSO₄) and centrifuged for 60 min at 3,000 $\times$ g for library in a swing-out centrifuge. The oxidation activities of expressed proteins with either (1R,2S)-**3** or isopropanol, as positive control, were measured in clear flat-bottom Nunc 96-well plates in a plate reader (SpectraMAX 190, Molecular devices, Sunnyvale, CA, USA) at 340 nm, 30 °C for 10-20 min. The conditions in each well for the measurement were 0.1 M sodium-phosphate, pH 8.0, 10 μ M ZnSO₄, 0.5 mM NAD⁺ and 10 mM

(1*R*,2*S*)-**3**. For each activity measurement, 40 μ l lysate was added in a total volume of 150 μ l. Clones were scored as hits dependent on the degree of increase in the reaction velocity as compared to the parent control. A re-screen was performed on isolated hits in triplicates. Overnight cultures of 2 ml 2TY and 100 μ g/ml ampicillin were made for all hits re-screened. Plasmid DNA was extracted and the ADH-A genes were sequenced in full. At least 1200 individual clones were screened from each library.

2.4. Enzyme expression and purification. Enzyme variants were expressed and purified as described before¹⁶ with the differences that expression of ADH-A variants and GroEL/ES were induced by addition of 0.04 % (w/v) L-arabinose and 1 mM IPTG, respectively. Cultures were incubated for 16-18 h at 30 °C. Cells were harvested by centrifugation at 5,000 $\times$ g and lysed either by French press or ultrasonication into Binding Buffer (20 mM sodium-phosphate, pH 7.5, 20 mM imidazole, 0.5 M NaCl). Lysate was cleared by centrifugation as described before. Purification was performed by immobilized metal ion chromatography using Ni(II) charged Chelating Sepharose (GE Healthcare, Uppsala, Sweden). Protein was adsorbed in batch at 4 °C for 1 h and the gel/lysate slurry was subsequently transferred to a column. The flow through was discarded and the gel was washed thrice by addition of 10 $\times$ gel volume of Washing Buffer (Binding Buffer containing 100 mM imidazole). Protein was eluted by addition of 2 $\times$ 5 ml Elution Buffer (Binding Buffer containing 300 mM imidazole). The eluted protein was desalted into Storage Buffer (0.1 M sodium-phosphate, pH 7.4, 10 μ M ZnSO₄). For protein crystallization the protein was desalted into 30 mM Tris-HCl, pH 7.4, 10 μ M ZnSO₄ and then concentrated. Protein purity was confirmed by SDS-PAGE stained with Coomassie Brilliant Blue-R250. The purity was in all cases deemed to be >98 %.

2.5. Steady state kinetics. Kinetic parameters were determined for variants of interest by measuring initial velocities in the presence of varying concentrations of (*S*)-**1**, (*R*)-(**2**), (*S*)-**2**, (1*R*,2*S*)-**3**, **4**, **5** and **8** (**Figure 1A**) in the presence of saturating levels of NAD⁺ or NADH as described previously.²² Curve fitting and model discrimination was performed using programs in the SIMFIT package (<http://www.simfit.org.uk/>). The steady state parameters k_{cat} and K_M were determined after fitting the Michaelis-Menten equation to the experimental data using MMFIT. In cases where cooperative behavior was observed in the substrate concentration dependence of the initial reaction velocities, a modified Hill equation²⁶ (**Eq. 1**) was fitted to the raw data, using program INRATE. The degree of cooperativity was defined by the Hill coefficient, n . The validity of the applied model was judged by F-tests applying a threshold of $P < 0.01$.

$$\frac{v_0}{[E]_{\text{tot}}} = \frac{k_{\text{cat}}[S]^n}{K_{0.5}^n + [S]^n} \quad (1)$$

2.6. Pre-steady state kinetics. Transient kinetics experiments were performed on an Applied Photo-physics (Leatherhead, UK) SX.20MV sequential stopped-flow spectrophotometer/fluorometer. All measurements were performed at 30 °C in 0.1 M sodium phosphate, pH 8.0, 10 μ M ZnSO₄. Averaged progression curves (≥ 5) were fitted to **Eq. 2**, a function of a single exponential with floating endpoint, for determination of the apparent rates, k_{obs} . F is the recorded fluorescence signal, A , the signal amplitude, t , time and C , the floating end point.

$$F = A \exp(-k_{\text{obs}}t) + C \quad (2)$$

2.6.1. Coenzyme binding. NADH binding to the enzymes was detected by monitoring the increase in NADH fluorescence occurring upon binding of the coenzyme to the active site. An excitation wavelength of 340 nm was used and fluorescent light was collected through a 420 nm cut-off filter. NAD⁺ binding was detected by monitoring the quenching of the intrinsic tryptophan fluorescence. Excitation was achieved by shining light of 290 nm, and the emitted light >320 nm was recorded. Nucleotide concentrations were always kept at least in 10-fold excess over the enzyme concentrations. The apparent rates, k_{obs} , were determined as described above. Kinetic parameters were determined by fitting the dependence of k_{obs} on the co-enzyme concentration to a linear (**Eq. 3**) function using program LINFIT.

$$k_{\text{obs}} = k_{\text{on}}[\text{NAD(H)}] + k_{\text{off}} \quad (3)$$

2.6.2. Alcohol oxidation. The change in fluorescence from the increase in the NADH concentration, as a result of alcohol oxidation was monitored. The oxidation reaction was measured with varying concentrations of alcohol and saturating concentration of NAD⁺ (0.4 mM). The enzyme concentrations were between 0.6 to 2.3 μ M. Averaged progression curves (≥ 5) were fitted **Eq. 2** to extract k_{obs} . The kinetic parameters, k_3 and K_2 were estimated by fitting the observed rates to **Eq. 4** by non-linear regression using MMFIT. K_2 is the equilibrium dissociation constant (k_{-2}/k_2) of the alcohol from the

ternary enzyme-substrate complex. Numbering of rate and equilibrium constants are based on the kinetic model in **Figure 2**.

$$k_{\text{obs}} = \frac{k_3[\text{R-OH}]}{K_2 + [\text{R-OH}]} \quad (4)$$

2.7. Synthesis and analysis of acyloins.

2.7.1. Oxidation of (1*R*,2*S*)-3. Wild type ADH-A or B1 or B1F4 (3 μM), were mixed with 280 mM (1*R*,2*S*)-3 and 16 mM NAD^+ . For nucleotide regeneration 280 mM sodium pyruvate and 2 U/ml lactate dehydrogenase were included. Reactions were carried out in 0.1 M sodium-phosphate, pH 8.0, with 10 μM ZnSO_4 , at 30 °C in the dark for 30 h.

2.7.2. Reduction of 8. Reactions were performed as in **2.7.1** but with 300 mM **8**, 0.4 mM NADH, 300 mM L-lactate (300 mM) and 2 U/ml lactate dehydrogenase.

2.7.3. Product analysis. The structures of reaction products were determined by NMR spectroscopy. Reaction mixtures were centrifuged at 5000 $\times g$ to pellet debris. End-point proton NMR (400 MHz) analysis was performed on the organic phase following extraction into deuterated chloroform (CDCl_3). NMR spectra are shown in **Figure S2**.

2.8. X-ray crystallography. The B1 and B1F4 variants were crystallized by hanging drop vapor diffusion at 20 °C. The drops were placed over a 1 ml reservoir well and consisted of 1 μl protein solution (8 mg/ml and 3 mg/ml of B1 and B1F4 respectively in 50 mM Tris-HCl, pH 8, 4 mM NAD^+) and 1 μl reservoir solution containing 18 % and 15 % (w/v) polyacrylic acid 5100 for B1 and B1F4 respectively, 100 mM Tris-HCl, pH 8.0 and 10 mM MgCl_2 . The crystals were cryo-protected by soaking in a solution of 19 % (w/v) polyacrylic acid 5100, 100 mM Tris-HCl, pH 8.0, 10 mM MgCl_2 and 10 % (v/v) glycerol for 10 seconds and flash frozen in liquid nitrogen. Crystallographic data were collected at 100 K at beamlines I04 or I24 of the Diamond Light Source (Didcot, UK). The data was autoprocessed and indexed using FAST DP²⁷ on site and further processed using COMBAT and AIMLESS from the CCP4 software suit.²⁸⁻³⁰ The data collection statistics are shown in **Table S2**.

The structures of both variants were solved by molecular replacement using PHASER³¹ with the ADH-A wild type structure (pdb ID: 3jv7²¹) as a search model. Manual model building as well as the addition of water molecules was performed with COOT³² and restrained and TLS refinement with REFMAC5.³³ For monitoring R_{free} , a set of randomly selected reflections (5%) was used. The

crystallographic models and data were deposited in the Protein Data Bank with accession codes 5od3 for B1 and 6fg0 for B1F4.

2.9. Molecular dynamics of substrate binding. Wild type ADH-A and its B1 and B1F4 variants were simulated using AMBER16³⁴ and the ff14SB force field³⁵ to describe the protein. The diols (*R*)- and (*S*)-**2** and (1*R*,2*S*)-**3** were parametrized using Antechamber³⁶ and the standard RESP protocol,³⁷ and the resulting parameters are provided in **Tables S3-S12** of the Supporting Information. Parameters for NAD⁺ and NADH were taken from the literature.^{38,39} As the system contains two divalent zinc ions per subunit, we used a cationic dummy model to describe these metal centers, based on our success with this model in previous studies.⁴⁰⁻⁴³ In the present work, dummy model parameters were obtained for a tetrahedral zinc model, based on the previous published data.^{44,45} Full details of the zinc parametrization are provided in the Supporting Information.

For wild type and B1, the B and C subunits of the tetramer in the crystal structures (3jv7 and 5od3, respectively) and the A and D chains with B1F4 (pdb code 6fg0) were used to simulate the protein in the dimeric form observed in solution. Where multiple rotamers of a given side chain were plausible, the rotamer with the higher occupancy was selected for the simulations, and where the different rotamers have equal probabilities, the first rotamer was selected. All crystallographic water molecules were then stripped from the system, before re-solvating the protein in an octahedral box of TIP3P water⁴⁶, with the closest solute atom at least 10 Å away from the box edges. Sodium and chloride ions were added to a 0.15 M concentration, and the substrates were placed manually in the active site to allow coordination of the hydroxyl moiety at the carbon involved in the hydride transfer towards the catalytic zinc. All structures were minimized in three steps with decreasing restraints on the protein system from 25 kcal/mol×Å² to zero, with 5000 minimization cycles at each step. Non-bonded interactions were explicitly calculated for all atoms within 8 Å of each other, with interactions beyond this cut-off being treated using the particle mesh Ewald (PME) approach. Following this, the systems were heated up over the course of 50 ps to the final temperature of 300 K in an NVT ensemble, again applying a modest restraint of 5 kcal/mol×Å² on the solute atoms. The step size from this point on was set to 2 fs, with SHAKE⁴⁷ applied to the hydrogen atoms in the system. Temperature control was achieved using the Langevin thermostat. Following this, the system was equilibrated under NPT conditions for additional 150 ps under the same conditions. During both steps, metal geometries were kept constrained to the crystal values using distance restraints of 20 kcal/mol/Å², with an equilibrium distance of 1.8 Å, as well as angle restraints for tetrahedral coordination with force constants of 15

kcal/mol/deg. This was followed by dropping the restraints down to zero over the course of 5 additional simulations, each 10 ps in length and reducing restraints by 1 kcal/mol each. The last equilibration step was performed for 10 ps without any restraints on the system, including the metal center. Additional distance restraints were applied to prevent E53 from coordinating the catalytic zinc during the during all equilibration steps except the final one. The restraints were chosen such as to add a 20 kcal/mol/Å penalty starting at a distance of 5.2 Å, with the well curvature set such as to increase the penalty to 50 kcal/mol/Å at 4.8 Å.

Following the equilibrations steps, the systems were further equilibrated for 40 ns in an NPT ensemble, with the glutamate restraints alternatively activated or disabled every 5 ns to achieve a stable coordination. After this, the systems were simulated for 60 ns without any restraints in an NPT ensemble at 300 K. All simulations were performed in triplicate from individually equilibrated structures.

Data analysis was performed on the snapshots captured from every 50 ps of the calculations, excluding the restrained equilibration. Protein and substrate RMSD values, as well as protein clusters, were calculated using CPPTRAJ⁴⁸ after backbone alignment. Clusters were centered on the residues surrounding the substrate bound in the active site, specifically residues at positions 383, 385, 407, 464, 498, 640, 694, 696 and 697, using the average linkage method and rms mass comparison with an epsilon of 4 and random sieve of 5. The production trajectories were further analyzed using R (version 3.4.4) to obtain cross-correlation information. For this, the concatenated coordinates were loaded into R and analyzed using the Bio3D package, with the two simulated subunits analyzed separately by first fitting to the C-alpha trace of the subunit, followed by PCA analysis. The script used for the analysis is available on request. Final analysis of atomic fluctuations was performed using CPPTRAJ as used before for RMSD analysis. For this, the coordinates were first fit to the average structure backbone atoms, before calculating the atomic fluctuation by residue. The final numbers are fluctuations in Å.

3. RESULTS AND DISCUSSION

3.1. Stereoselectivity of ADH-A in oxidation of vicinal diols. Wild type ADH-A catalyzes the oxidation of the benzylic carbon of vicinal diol **2** with a 14-fold preference for the *R*-enantiomer, as judged by the corresponding $k_{\text{cat}}/K_{\text{M}}$ values (**Table 1**). Given the relatively poor activity with (*R*)-**2**, as compared to the non-hydroxylated derivative (*S*)-**1**, it was not surprising to find that (1*R*,2*S*)-**3** is oxidized at a comparably low steady state rate. The velocity dependence on substrate concentration

with this compound exhibits clear negative cooperativity with a Hill coefficient of 0.75. This behavior is indicative of stabilization of a less active enzyme form ('T' in classical MWC nomenclature⁴⁹) with increasing concentration of (1*R*,2*S*)-**3**. The enzyme preferentially oxidizes the benzylic C-1 carbon producing acyloin (*S*)-**6** in a ratio of 3.6:1 over (*R*)-**7** as judged by NMR spectroscopy (**Table 2**). The catalytic activity with the 1,2-diketone **8** is high; k_{cat} is 62 s⁻¹ and the $k_{\text{cat}}/K_{0.5}$ value places this substrate as the most efficiently transformed by this enzyme so far (**Table 1**). The reduction of **8** is preferentially at C-2 forming acyloin **6** (**Table 2**).

We complemented the kinetics analyses with molecular dynamics (MD) simulations. The 80-fold lower k_{cat} with (*S*)-**2** as compared to (*R*)-**2**, can be rationalized from the simulations as the hydride to be transferred is pointing away from the nicotinamide C-4 (**Figure 3B**). Furthermore, the (1*R*,2*S*)-**3** diol binds tightly within the active site in an orientation favorable for hydride transfer from the benzylic C-1 carbon, with a distance of approximately 2.7 Å in the main cluster (**Figure 3C**). The corresponding distance for transfer from C-2 is similar but with increased RMSD values for the substrate (**Figure 3D**). The energetic difference between reaction at either C-1 or C-2, calculated from the product ratios of 3.6, is only 0.8 kcal/mol, so either binding mode may be trapped in the MD simulation. In the case of 1,2-diketone **8**, there was no convergence towards a stable structure cluster. Instead, none of the obtained averaged structures contains **8** bound at the active site in manner that is compatible with hydride transfer from C-4 of NADH to any of the carbonyl carbons of the substrate. This suggests that this flat and rigid molecule is not bound in a specific manner in the active site and that the relationships between rates of sampling productive binding modes will decide regioselectivity and product configuration(s). The experimentally observed regioselectivity in reduction of **8** was, however, caught in the simulation with one generated of the ADH-A mutants (*vide infra*). Our MD results suggest therefore that different combinations of substrate and enzyme lead to distinct binding modes and complex stabilities; subtle changes in the dynamic behavior of certain parts of the active site are apparently causing the differences in stereopreference.

3.2. Change in regio- and stereoselectivity by active-site substitutions. With the goal to isolate ADH-A variants with increased catalytic activity with (1*R*,2*S*)-**3** we conducted targeted saturation mutagenesis of four active-site residues. Targeted residues were chosen based on their expected contribution to shaping the alcohol/ketone binding compartment of the active site, thereby influencing substrate selectivity. To facilitate screening for active enzyme variants the number of mutated residues per constructed library were limited to two spatially adjacent residues but probing different parts of the

substrate binding site (**Figure 4**). A parallel aim with the laboratory evolution effort was to test if regioselectivity in oxidation may be linked to catalytic activity, and if so, which structural modifications of the active site were required?

The chosen residues were grouped in pairs; site ‘A’ consisting of Y294 and W295 and site ‘B’, of Y54 and L119 (**Figure 4A**). However, no enzyme variant from the A-site library displayed increased activity with (1*R*,2*S*)-**3** under the applied screening conditions. From the B-library, however, one mutant, dubbed ‘B1’ displayed improved catalytic activity in the screen, when compared to the parent enzyme (**Table S13**). The apparent improvement in enzyme activity is due to a two-fold increase in k_{cat} for (1*R*,2*S*)-**3** and a shift from negative to positive cooperativity with this substrate ($n=1.2$). Since the screening of the constructed enzyme library was performed in the presence of relatively high concentration of diol substrate (10 mM), the higher turnover number in combination with the positive cooperativity, resulted in an observed velocity increase in the screening assay, although the actual catalytic efficiency, if expressed as $k_{\text{cat}}/K_{0.5}$, is actually lower than that of the wild type (**Table 1**). The kinetic parameters, however, are not directly comparable since the Y54G, L119Y substitutions also cause the B1 variant to shift regiopreference from oxidation of the benzylic C-1 to C-2 (**Figure 5**). As a consequence, B1 catalyzed oxidation of (1*R*,2*S*)-**3** produces (*R*)-**7** as the major product (3.4:1) (**Table 2**). Thus, (1*R*,2*S*)-**3** is treated as two different substrates by the wild type and the B1 variant.

In the MD simulations, if (1*R*,2*S*)-**3** is deprotonated on the C-1 hydroxyl, it binds in an orientation in the B1 active site that leads to a stronger interaction between the substrate phenyl ring and the nicotinamide ring of the coenzyme, but increases the distance between the hydride and the C-4 of the co-factor to 5 Å (**Figure 6A**). The binding mode for the C-2-deprotonated substrate, albeit different from the corresponding binding in the wild type enzyme (compare **Figures 3D** and **6B**), positions the substrate in a manner that may allow for hydride transfer with a distance to the coenzyme of <3 Å (**Figure 6B**). The MD simulations are thus in agreement with the experimental observations that show a regiopreference for oxidation of C-2. Since none of the substituted residues can be envisioned to interact directly in either binding mode, changes in the dynamic nature of the active site may be a possible reason for the change in regiopreference with (1*R*,2*S*)-**3**.

With the mono-substituted vicinal diols **2** as substrates, the wild type’s 14-fold preference for the *R*-enantiomer over (*S*)-**2**, is shifted. The B1 variant instead displays a 2-fold preference for (*S*)-**2** as substrate, mainly caused by an increase in $k_{\text{cat}}^{(S)-2}$ with a parallel decrease in $k_{\text{cat}}^{(R)-2}$ (**Table 1**) while the K_M for either enantiomer increases to similar extents (6-8 – fold). These combined changes result in the

overall differences in k_{cat}/K_M . In spite of the decrease in turnover numbers with (*R*)-**2**, k_{cat} for this enantiomer still exceeds $k_{\text{cat}}^{(S)-2}$ by >four-fold. In MD simulations with these substrates the binding behavior is similar to what is observed with the wild type (**Figure S3**) and largely agrees with the observed kinetics. When including **8** and NADH in the simulation, the MD results largely follow that of the wild type as does the activity with this substrate. However, one cluster displays **8** bound inside the active site relatively close to the cofactor (7 Å) although outside the range expected to be compatible with efficient hydride transfer. The regioselectivity, as determined from the product structure, is similar to the wild type with reduction preferentially at C-2 producing **6** (**Table 2**).

Although the overall crystal structure of the B1 variant is essentially indistinguishable from the wild type structure (**Figure 7A**), with an RMSD of 0.26 Å for superimposition of the α -carbons, the residue changes at positions 54 and 119 affect the active-site cavity. The substitutions act truly synergistically (**Figure 7B**) in reshaping the active site distal to the reactive part of the coenzyme (**Figure 7C, D**). Although the Y54G, L119Y substitutions appear to have a relatively modest effect on overall active-site architecture, they result in local changes in intramolecular interactions: in the wild type, the phenyl ring of Y54 interacts via π - π bonds with the backbone amide of L119. This contact is lost in the B1 mutant which instead forms a new hydrogen bond with the main chain amide of A53. It should be mentioned that these interactions are not present in all snapshot structures of the MD simulations which suggests dynamic formation and disruption of these bonds. Although the topology of the active site is not altered substantially in the B1 crystal structure the functional consequences of shifts in regioselectivity with (1*R*,2*S*)-**3** and in enantioselectivity from (*R*)-**2** to (*S*)-**2** are clear.

The B1 gene was subsequently recruited to parent a second round of mutagenesis and selection. A library, ‘B1F’, was constructed, encoding randomized codons for residues F43 and F282, where F282 enters the active site from the neighboring subunit (**Figure 4B**). The rationale behind the design of this library was to introduce structural variability spatially opposite to the Y54G, L119Y substitutions. A number of variants that exhibited apparently higher activity with (1*R*,2*S*)-**3** under the screening conditions were identified (**Table S13**). Three of these F₂ variants, B1F1, B1F4 and B1F5 were expressed at larger scale, purified and characterized. B1F1 (F43H on the B1 background), shows comparable activity to B1 with (1*R*,2*S*)-**3**, and also the other tested substrates, whereas B1F5 (B1, F43S, V194I, F282P) displays *lower* activities also with (1*R*,2*S*)-**3** (**Table 1**). However, both these variants, and especially B1F5, are produced at higher-than-average levels which can explain why they

were scored as hits during the screen in which the measured reaction velocities were not normalized for protein expression levels.

Variant B1F4 (B1, F43T, F282W) displays improved activity with (1*R*,2*S*)-**3** also after its purification. The turnover number is increased 2.3-fold as compared to the parent B1 and 4.8-fold as compared to wild type ADH-A (**Table 1**). The activity with (*S*)-**2** is comparable to B1 while $k_{\text{cat}}^{(R)-2}$ is increased to a similar degree as with (1*R*,2*S*)-**3**. B1F4 also displays positive cooperativity in the oxidation of (1*R*,2*S*)-**3** ($n=1.4$). The regiopreference of B1F4 mimics that of the parent B1 but with a further accentuated preference for oxidation at C-2 (**Table 2** and **Figure 5**).

When analyzing the primary cluster obtained by clustering the structures obtained from the MD simulations (based on occupancies of greater than 90 %), (1*R*,2*S*)-**3** is bound in such a conformation as to facilitate hydride transfer primarily from C-1 to NAD⁺, although also the C-2 hydride is within 3.5 Å from the coenzyme (**Figure 6C, D**). Thus, the MD simulations, in this case, cannot capture the shift in regioselectivity observed experimentally. This is possibly due to the fact that the adjustments in substrate binding mode that are required to achieve the observed shift in regiopreference from C-1 to the closely positioned C-2, can be accommodated by the changes in the overall structure dynamics.

The catalyzed reduction of diketone **8** is especially effective with a k_{cat} of 160 s⁻¹, a 2.5-fold increase as compared to wild type ADH-A. The regiopreference in the catalyzed reductive reaction is, similar to both other enzymes, towards C-2 (**Table 2**). Again, the primary cluster found when analyzing the MD simulations as above shows that with B1F4 and **8** the substrate is bound in a conformation that would preferentially favor hydride transfer to C-2, in accordance with the experimental results (**Figure 8**).

The crystal structure of B1F4 is overall very similar to the wild type and B1 structures (**Figure 7A**) with an RMSD of the corresponding subunit's alpha carbons of compared chains of 0.18-0.26 Å for superimposition with B1, and 0.27-0.33 Å for superimposition with the wild type. The F43T and F282W substitutions creates a somewhat wider entrance into the active site (**Figure 7E**), but the differences are minor. T43 engages in a new water-linked hydrogen bond network involving the carbonyl oxygen of I271 and the 2' and 3' hydroxyls of the ribose of the nicotinamide moiety of the coenzyme (**Figure 9**). A similar interaction network has been observed in an F43S ADH-A variant isolated in another directed evolution effort²² and the same F43S substitution is also present in a number of F₂ variants isolated in this work (**Table S13**). The F43H substitution in B1F1-B1F3 has been observed before^{12,22} and the H43 always hydrogen-bonds to the same ribose hydroxyl groups. The

significance of these new protein-coenzyme interactions is not yet clear but could possibly compensate for other destabilizing mutations.

3.3. Rate limitation for k_{cat} and destabilization of nonproductive substrate binding. The two- and five-fold increases in turnover of (1*R*,2*S*)-**3** by B1 and B1F4, respectively, are, paradoxically, not due to changes in rate-limiting reaction steps. We have previously shown that nucleotide release is rate determining in the ADH-A catalyzed transformations of the preferred substrates (*S*)-**1** and **4**.¹⁶ The deuterium kinetic isotope effect for k_{cat} in the oxidation of isopropanol is significantly smaller ($^{\text{D}}k_{\text{cat}}=1.5$) than that of the oxidation step ($^{\text{D}}k_3=3.6$)²² which further suggests that (a) step(s) downstream of the chemical transformation is(are) rate determining.

Here, the oxidation rates of (1*R*,2*S*)-**3** are also considerably faster than the corresponding turnover numbers (**Table 3**) as are the release rates of NADH, which are 65 and 24-fold faster than $k_{\text{cat}}^{(1R,2S)\text{-}3}$ for B1 and B1F4, respectively. The comparable rates of alcohol oxidation and NADH release result in lower degrees of accumulation of the binary E•NADH complexes, and indeed, the amplitude of the kinetic ‘burst’ of NADH formation during the pre-steady state phase of the reaction is significantly lower as compared to *e.g.* (*S*)-**1** (**Figure S4**).

Since both k_3 and k_5 exceeds the turnover rates by more than one order of magnitude, the observed changes in the kinetic parameters must be caused also by other mechanisms. The parallel increases in k_{cat} and $K_{0.5}$ in the B1 and B1F4 catalyzed oxidation reactions suggest two things: (1) the frequency of binding modes that allow for formation of productive ternary enzyme-substrate complexes has increased at the expense of binding affinity for *unproductive* ground-state ternary complex(es), as reflected in higher $K_{0.5}$ values. This destabilization increases the probability of formation of productive binding modes, observed as elevations of k_{cat} . (2) productive binding modes require a shift in the positioning of the diol substrate so that the C-2 hydride can now react with NAD⁺. In contrast, in the wild type enzyme the substrate is bound in a less reactive manner in which the benzylic (C-1) hydride is preferentially transferred.

The nature of the *productive* ternary complex is illusive; is the increase in k_{cat} simply a reflection of a larger proportion of enzyme-substrate complexes with geometries that facilitates the climb to the transition state, similar to the proximity effect of intramolecular catalysis⁵⁰ or does it represent an increase in ternary ‘near attack complexes’ (NACs) where the alcohol and coenzyme are bound in a TS-like geometry, thereby lowering the energetic barrier for transformation into product?⁵¹⁻⁵⁹ Improved stabilization of a TS-like intermediate, such as a NAC, is expected to increase the rate of oxidation. The

rates of oxidation (k_3) are indeed, albeit moderately, faster in the B1 and B1F4 variants (**Table 3**). The fact that these enzymes favor oxidation at C-2, however, complicates the comparison since the measured rates describe different reactions.

The underlying mechanism responsible for the increase in turnover numbers is not known but the phenomenon can be explained by the theory of conformational sampling and selection in catalysis. Here, the enzyme structure undergoes dynamic changes between substrates that may exhibit distinct properties regarding catalytic efficiencies and substrate selectivities.^{60,61} Introduced mutations (or changes in environmental conditions⁶²) can shift the equilibria between substrates to unveil new functional properties.^{63,64} The increase in k_{cat} and the shifts in regioselectivities in the oxidation of (1*R*,2*S*)-**3** displayed here by the isolated ADH-A variants are due to the structural changes caused by the introduced substitutions but the distal location of these modifications in relation to the site of hydride transfer makes it unlikely that they are directly causing the observed changes in regioselectivities. An alternative suggestion could be that the equilibrium mixtures of the conformational substrates of these enzymes have been shifted to instead favor oxidation of C-2 rather than C-1, which is in itself a very subtle change. Thus, the increase in *productively* bound ternary enzyme-substrate complexes may be a reflection of a shift in the distribution of conformational substrates. The analysis of structural fluctuations during the MD simulation shows that the regions of the tested enzymes that display the largest differences in fluctuations of the C- α carbons between oxidation of C-1 or C-2 in (1*R*,2*S*)-**3** include the mutated residues 43, 54 and 119 (**Figure 10**). We note also that the issue of non-productive binding conformations in enzyme catalysis has been taken up by several other workers in different contexts, in addition to the references mentioned above (for some examples see *e.g.* refs. 65-68).

4. CONCLUSIONS

Nonproductive binding of substrate in the ground state has not been widely presented as an intrinsic feature of enzyme catalysis and substrate discrimination although it may be a general phenomenon.⁶⁹ Increased understanding of its role in enzyme selectivity will be important for efficient engineering of biocatalysts and for the fundamental understanding of enzyme function. In this study, relatively minor changes in the topology of the ADH-A active site, primarily caused by the altered position of a tyrosine side chain, affects the substrate selectivity: the regiopreference towards a vicinal diol containing two

sec-alcohol functionalities is shifted and the enantiodiscrimination between (*R*) and (*S*)-1-phenylethane-1,2-diol is affected.

Competing interests

The authors declare no competing interests.

Funding information

This work was supported by Stiftelsen Olle Engkvist Byggmästare (MW). The support from COST action 1303 Systems Biocatalysis is also gratefully acknowledged, as well as the Swedish Research Council (2015-04928) and the Knut and Alice Wallenberg Foundation (KAW 2013.0124, Wallenberg Academy Fellowship to SCLK). The authors also thank Diamond Light Source for beamtime (proposal nr. Mx11171 and mx15868), and the staff of beamlines I04 and I24 for assistance with crystal testing and data collection. Access was supported in part by the EU FP7 infrastructure grant BIOSTRUCTX (contract no. 283570). All computational work was performed on the Abisko and Kebnekaise clusters at HPC2N, through the generous allocation of computational time by the Swedish National Infrastructure for Computing (SNIC, 2015/16-12 and 2016/34-27).

Author contributions

[†] DM and TRE contributed equally to this work.

TRE, EH, HH and MW conducted enzyme kinetic analyses. DM and HH conducted the protein crystallography. DM and DD solved the protein crystal structures. PB, ML, DP and SCLK conducted and analyzed the computational simulations. All the authors discussed the results and contributed to writing of the manuscript.

Supporting Information

- (1) Description of strategy and oligonucleotide sequences used in the construction of library ‘B1F’.
- (2) Description of the parametrization of the catalytic zinc ion for the molecular dynamics simulations.
- (3) Supporting figures: **Figure S1**, cloning strategy; **Figure S2**, Extended NMR information and spectra of compound mixtures following ADH-A-catalyzed oxidation of (1*R*,2*S*)-**3** and reduction of **8**. **Figure S3**, Highest populated clusters after MD simulations with either (*R*)-**2** or (*S*)-**2** bound at the

active site of variant B1. **Figure S4**, Comparison of signal amplitudes after rapid mixing of enzyme and substrate during the pre-steady state phase of catalyzed reactions.

(4) All non-standard force field parameters used in the simulations.

Abbreviations ADH, alcohol dehydrogenase; ISM, iterative saturation mutagenesis; IPTG, β -D-thiogalactopyranoside; HPLC, high-performance liquid chromatography; NMR, nuclear magnetic resonance; RMSD, root mean square deviation; MD, molecular dynamics.

Acknowledgements

We are grateful to Andreas Söderlind for his contribution to the work.

References

- (1) Davis, F. A.; Chen, B.-C. Asymmetric Hydroxylation of Enolates with N-Sulfonyloxaziridines. *Chem. Rev.* **1992**, *92*, 919-934.
- (2) Wessjohann, L. Epothilones: Promising Natural Products with Taxol-Like Activity. *Angew. Chem. Int. Ed.* **1997**, *36*, 715-718.
- (3) Agudo, R.; Roiban, G.-D.; Lonsdale, R.; Ilie, A.; Reetz, M. T. Biocatalytic Route to Chiral Acyloins: P450-Catalyzed Regio- and Enantioselective α -Hydroxylation of Ketones. *J. Org. Chem.* **2015**, *80*, 950-956.
- (4) Bortolino, O.; Casanova, E.; Fantin, G.; Medici, A.; Poli, S.; Hanau, S. Kinetic Resolution of *vic*-Diols by *Bacillus stearothermophilus* diacetyl Reductase. *Tetrahedron Asym.* **1998**, *9*, 647-651.
- (5) Zhang, J.; Xu, T.; Li, Z. Enantioselective Biooxidation of Racemic *trans*-Cyclic Vicinal Diols: One-Pot Synthesis of Both Enantiopure (*S,S*)-Cyclic Vicinal Diols and (*R*)- α -Hydroxy Ketones. *Adv. Synth. Catal.* **2013**, *355*, 3147-3153.
- (6) Souza Brenelli, E. C.; Moran, P. J. S.; Rodrigues, J. A. R. Preparation of (-)-(1*R*,2*S*)-1-Phenylpropane-1,2-Diol by Fermenting Baker's Yeast Reduction of 1-Phenyl-1,2-propanedione. *Synth. Commun.* **1990**, *20*, 261-266.

- (7) Takeshita, M.; Sato, T. Synthesis of Optically-Active 1-Phenyl-1,2-Propanediol by Use of Baker's-Yeast. *Chem. Pharm. Bull.* **1989**, *37*, 1085-1086.
- (8) Martins, R. S.; Zampieri, D. S.; Rodrigues, J. A. R.; Carvalho, P. O.; Moran, P. J. S. Stereoselective Oxidation of 1-Phenyl-1,2-Propanediol Mediated by Microorganisms. *ChemCatChem* **2011**, *3*, 1469-1473.
- (9) Kulig, J.; Simon, R. C.; Rose, C. A.; Masood Husain, S.; Häckh, M.; Lüdeke, S.; Zeitler, K.; Kroutil, W.; Pohl, M.; Rother, D. Stereoselective Synthesis of Bulky 1,2-Diols With Alcohol Dehydrogenases. *Catal. Sci. Technol.* **2012**, *2*, 1580-1589.
- (10) Chen, X.; Mei, T.; Cui, Y.; Chen, Q.; Liu, X.; Feng, J.; Wu, Q.; Zhu, D. Highly Efficient Synthesis of Optically Pure (*S*)-1-Phenyl-1,2-Ethanediol by a Self-Sufficient Whole Cell Biocatalyst. *Chem. Open* **2015**, *4*, 483-488.
- (11) Sun, Z.; Lonsdale, R.; Ilie, A.; Li, G.; Zhou, J.; Reetz, M. T. Catalytic Asymmetric Reduction of Difficult-to-Reduce Ketones: Triple-Code Saturation Mutagenesis of an Alcohol Dehydrogenase. *ACS Catal.* **2016**, *6*, 1598-1605.
- (12) Hamnevik, E.; Maurer, D.; Enugala, T. R.; Chu, T.; Löfgren, R.; Dobritzsch, D.; Widersten, M. Directed Evolution of Alcohol Dehydrogenase for Improved Stereoselective Redox Transformations of 1-Phenylethane-1,2-Diol and Its Corresponding Acyloin. *Biochemistry* **2018**, *57*, 1059-1062.
- (13) Stampfer, W.; Kosjek, B.; Moitzi, C.; Kroutil, W.; Faber, K. Biocatalytic Asymmetric Hydrogen Transfer. *Angew. Chem. Int. Ed.* **2002**, *41*, 1014-1017.
- (14) Stampfer, W.; Kosjek, B.; Faber, K.; Kroutil, W. Biocatalytic Asymmetric Hydrogen Transfer Employing *Rhodococcus ruber* DSM 44541. *J. Org. Chem.* **2003**, *68*, 402-406.
- (15) Kosjek, B.; Stampfer, W.; Pogorevc, M.; Goessler, W.; Faber, K.; Kroutil, W. Purification and Characterization of a Chemotolerant Alcohol Dehydrogenase Applicable to Coupled Redox Reactions. *Biotechnol. Bioengineer.* **2004**, *86*, 55-62.
- (16) Hamnevik, E.; Blikstad, C.; Norrehed, S.; Widersten, M. Kinetic Characterization of *Rhodococcus ruber* DSM 44541 Alcohol Dehydrogenase A. *J. Mol. Catal. B. Enzym.* **2014**, *99*, 68-78.
- (17) Lindberg, D.; de la Fuente Revenga, M.; Widersten, M. Deep Eutectic Solvents (DESs) Are Viable CoSolvents for Enzyme-Catalyzed Epoxide Hydrolysis. *J. Biotechnol.* **2010**, *147*, 169-171.

- (18) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angew. Chemie Int. Ed.* **2005**, *44*, 4192-4196.
- (19) Reetz, M. T.; Carballeira, J. D. Iterative Saturation Mutagenesis (ISM) for Rapid Directed Evolution of Functional Enzymes. *Nat. Prot.* **2007**, *2*, 891-903.
- (20) Reetz, M. T.; Kahakeaw, D.; Sanchis, Shedding Light on the Efficacy of Laboratory Evolution Based on Iterative Saturation Mutagenesis. *J. Mol. BioSyst.* **2009**, *5*, 115-122.
- (21) Karabec, M.; Łyskowski, A.; Tauber, K. C.; Steinkellner, G.; Kroutil, W.; Grogan, G.; Gruber, K. Structural Insights into Substrate Specificity and Solvent Tolerance in Alcohol Dehydrogenase ADH-‘A’ from *Rhodococcus ruber* DSM 44541. *Chem. Commun.* **2010**, *46*, 6314-6316.
- (22) Hamnevik, E.; Enugala, T. R.; Maurer, D.; Ntuku, S.; Oliveira, A.; Dobritzsch, D.; Widersten, M. Relaxation of Nonproductive Binding and Increased Rate of Coenzyme Release in an Alcohol Dehydrogenase Increases Turnover with a Nonpreferred Alcohol Enantiomer. *FEBS J.* **2017**, *284*, 3895-3914.
- (23) Tang, L.; Gao, H.; Zhu, X.; Wang, X.; Zhou, M.; Jiang, R. Construction of “Small-Intelligent” Focused Mutagenesis Libraries Using Well-Designed Combinatorial Degenerate Primers. *Biotechniques* **2012**, *52*, 149-158.
- (24) Dale, G. E.; Schönfeld, H.-J.; Langen, H.; Stieger, M. Increased Solubility of Trimethoprim-Resistant Type SI DHFR from *Staphylococcus aureus* in *Escherichia coli* Cells Overproducing the Chaperonins GroEL and GroES. *Protein Eng.* **1994**, *7*, 925-931.
- (25) Elfström, L. T.; Widersten, M. The *Saccharomyces cerevisiae* ORF YNR064c Protein Has Characteristics of an ‘Orphaned’ Epoxide Hydrolase. *Biochim. Biophys. Acta* **2005**, *1748*, 213-221.
- (26) Hill, A. V. The Possible Effects of the Aggregation of the Molecules of Haemoglobin on its Dissociation Curves. *J. Physiol.* **1910**, *40*, iv–vii.
- (27) Winter, G.; McAuley, K. E. Automated Data Collection for Macromolecular Crystallography. *Methods* **2011**, *55*, 81-93.
- (28) Evans, P. Scaling and Assessment of Data Quality. *Acta Cryst.* **2006**, *D62*, 72-82.
- (29) Collaborative Computational Project, Number 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst.* **1994**, *D50*, 760-763.

- (30) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Cryst.* **2011**, *D67*, 235-242.
- (31) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Cryst.* **2007**, *40*, 658-674.
- (32) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Cryst.* **2010**, *D66*, 486-501.
- (33) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* **1997**, *D53*, 240-255.
- (34) Case, D. A.; Betz, R. M.; Botello-Smith, W.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. **2016**, AMBER 2016 Reference Manual (*Covers Amber16 and AmberTools16*), University of California, San Francisco.
- (35) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.
- (36) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Modell.* **2006**, *25*, 247-260.
- (37) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of the Multimolecule and Multiconformational RESP Methodology to Biopolymers: Charge Derivation for DNA, RNA, and Proteins. *J. Comp. Chem.* **1995**, *16*, 1357-1377.
- (38) Pavelites, J. J.; Bash, P. A.; Mackerell, Jr., A. D. A Molecular Mechanics Force Field for NAD⁺ NADH, and the Pyrophosphate Groups of Nucleotides. *J. Comput. Chem.* **1997**, *18*, 221-239.
- (39) Walker, R.C.; de Souza, M. M.; Mercer, I. P.; Gould, I. R.; Klug, D. R. Large and Fast Relaxations Inside a Protein: Calculation and Measurement of Reorganization Energies in Alcohol Dehydrogenase. *J. Phys. Chem. B* **2002**, *106*, 11658-11665.

- (40) Barrozo, A.; Duarte, F.; Bauer, P.; Carvalho, A. T. P.; Kamerlin, S. C. L. Cooperative Electrostatic Interactions Drive Functional Evolution in the Alkaline Phosphatase Superfamily. *J. Am. Chem. Soc.* **2015**, *137*, 9061-9076.
- (41) Purg, M.; Pabis, A.; Baier, F.; Tokuriki, N.; Jackson, C.; Kamerlin, S. C. L. Probing the Mechanisms for the Selectivity and Promiscuity of Methyl Parathion Hydrolase. *Phil. Trans. R. Soc. A* **2016**, *374*, 20160150.
- (42) Wallin, C.; Kulkarni, Y. S.; Abelein, A.; Jarvet, J.; Liao, Q.; Strodel, B.; Olsson, L.; Luo, J.; Abrahams, J. P.; Sholts, S. B.; Roos, P. M.; Kamerlin, S. C. L.; Gräslund, A.; Wärmländer, S. K. T. S. Characterization of Mn(II) Ion Binding to the Amyloid- Peptide in Alzheimer's Disease. *J. Trace Elem. Med. Biol.* **2016**, *38*, 183-193.
- (43) Blaha-Nelson, D.; Krüger, D. M.; Szeler, K.; Ben-David, M.; Kamerlin, S. C. L. Active Site Hydrophobicity and the Convergent Evolution of Paraoxonase Activity in Structurally Divergent Enzymes: The Case of Serum Paraoxonase 1. *J. Am. Chem. Soc.* **2017**, *139*, 1155-1167.
- (44) Duarte, F.; Bauer, P.; Barrozo, A.; Amrein, B. A.; Purg, M.; Åqvist, J.; Kamerlin, S. C. L. Force Field Independent Metal Parameters Using a Nonbonded Dummy Model. *J. Phys. Chem. B* **2014**, *118*, 4351-4362.
- (45) Liao, Q. H.; Kamerlin, S. C. L.; Strodel, B. Development and Application of a Nonbonded Cu²⁺ Model that Includes the Jahn–Teller Effect. *J. Phys. Chem. Lett.* **2015**, *6*, 2657-2662.
- (46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.
- (47) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341.
- (48) Roe, D. R.; Cheatham, III, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
- (49) Monod, J.; Wyman, J.; Changeux, J.-P. On the Nature of Allosteric Transitions: A Plausible Model. *J. Mol. Biol.* **1965**, *12*, 88-118.
- (50) Bruice, T. C.; Pandit, U. K. Intramolecular Models Depicting the Kinetic Importance of “Fit” in Enzymatic Catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **1960**, *46*, 402-404.
- (51) Lightstone, F. C.; Bruice, T.C. Ground State Conformations and Entropic and Enthalpic Factors in the Efficiency of Intramolecular and Enzymatic Reactions. 1. Cyclic Anhydride Formation by

- Substituted Glutarates, Succinate, and 3,6-Endoxo- Δ^4 -tetrahydrophthalate Monophenyl Esters. *J. Am. Chem. Soc.* **1996**, *118*, 2595-2605.
- (52) Torres, R. A.; Bruice, T. C. Molecular Dynamics Study Displays Near In-Line Attack Conformations in the Hammerhead Ribozyme Self-Cleavage reaction. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11077-11082.
- (53) Rajagapolan, P. T. R.; Benkovic, S. J. Preorganization and Protein Dynamics in Enzyme Catalysis. *Chem. Record* **2002**, *2*, 24-36.
- (54) Hur, S.; Bruice, T. C. The Near Attack Conformation Approach to the Study of the Chorismate to Prephenate Reaction. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12015-12020.
- (55) Giraldo, J.; Roche, D.; Rovira, X.; Serra, J. The Catalytic Power of Enzymes: Conformational Selection or Transition State Stabilization? *FEBS Lett.* **2006**, *580*, 2170-2177.
- (56) Griffin, J. L.; Bowler, M. W.; Baxter, N. J.; Leigh, K. N.; Dannatt, H. R. W.; Hounslow, A. M.; Blackburn, G. M.; Webster, C. E.; Cliff, M. J.; Waltho, J. P. Near Attack Conformers Dominate β -Phosphoglucomutase Complexes Where Geometry and Charge Distribution Reflect Those of Substrate. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 6910-6915.
- (57) Burchowsky, D.; van Erde, A.; Ökvist, M.; Kienhöfer, A.; Kast, P.; Hilvert, D.; Krenkel, U. Electrostatic Transition State Stabilization Rather Than Reactant Destabilization Provides the Chemical Basis for Efficient Chorismate Mutase Catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 17516-17521.
- (58) Sadiq, S. K.; Coveney, P. V. Computing the Role of Near Attack Conformations in an Enzyme-Catalyzed Nucleophilic Bimolecular Reaction. *J. Chem. Theory Comput.* **2015**, *11*, 316-324.
- (59) Warshel, A.; Bora, R. P. Defining and Quantifying the Role of Dynamics in Enzyme Catalysis. *J. Chem. Phys.* **2016**, *144*, 180901.
- (60) Ma, B.; Nussinov, R. Enzyme Dynamics Point to Stepwise Conformational Selection in catalysis. *Curr. Opin. Chem. Biol.* **2010**, *14*, 652-659.
- (61) Ramanathan, A.; Savol, A.; Burger, V.; Chennubhotla, C. S.; Agarwal, P. K. Protein Conformational Populations and Functionally Relevant Substates. *Acc. Chem. Res.* **2014**, *47*, 149-156.
- (62) Klinman, J. P.; Kohen, A. Evolutionary Aspects of Enzyme Dynamics. *J. Biol. Chem.* **2014**, *289*, 30205-30212.

- (63) James, L. C.; Tawfik, D. S. Conformational Diversity and Protein Evolution – a 60-Year-Old Hypothesis Revisited. *Trends Biochem. Sci.* **2003**, 28, 361-368.
- (64) Tokuriki, N.; Tawfik, D. S. Protein Dynamism and Evolvability. *Science* **2009**, 324, 203-207.
- (65) Glick, B. R.; Brubacher, L. J. Evidence for Nonproductive Binding Subsites Within the Active Site of Papain. *Can. J. Biochem.* **1974**, 52, 877-883.
- (66) Natarajan, S. K.; Sierks, M. R. Minimizing Nonproductive Substrate Binding: A New Look at Glucoamylase Subsite Affinities. *Biochemistry* **1997**, 36, 14946-14955.
- (67) Stojan, J. Non-Productive Binding of Butyryl(thio)choline in the Active Site of Vertebrate Acetylcholinesterase. *Chem. Biol. Interact.* **2010**, 187, 128-134.
- (68) Gao, D.; Chundawat, S. P. S.; Sethi, A.; Balan, V.; Gnanakran, S.; Dale, B. E. Increased Enzyme Binding to Substrate is Not Necessary for More Efficient Cellulose Hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, 110, 10922-10927.
- (69) Bar-Even, A.; Milo, R.; Noor, E.; Tawfik, D. S. The Moderately Efficient Enzyme: Futile Encounters and Enzyme Floppiness. *Biochemistry* **2015**, 54, 4969-4977.
- (70) The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

TABLES

Table 1. Steady state kinetic parameters

	k_{cat} (s ⁻¹)	$K_{0.5}$ (mM)	$k_{\text{cat}}/K_{0.5}$ (s ⁻¹ M ⁻¹)	n^a	k_{cat} (s ⁻¹)	$K_{0.5}$ (mM)	$k_{\text{cat}}/K_{0.5}$ (s ⁻¹ M ⁻¹)	n
substrate	wild type				B1 (Y54G, L119Y)			
(S)-1	80±20 ^b	0.63±0.05 ^b	130 000±30 000 ^b	(1)	11±1	32±6	340±30	(1)
(R)-2	0.73±0.01 ^b	17±0.6 ^b	44±8 ^b		0.17±0.008	130±10	1.3±0.06	(1)
(S)-2	0.0094±0.0005 ^b	3.0±0.7 ^b	3.1±0.6 ^b		0.039±0.001	17±2	2.3±0.2	(1)
(1R,2S)-3	0.48±0.01	11±1	43±5	0.75±0.03	0.98±0.03	33±2	30±2	1.2±0.06
4	36±0.8 ^b	1.2±0.09 ^b	30 000±2 000 ^b	(1)	7.3±0.6	4.8±1	1 500±300	(1)
5	2.0±0.08 ^b	3.7±0.4 ^b	550±30 ^b		0.54±0.02	15±0.9	36±1	(1)
8	62±0.6	0.31±0.01	200 000±9 000	1.3±0.06	85±0.5	0.32±0.009	260 000±8 000	0.93±0.02
	B1F1 (F43H, Y54G, L119Y)				B1F4 (F43T, Y54G, L119Y, F282W* ^c)			
(S)-1	22±0.8	47±3	460±40	1.1±0.02	37±3	73±10	500±100	0.77±0.02
(R)-2	0.96±0.09	940±100	1.0±0.05	(1)	0.45±0.006	140±4	3.2±0.004	(1)
(S)-2	0.022±0.002	28±5	1.3±0.3	1.3±0.2	0.072±0.008	60±20	1.2±0.4	0.79±0.06
(1R,2S)-3	0.76±0.01	41±1	19±0.6	1.3±0.03	2.3±0.04	35±1	66±3	1.4±0.05
4	5.7±0.03	5.9±0.1	980±20	1.5±0.09	26±0.6	8.5±0.5	3 000±200	1.2±0.05
5	0.62±0.03	39±3	16±2	1.1±0.03	0.98±0.2	25±1	38±1	(1)
8	84±2	1.8±0.1	48 000±3 000	1.2±0.07	160±2	0.51±0.02	320 000±2 000	(1)
	B1F5 (F43S, Y54G, L119Y, V194I, F282P*)							
(1R,2S)-3	0.41±0.009	48±2	8.4±0.4	1.4±0.05				
8	49±0.7	1.8±0.09	28 000±1 000	(1)				

^a n , Hill coefficient. $K_{0.5} = K_M$ if $n=1$. ^bData from ref. 16. ^c*, Residue from the neighboring subunit.

Table 2. Relative amounts of formed products following oxidation of (1*R*,2*S*)-**3** or reduction of **8**.

¹ H NMR integrals (%) ^a							
oxidation of (1 <i>R</i> ,2 <i>S</i>)- 3			reduction of 8				
	wild type	B1	B1F4		wild type	B1	B1F4
(1 <i>R</i> ,2 <i>S</i>)- 3	72	70	32	8	78	82	71
(<i>S</i>)- 6	22	7	12	6	21	17	27
(<i>R</i>)- 7	6	24	55	7	1	1	2
8	- ^b	-	-	3	-	-	-

^a Calculated from the relative abundance of the distinguishing methyl protons. ^b -, not detected. NMR spectra are shown in **Figure S2**.

Table 3. Kinetic rates and equilibrium dissociation constants

Substrate	k_1 ($\text{s}^{-1}\mu\text{M}^{-1}$)	k_{-1} (s^{-1})	K_1^a (μM)	k_3^b (s^{-1})	K_2^b (mM)	k_5 (s^{-1})	k_{-5} ($\text{s}^{-1}\mu\text{M}^{-1}$)	K_5^c (μM)
wild type								
NAD ⁺	6.8±0.5	110±10	17±2					
NADH						51±6 ^d	5.1±0.1 ^d	10±0.1 ^d
(S)-1				630±40 ^e	2.5±0.5 ^e			
(R)-2				42±8	87±30			
(1R,2S)-3				20±9	62±40			
B1 (Y54G, L119Y)								
NAD ⁺	9.7±0.4	190±20	20±2					
NADH						64±20	12±0.5	5.3±2
(1R,2S)-3				77±20	91±30			
B1F4 (F43T, Y54G, L119Y, F282W ^{*f})								
NAD ⁺	6.9±0.3	110±9	16±1					
NADH						56±30	9.6±1	5.9±3
(1R,2S)-3				68±20	41±20			

^a Calculated from the ratios of k_{-1}/k_1 . ^b From fitting Eq. 4 to the observed transient rates. ^c Calculated from the ratios of k_5/k_{-5} . ^d Data from ref. 22. ^e Data from ref. 16. ^f *, Residue from the neighboring subunit.

Figures

Figure 1. (A) Alcohols and ketones **1-5** and **8** were tested as substrates. Acyloins **6** and **7** are the oxidation products from **(1R,2S)-3** following hydride transfer from either carbon 1 or 2, respectively. (B) **(1R,2S)-3** is produced by epoxide hydrolase catalyzed hydrolysis of the corresponding **(1S,2S)-epoxide**.

Figure 2. Model of the kinetic mechanism for alcohol oxidation by ADH-A.

Figure 3. Structures of MD snapshots of most populated clusters obtained from the calculations with wild type ADH-A and (R)-2 (A), (S)-2 (B) and (1R,2S)-3 (C, D). (R)-2 binds in a configuration that can facilitate hydride transfer from C-1 of the diol to the nicotinamide C-4 (A) whereas with (S)-2, the corresponding hydrogen points away from the nicotinamide ring (B, hydrogen highlighted by a red dotted circle). (1R,2S)-3 is bound in an orientation for hydride transfer from the benzylic C-1 carbon with a distance of approximately 2.7 Å in the main cluster (C). The corresponding distance for transfer from C-2 is similar (D) but with slightly increased RMSD values for the substrate. See the Materials and Methods section for details on the methodology and the Supporting Information for details on parametrization of the active-site tetrahedral Zn²⁺.

Figure 4. Residues exposed to saturation mutagenesis. The surface shows the volume of the active site cavities. (A) In the wild type enzyme (pink), Y294 and W295 make up site ‘A’ and Y54 and L119, site ‘B’. (B) In the B1 variant (Y54G, L119Y), F43 and F282* (from the neighboring subunit) make up site ‘F’. Image created with PyMOL ver. 1.8.7⁷⁰ with the atomic coordinates in 3jv7²² (wild type) and 5od3 (B1).

Figure 5. Outcome of oxidation of (1R, 2S)-3 by wild type ADH-A or the B1 or B1F4 variants. The introduced substitution changes the regiopreference from C-1 to C-2.

Figure 6. Molecular dynamics simulation performed to describe the interactions of (1*R*,2*S*)-**3** (bright green) within the active sites of B1 (dark green) or B1F4 (blue). **(A)** In the B1 variant, if the diol is deprotonated on the C-1 hydroxyl, it becomes bound in an orientation that leads to a stronger interaction between the substrate phenyl ring and the nicotinamide ring, resulting in an increased distance between the hydride and the co-factor. **(B)** For the C-2-deprotonated substrate the distance between the hydride and the coenzyme in the most populated MD cluster is <3 Å, a distance in agreement with hydride transfer. The flexibility of F43 seen in the simulations with the wild type enzyme is observed also with B1. In the case of B1F4, the most stable structure cluster with (1*R*,2*S*)-**3** bound in the active site places C-1 in the most favorable position for oxidation **(C)**, while the distance for hydride transfer from C-2 is substantially larger and at an unfavorable angle **(D)**. Note that, as indicated in the Materials and Methods section, a multisite model was used to describe the Zn²⁺ ions in our simulations, shown here by a tetrahedron.

Figure 7. (A) Superimposition of the protein backbones of the A-chains of the wild type ADH-A (pink, 3jv7), B1 (dark green, PDB code 5od3) and B1F4 (blue, PDB code 6fg0) crystal structures. The catalytic and structural zinc ions are shown as spheres and their protein ligands and the bound coenzyme are shown as sticks. (B) Conformations of amino acid residues 54 and 119 in the wild type enzyme (pink) and the B1 variant (green). The Y54G substitution allows Y119 (in the B1 variant) to adopt a conformation in which the phenol side-chain occupies the freed space. (C-E), Surface representations of the active-site cavities of the wild type (C), B1 (D) and B1F4 (E) enzymes. The substitutions in the first-generation B1 variant reshapes the binding pocket distal to the reactive C-4 of the coenzyme NAD⁺. The F43T and F282W* substitutions result in a slightly wider opening to the active site. The coenzymes are shown in stick representation and the catalytic zinc ions as spheres.

Figure 8. Most stable structure cluster after molecular dynamics simulation of B1F4 with the diketone **8** (green). The substrate is bound in a position that could facilitate transfer of the *pro*-(*R*) hydride of NADH to C-2 of **8**. Our multisite Zn^{2+} model is shown here as a tetrahedron (see the Materials and Methods section for more details).

Figure 9. Active site of B1F4. The side-chain O- γ of the inserted T43 hydrogen is bound to the coenzyme ribose 2' and 3' hydroxyls via a water molecule. Zinc ligands are shown in sticks with white carbon atoms, and active-site residues subjected to mutagenesis are shown as sticks with blue carbon atoms. The image was created from the atomic coordinates of PDB entry 6fg0 with PyMOL ver. 1.8.⁷⁰.

Figure 10. Root mean square fluctuations of C- α during molecular dynamics simulations of complexes of enzymes with bound (1*R*, 2*S*)-3. **(A)** Diol (1*R*,2*S*)-3 bound in the active site for oxidation of C-1. Regions of relatively higher fluctuations, excluding the N- and C-termini, together with the Rossman domain are indicated by colored bars. Their locations in the folded structure are shown in **(D)**. **(B)** Diol (1*R*,2*S*)-3 bound for oxidation of C-2. **(C)** Differences in RMSF between the two enzyme-substrate complexes. The region displaying the largest differences between different substrate binding modes is indicated by a blue bar and its location in the folded structure is shown in **(E)**. See the Materials and Methods section for a detailed description of how these values were calculated.