Modeling-Assisted Design of Thermostable Benzaldehyde Lyases from *Rhodococcus erythropolis* for Continuous Production of α-Hydroxy Ketones

Martin Peng,[a] Dominik L. Siebert,[a] Martin K. M. Engqvist,[b] Christof M. Niemeyer,[a] and Kersten S. Rabe*[a]

Enantiopure α-hydroxy ketones are important building blocks of active pharmaceutical ingredients (APIs), which can be produced by thiamine-diphosphate-dependent lyases, such as benzaldehyde lyase. Here we report the discovery of a novel thermostable benzaldehyde lyase from *Rhodococcus erythropolis* R138 (*Re*BAL). While the overall sequence identity to the only experimentally confirmed benzaldehyde lyase from *Pseudomonas fluorescens* Biovar I (*Pf*BAL) was only 65%, comparison of a structural model of *Re*BAL with the crystal structure of *Pf*BAL revealed only four divergent amino acids in the substrate binding cavity. Based on rational design, we generated two *Re*BAL variants, which were characterized along with the wild-type enzyme in terms of their substrate spectrum, thermostability and biocatalytic performance in the presence of different co-solvents. We found that the new enzyme variants have a significantly higher thermostability (up to 22 °C increase in T_{50}) and a different co-solvent-dependent activity. Using the most stable variant immobilized in packed-bed reactors via the SpyCatcher/SpyTag system, (*R*)-benzoin was synthesized from benzaldehyde over a period of seven days with a stable space-time-yield of 9.3 mmol·L^{-1}·d^{-1}. Our work expands the important class of benzaldehyde lyases and therefore contributes to the development of continuous biocatalytic processes for the production of α-hydroxy ketones and APIs.

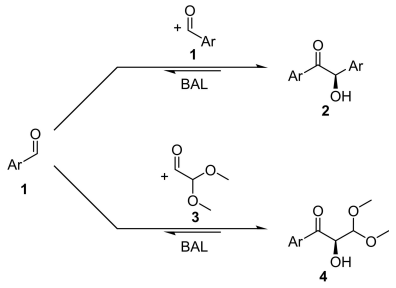
Introduction

The production of active pharmaceutical ingredients (APIs) and natural compounds with multiple stereocenters is a highly topical field in research and development. Chiral α-hydroxy ketone moieties are key structural units of natural products and APIs of industrial relevance as they can be readily converted, for example, into diols, amino alcohols or esters.[1] The enantioselective synthesis of α-hydroxy ketones can be realized enzymatically by using oxidoreductases, like alcohol dehydrogenases[2] or benzil reductase,[3] or else, via thiamine-diphosphate (ThDP)-dependent lyases, such as pyruvate decarboxylase, benzoylformate decarboxylase or benzaldehyde lyase.[1]

Benzaldehyde lyases (BAL, EC 4.1.2.38) catalyze the stereoselective homo-coupling and cross-coupling of two aldehydes via benzoin condensation (Scheme 1) by employment of ThDP and Mg^{2+} as cofactors. The BAL from *Pseudomonas fluorescens*[4,5] (*Pf*BAL) is the only experimentally confirmed representative in this enzyme class recorded in the BRENDA[6] and TEED[7] database to date (for more detailed analysis of TEED database see supporting information page 20). It has an broad substrate spectrum[8,9] and can also catalyze C–N bond formations,[10] cyclizations[11] and intramolecular Stetter reactions.[12] Moreover, via rational engineering of its binding site, *Pf*BAL was successfully altered to perform benzoylformate decarboxylase[13] or formolase reactions.[14] However, current applications are limited by the relatively low stability of this enzyme. For example, *Pf*BAL loses 40% of its initial activity within 27 h in phosphate reaction buffer at pH 8 and 25 °C.[15]

[a] M. Peng, D. L. Siebert, Prof. Dr. C. M. Niemeyer, Dr. K. S. Rabe
Karlsruhe Institute of Technology (KIT)
Institute for Biological Interfaces (IBG 1)
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen (Germany)
E-mail: kersten.rabe@kit.edu
[b] Dr. M. K. M. Engqvist
Chalmers University of Technology
Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Kemivägen 10, 412 96 Gothenburg (Sweden)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202100468



Scheme 1. Benzaldehyde lyase (BAL)-catalyzed formation of (*R*)-benzoin derivate (**2**) as homo-coupling product of aromatic aldehydes (**1**) and formation of (*R*)-2-hydroxy-3,3-dimethoxy-1-arylpropan-1-one (**4**) as cross-coupling product of aromatic aldehydes (**1**) and 2,2-dimethoxyacetaldehyde (**3**) investigated in this work. Ar: aromatic residue. For distinctive structures (**1a**–**p**), see Scheme 2.

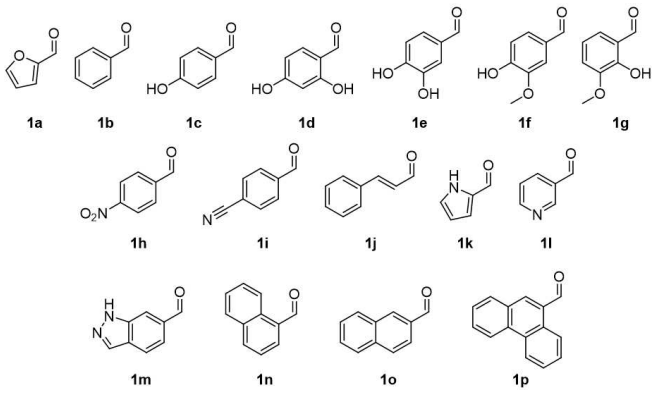
This is especially critical for the realization of continuously operated flow reactors. Hence, continuous flow reactors using this enzyme based on retention by membranes,[16,17] entrapment in polyvinyl alcohol[18] or adsorption to silica particles[19] showed limited success and were technically laborious. Likewise, several strategies for the direct, site-selective immobilization of *Pf*BAL on solid supports by employment of genetically-encoded tag systems, like the hexahistine (His_6)-tag,[19,20] HaloTag[21] or aggregation-inducing tags,[22] did not show the desired performance and only the His_6-tag variant could be used in a fluidic setup.[19]

Results and Discussion

To overcome the above limitations and to identify BAL enzymes of higher stability that promise improved production of α-hydroxy ketones in a continuous reaction, we used the computational prediction tool Tome, which was recently developed by combining datasets of amino acid composition of individual enzymes with the optimal growth temperature of organism in a machine learning model.[23] From the resulting set of sequences, we selected three candidate enzymes. A potential BAL from *Rhodococcus erythropolis* (*Re*BAL), which was selected based on the highest sequence identity (see Table S1). Although *R. erythropolis* is a mesophilic organism, individual enzymes may exhibit much higher thermostability than the growth temperature of the organism, as previously reported.[24] Furthermore, one potential BAL each was selected from *Sphaerobacter thermophilus* (*St*BAL) and *Thermobacillus composti* (*Tc*BAL). The enzymes were obtained by heterologous expression in *Escherichia coli* (Figure S1A, sequence data see Table S2) and the raw cell lysates were initially screened for aromatic ligase activity, using a literature-reported photometric assay[25] based on the conversion of 2-furaldehyde (**1a**, Scheme 2) into 2,2′-furoin (**2a**) (see also Figure S1B).

The enzyme from *R. erythropolis* showed the desired activity with **1a** (Figure S1B), whereas the enzymes from *S. thermophilus* and *T. composti* did not show any significant activity with this substrate. This finding was reasonable since the selection of these enzymes was based on sequence homology alone. In order to better understand the results, we performed a multiple sequence alignment of the three enzymes along with other *Pf*BAL-related enzymes described in the literature. The resulting phylogenetic tree (Figure S2) revealed that the inactive *St*BAL and *Tc*BAL are indeed closer related to acetohydroxyacid synthase (AHAS, E.C. 4.1.3.18) from *Saccharomyces cerevisiae* then to *Pf*BAL. This result suggested that the potential catalytic activity of *St*BAL and *Tc*BAL may include the decarboxylation of pyruvate like AHAS,[26] rather than the desired benzoin condensation. In contrast and in line with the activity observed for *Re*BAL, this enzyme is more closely related to *Pf*BAL on the sequence level. This finding also reflects the capability of both organisms of origin to metabolize aromatic compounds. As such, the organism of origin of *Pf*BAL (*P. fluoreszens* Biovar I) was specifically enriched in medium with anisoin as the sole carbon and energy source,[4] and the gene encoding for *Re*BAL is originally located on the plasmid pLRE138 from *R. erythropolis* R138 that harbors several genes that were classified as involved in uptake and metabolism of aromatic compounds.[27] The two enzymes are of comparable length (*Re*BAL, 557 amino acids and *Pf*BAL, 562 amino acids) and share 283 identical and 83 similar amino acids with only 9 gaps, indicating that 65% of the sequence is conserved between the two variants (Figure S3). Considering this high sequence similarity, we generated a 3D model of the protein structure of *Re*BAL (Figure 1) based on the crystal structure of *Pf*BAL[28] using SWISS Model.[29–33] The model suggests the same homotetrameric structure as found for the *Pf*BAL,[34] which can be described as a dimer of a dimers with four active sites (Figure 1, orange circles) located at the interface.

Comparison of the model of *Re*BAL (Figure 2A) with the structure of *Pf*BAL (Figure 2B) with respect to the residues considered to mainly constitute the binding site[35] indicated that these residues were highly conserved. The only exceptions are Q112, E393, S416 and T479 in *Re*BAL, which are located in



Scheme 2. Distinctive aromatic aldehydes (**1a**–**p**) investigated in this work.

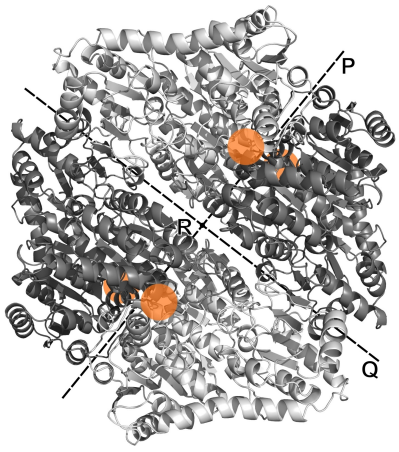


Figure 1. Proposed molecular structure of the benzaldehyde lyase *Re*BAL from *Rhodococcus erythropolis*. The model is based on the crystal structure of the tetrameric benzaldehyde lyase *Pf*BAL from *Pseudomonas fluorescens* (PDB: 3D7K),[28] with three twofold axes P, Q and R. The locations of the four active sites are indicated as orange circles (two in the front plane and two in the back).

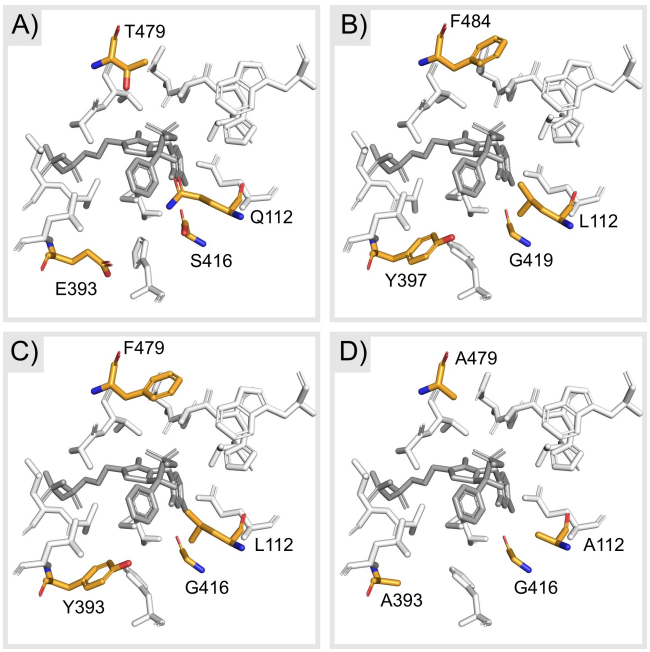


Figure 2. Structure model of A) wild-type *Re*BAL and B) crystal structure of *Pf*BAL (PDB: 3D7K)[28] as well as models of C) *Re*BAL$_{mat}$ and D) *Re*BAL$_{wid}$. Models were generated using SWISS Model.[29–33] Stick models of all amino acid residues of *Pf*BAL, wild-type, matched and widened *Re*BAL within a range of 6 Å of the catalytic site containing thiamine diphosphate (ThDP) and benzoylphosphonic acid methyl ester (MBP) (gray). Amino acid residues, which differ between *Pf*BAL and *Re*BAL and were mutated in *Re*BAL$_{mat}$ and *Re*BAL$_{wid}$, are represented as orange sticks. Coordinates of the ThDP-MBP adduct in the modeled structures were superimposed from the crystal structure of *Pf*BAL.[28] For a surface representation of the binding sites, see Figure S4.

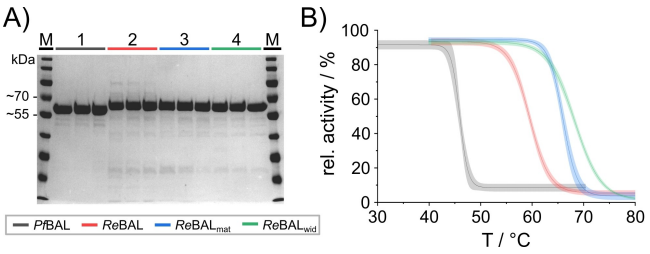


Figure 3. A) His$_6$-tagged *Pf*BAL (60 kDa, 1), *Re*BAL (61 kDa, 2), *Re*BAL$_{mat}$ (61 kDa, 3) and *Re*BAL$_{wid}$ (61 kDa, lanes 4) were produced as pure enzymes in biological triplicates and analyzed using a Coomassie stained 15% SDS tris-glycine polyacrylamide gel. M: PageRuler Prestained Protein Ladder (Thermo Scientific). B) Average thermal inactivation curves with 95% confidence interval of all four purified BAL enzymes at an identical concentration of 5 μM after 10 min incubation in 50 mM KP$_i$ reaction buffer pH 8 containing 2.5 mM $MgSO_4$ and 0.25 mM ThDP (for individual data sets see Figures S6–S9). Analyses were carried out in technical duplicates and biological triplicates.

an equivalent position as L112, Y397, G419 and F484 in *Pf*BAL as suggested from alignment data (Figure S3). These residues lead to the formation of a larger and, at the same time, more polar binding site in *Re*BAL, suggesting that the two enzymes have different substrate scopes. However, no information on the influence of the corresponding residues of *Pf*BAL L112, Y397, G419 on the catalytic properties of the enzyme are available, except of position 484, where mutations have led to an about 70% reduction of k_{cat} for benzoin lysis.[15]

We therefore decided to not only compare the activity and selectivity of the wild-type *Re*BAL with the well characterized *Pf*BAL, but to also generate *Re*BAL variants with either a binding site matched to *Pf*BAL (matched, *Re*BAL$_{mat}$, Figure 2C) or an even larger but less polar binding site (widened, *Re*BAL$_{wid}$, Figure 2D). To this end, we used site-directed mutagenesis to generate the *Re*BAL$_{mat}$ variant bearing mutations Q112L, E393Y, S416G and T479F as well as the *Re*BAL$_{wid}$ variant with mutations Q112A, E393A, S416G and T479A (for the corresponding primer sequences see Table S3). The formation of larger binding sites was also validated using CASTp[36] (Figure S5).

After heterologous expression, we purified *Re*BAL, *Re*BAL$_{mat}$ and *Re*BAL$_{wid}$ as well as *Pf*BAL in biological triplicates (Figure 3A) and compared their thermostability. To this end, the temperature-dependent loss of activity was characterized by incubating 5 μM enzyme in solutions containing 0.25 mM ThDP and 2.5 mM $MgSO_4$ as cofactors for 10 min at defined temperatures in the range of 30–80 °C. The above described 2-furaldehyde-based activity assay was then used to determine the remaining enzyme activity at 30 °C. T_{50} values, defined as the temperature at which 50% of the initial enzyme activity still remained, were determined from the plots of the residual activities against the incubation temperatures (Figures S6–S9). The analysis (Figure 3B) showed a significantly increased T_{50} value for *Re*BAL (59.7 °C ± 0.6) as compared to *Pf*BAL (45.9 °C ± 0.4). Interestingly the *Re*BAL$_{mat}$ (66.0 °C ± 0.2) and the *Re*BAL$_{wid}$ (68.4 °C ± 0.3) variants showed an even higher T_{50} value. These findings could be explained by the location of the binding site at the dimer interface, where an exchange of amino acids might also affect the interaction between both monomers resulting in an increased overall stability as it has been previously reported for other enzyme classes.[37]

In order to compare the activity of the here investigated enzymes with published data, we directly determined the turnover number (TON) for the conversion of benzaldehyde (**1 b**) into benzoin (**2 b**, Scheme 1) for all four BAL variants using an HPLC assay (Figure S10). For the *Pf*BAL we obtained a TON of 234 s^{-1} ± 11, which is comparable to previously reported values.[21,38] For *Re*BAL we found an about 15 times lower conversion of **1 b** of only 15 s^{-1} ± 1. As expected, matching the amino acids in the binding site of *Re*BAL to the residues found in *Pf*BAL led to a twofold increase of the TON to 36 s^{-1} ± 1 for the *Re*BAL$_{mat}$ variant, while the larger binding site of *Re*BAL$_{wid}$ showed a slightly higher TON of 45 s^{-1} ± 1. This finding can be correlated with the replacement of polar residues in the binding site with nonpolar residues, which should lead to an increased affinity for the nonpolar (aromatic) substrate. This hypothesis is in line with the activity of *Re*BAL$_{wid}$ having the smallest number of polar residues and showing the highest turnover of the *Re*BAL variants. However, the significantly lower activity of the *Re*BAL variants is compensated by their thermostability at elevated temperature. For example, the residual activity of *Re*BAL significantly exceeded that of *Pf*BAL after incubation at 50 °C for 1 h (Figure S11).

In order to further explore the differences in the molecular structure of the binding site on the enzymatic properties, we characterized the four BAL variants regarding their individual substrate spectrum using 16 aromatic aldehydes (**1a**–**p**, Scheme 2) for aromatic homo-coupling and aromatic-aliphatic cross-coupling reactions with 2,2-dimethoxyacetaldehyde (**3**, Scheme 1). Since many of these substrates are only poorly soluble in aqueous buffers, we used 20% DMSO as co-solvent, which had been previously reported for the *Pf*BAL.[15] The conversion of substrate and formation of product was determined by thin-layer chromatography (TLC, Figure S12) after 20 h incubation at 30 °C (Table 1). Compared to *Pf*BAL, *Re*BAL showed a similar but less broad substrate spectrum for the homo-coupling reactions with considerably lower formation of **2h**, **2i**, **2l** and **2p**. This indicates that *Re*BAL does not accept substrates with polar, aprotic *para*-substitutions as well as pyridine compounds. As expected, matching the binding site of *Re*BAL (*Re*BAL$_{mat}$) to the residues found in *Pf*BAL restored an almost identical substrate pattern while losing reactivity against **2d**. In contrast, the simple widening of the binding site by the introduction of alanine and glycine residues did not significantly alter the substrate spectrum. When compared to the wild-type *Re*BAL, the *Re*BAL$_{wid}$ no longer accepts **2d** but gains some reactivity towards **2p**.

However, the impact of the mutations was found to be different for the cross-coupling reactions with substrate **3**. While *Re*BAL performed quite poorly in these reactions, only accepting **1h** with >10% conversion, the introduction of the mutations into the *Re*BAL binding site improved the ability to catalyze the cross-coupling. Again, the substrate scope of the *Re*BAL$_{mat}$ showed a high similarity to *Pf*BAL albeit with an overall lower conversion. Like in the case of the homo-coupling reactions, here the *Re*BAL$_{wid}$ only showed a slightly altered reactivity compared to the wild-type *Re*BAL with better acceptance of substrate **1a**, **1b**, **1j** and **1o**, although with loosing acceptance of **1h** and **1l**. These results clearly illustrate that the substrate spectrum of the wild-type *Re*BAL can be aligned with the *Pf*BAL enzyme by replacing specific amino acids within the binding site. Of note, similar as *Pf*BAL, these *Re*BAL variants also showed a clear (*R*)-selectivity for the formation of the homo- and cross-coupling products of **1b**. The stereoselectivity of the products was confirmed by determining the enantiomeric excess of **2b** (Figure S13) as well as preparation and analysis of Mosher ester derivatives[39] of the reaction product **4b** with Mosher acid chloride (Figures S14–S16) using the known stereoselectivity of *Pf*BAL with **4b** as reference.[8]

In order to expand the range of hydrophobic substrates of the BALs variants, we also wanted to investigate the effects of possible co-solvents. To this end, we analyzed the change in reactivity of the four BAL variants in ten different reaction mixtures (Figure 4A). Since we found only minor differences for the reaction in co-solvent-free buffer at either pH 7 or pH 8, we chose pH 8 as the reference for the assessment of the effect of co-solvent addition, since it was employed also in earlier studies (for direct comparison of the activities of the different enzymes with each other, see Figure 4B).[15,38,40] We investigated reaction mixtures containing 40% choline-based deep eutectic solvents (DES) because this type of solvents has been reported to enable the conversion of hydrophobic substrates.[41] We found an overall lower conversion for all four enzymes in the presence of choline : glycerol (G-DES). Addition of 20% tert-butanol, 5% tetrahydrofuran (THF), 5% 2-methyltetrahydrofuran (2-MTHF) or 5% acetonitrile led to similar or a decreased enzymatic activity, and 20% dimethyl sulfoxide (DMSO) led to the already reported increased performance of *Pf*BAL.[15,38,42,43] Interestingly, *Re*BAL was not significantly affected by this solvent and the effect of 20% DMSO on enzymatic performance appears to be selective for the architecture of the binding site of *Pf*BAL, as the activity

Table 1. Evaluation of BAL-catalyzed conversion of **1a**–**p** into homo-coupling product **2** and aromatic-aliphatic cross-coupling products **4** with **3** (see Scheme 1) by TLC analysis (see Figure S12) after 20 h.[a] The numbers indicate %-values of total conversion, determined via grayscale analysis of spot intensities at 245 nm (Figure S11A). In some cases (*), the TLC method used could not clearly confirm product formation[b] or product formation could not be quantified.[c] In other cases, production formation could be clearly excluded (n.d., not detectable).

1 Entry	2 *Pf*BAL	2 *Re*BAL	2 *Re*BAL$_{mat}$	2 *Re*BAL$_{wid}$	4 *Pf*BAL	4 *Re*BAL	4 *Re*BAL$_{mat}$	4 *Re*BAL$_{wid}$
a	34	57	62	55	87	n.d.	5	20
b	60	68	66	64	87	6	20	16
c	41	17	22	16	87	n.d.	n.d.	n.d.
d	*[b,c]	*[b,c]	n.d.	n.d.	*[b]	*[b]	*[b]	*[b]
e	*[b]	*[b]	*[b]	*[b]	49	*[b]	*[b]	*[b]
f	35	17	22	18	94	*[b]	*[b,c]	*[b,c]
g	59	25	17	10	89	*[b,c]	*[b]	*[b,c]
h	89	4	71	11	*[b]	45	*[b]	*[b]
i	84	3	19	5	33	n.d.	9	n.d.
j	n.d.	n.d.	n.d.	n.d.	9	n.d.	11	6
k	3	n.d.	3	1	36	n.d.	n.d.	n.d.
l	99	16	45	11	46	2	3	n.d.
m	56	52	53	38	81	n.d.	n.d.	n.d.
n	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
o	75	63	66	70	51	4	43	10
p	28	9	13	20	n.d.	n.d.	n.d.	n.d.

[a] Reaction conditions: 5–10 mM aromatic aldehyde (**1a**–**p**) and 100 mM **3** in 50 mM KP_i buffer (0.25 mM ThDP, 2.5 mM $MgSO_4$, 20% DMSO, pH 8), 2 h at 30 °C; BAL final concentration: 5 μM; Reaction volume: 100 μL. [b] Possible overlay of homo-coupling, oxidation or autocatalytic product on the TLC. [c] Indication of product conversion due to appearance of spots upon irradiation at 365 nm (see Figure S12B).

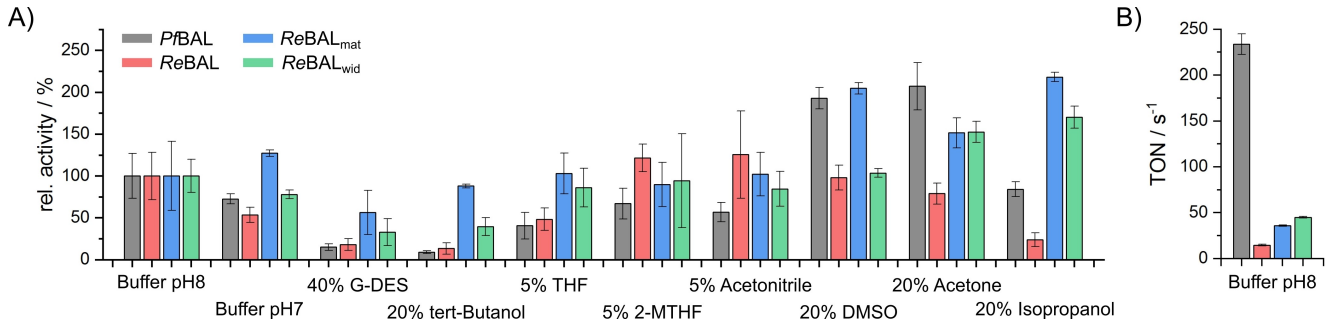


Figure 4. Conversion of **1 b** into **2 b** by *Pf*BAL (gray bars), *Re*BAL (red bars), *Re*BAL$_{mat}$ (blue bars) and *Re*BAL$_{wid}$ (green bars) after 1 h. A) Activities in the presence of different co-solvents (v/v) relative to the formation in potassium phosphate (KP_i) buffer pH 8. *Pf*BAL (20 nM) and *Re*BAL enzymes (200 nM) with 0.25 mM ThDP were incubated in mixtures of 50 mM KP_i buffer pH 8 and 2.5 mM $MgSO_4$ with different co-solvents and 30 mM **1 b** at 30 °C. Note that the activities are given for each BAL variant relative to the activity of the same variant at pH 8 without co-solvent to allow direct comparison of the effect of co-solvents on the individual BAL variants. To allow a direct comparison of the activities of the different enzymes with each other, the turnover number for all four BAL variants in potassium phosphate (KP_i) buffer pH 8 are shown in B). Analyses were carried out at least in technical and biological duplicates. Solvents: KP_i buffer pH 8, KP_i buffer pH 7, 40% choline : glycerol DES (G-DES), 20% *tert*-butanol, 5% tetrahydrofuran (THF), 5% 2-methyltetrahydrofuran (2-MTHF), 5% acetonitrile, 20% dimethyl sulfoxide (DMSO), 20% acetone, 20% isopropanol.

of *Re*BAL$_{mat}$ increased while *Re*BAL$_{wid}$ remained unaffected. The activity of *Pf*BAL also increased with the addition of 20% acetone, which had no significant effect on *Re*BAL but also increased the activity of *Re*BAL$_{mat}$ and *Re*BAL$_{wid}$. Surprisingly, the presence of 20% isopropanol also significantly increased the activity of *Re*BAL$_{mat}$ and *Re*BAL$_{wid}$. These results suggest a better interaction between the substrate and the less polar binding site architecture of *Re*BAL$_{mat}$ and *Re*BAL$_{wid}$ in the presence of these solvents. This is consistent with previously reported data in which a modulation of catalytic properties for thiamine-diphosphate-dependent enzymes was found due to a direct interaction of small co-solvent molecules with the nonpolar binding site.[42] In addition, the effects observed here may also be related to an interdependent effect of binding site architecture and surface hydrophobicity that affects enzyme performance (for a detailed discussion, see Figure S17). The finding of an increased performance with isopropanol was very promising for the application of the enzymes in flow reactors since the use of DMSO as a co-solvent can lead to problems during the workup procedure,[44] whereas isopropanol can be easily removed under reduced pressure.

To evaluate whether the newly identified *Re*BAL is suitable for the application in a continuous flow reactor, we chose the homo-coupling of **1 b** as a model reaction and the SpyCatcher/SpyTag system[45] for site-selective enzyme immobilization in a packed-bed reactor setup. This strategy allows an efficient and covalent immobilization under mild reaction conditions onto beads,[46] which has already been exploited, for example, to generate meso diols from ketones using alcohol dehydrogenases.[47] Prolonged reaction times are desirable in a packed-bed reactor, requiring stable enzymes. Here we used the ReBAL$_{wid}$ variant, since higher thermostability is often also associated with greater overall robustness.[48] To enable site-selective, covalent immobilization for the flow biocatalysis, the SpyTag (ST) peptide was genetically fused to the C-terminus of the enzyme to generate a *Re*BAL$_{wid}$-ST (for the corresponding primer sequences see Table S3), which can then be coupled to a solid support bearing the complementary SpyCatcher (SC) unit.[47]

After heterologous expression in *E. coli* and purification, we initially confirmed that the activity of the *Re*BAL$_{wid}$-ST is comparable to the other untagged *Re*BAL variants (Figure 5A). The ReBAL$_{wid}$-ST was then immobilized onto microbeads bearing the SC protein[49] and the resulting *Re*BAL$_{wid}$-ST@SC-Beads were loaded into custom-made PTFE microreactors (Figure S18) with a reactor volume of 200 µL. We selected 20% isopropanol as the co-solvent since it enhanced the performance of the enzyme (Figure 4A). Two packed-bed reactors were perfused in parallel with 30 mM **1 b** in 80% cofactor-containing KP_i buffer:20% isopropanol at a constant flow rate of 1 $\mu L \cdot min^{-1}$ (Figure 5B). The reactor outflows were sampled at defined time intervals and analyzed by reverse-phase HPLC. Under these reaction conditions, in total 20 mL of 1.3 ± 0.3 mM **2 b** were steadily generated over seven days (Figure 5C) with a space-time yield (STY) of 9.3 $mmol \cdot L^{-1} \cdot d^{-1}$. In comparison to other studies using *Pf*BAL in a continuous reaction system[16–19] the system developed here has the clear advantages of enabling long process times and employing simple process conditions such as homogeneous reaction media, easily scalable reactor design[50] e.g. via numbering-up.[51]

Conclusion

In conclusion, our work emphasizes that modeling-assisted rational design of *Re*BAL enzymes in combination with robust immobilization strategies is a powerful approach to create novel biocatalytic processes for continuous production of valued-added molecules. Specifically, we herein demonstrated the immobilization of a thermostable benzaldehyde lyase in a continuous packed-bed microreactor for the production of α-hydroxy ketones with constant activity over days. In this context, we screened three previously uncharacterized ThDP-dependent enzymes for aromatic ligase activity and identified

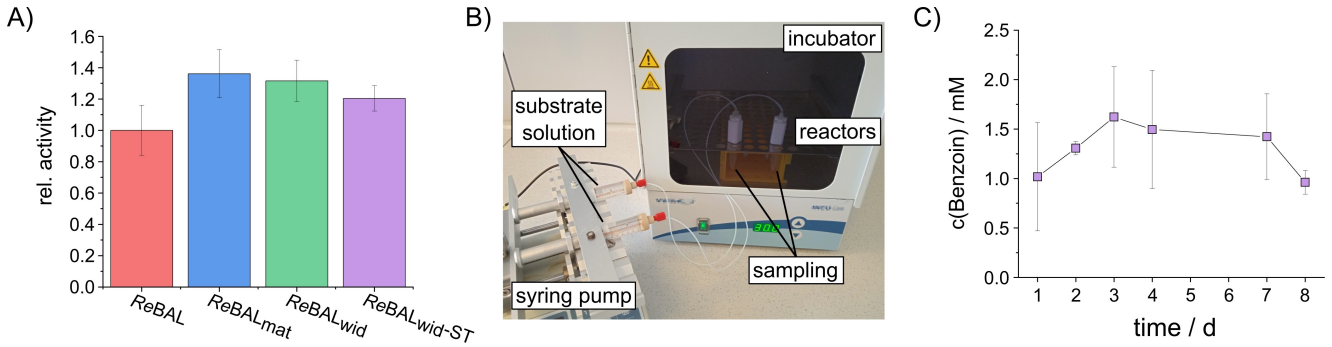


Figure 5. A) Enzymatic activity of *Re*BAL, *Re*BAL$_{mat}$ and *Re*BAL$_{wid}$ as well as the *Re*BAL$_{wid}$-ST at 30 °C compared with the activity of wild-type *Re*BAL, determined by using the 2-furaldehyde (**1 a**) assay. B) Flow reactor setup of two parallel reactors. The flow reactors on top of 1.5 mL collection tubes are placed inside an incubator at 30 °C and are perfused with substrate solution using a syringe pump. For more detailed information on the reactor assembly, see Figure S18. C) Concentration of **2 b** quantified by HPLC in the outflow of the two packed-bed reactors at a flow rate of 1 μL min^{-1}. Reaction condition: 30 mM **1 b** in 50 mM KP$_i$ buffer (0.25 mM ThDP, 2.5 mM $MgSO_4$, 20% isopropanol, pH 8) at 30 °C.

*Re*BAL as a novel benzaldehyde lyase. After modeling the protein structure of *Re*BAL, two rationally designed variants with altered binding sites were generated. Compering these three benzaldehyde lyase variants with the literature described *Pf*BAL, we showed that *Re*BAL is more thermostable while being less active and having a narrower substrate spectrum. However, the *Re*BAL$_{mat}$ variant whose binding site was designed to match with *Pf*BAL recovered the substrate scope. The most thermostable *Re*BAL$_{wid}$ variant was then immobilized for the continuous flow biocatalysis where it showed a stable conversion over seven days. Our experiments expand the availability of enzymes for the continuous production of α-hydroxy ketones and therefore can enhance the usability of this important enzyme class in biotechnological production processes. Since thermostable enzymes in general tolerate more mutations,[52] we envision that the more stable *Re*BAL backbone identified in this work can be further modified using protein engineering to create completely novel reactivities and produce important enantiopure building blocks for the synthesis of APIs.

Experimental Section

Experimental details can be found in the Supporting Information.

Acknowledgements

We thank the Institute for Astroparticle Physics (IAP) for manufacturing the PTFE microreactors. The work was supported through the Helmholtz program "Materials Systems Engineering" under the topic "Adaptive and Bioinstructive Materials Systems". Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: C–C coupling · enzyme catalysis · immobilization · lyases · protein engineering

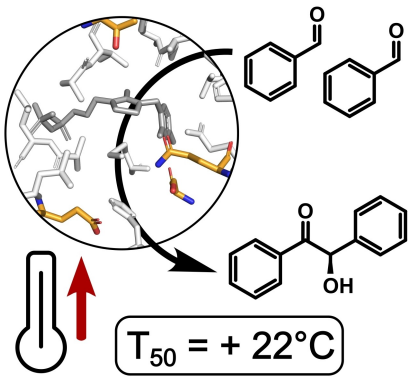
[1] P. Hoyos, J.-V. Sinisterra, F. Molinari, A. R. Alcántara, P. Domínguez de María, *Acc. Chem. Res.* **2010**, *43*, 288–299.
[2] R. S. Martins, D. S. Zampieri, J. A. R. Rodrigues, P. O. Carvalho, P. J. S. Moran, *ChemCatChem* **2011**, *3*, 1469–1473.
[3] M. Rabuffetti, P. Cannazza, M. L. Contente, A. Pinto, D. Romano, P. Hoyos, A. R. Alcantara, I. Eberini, T. Laurenzi, L. Gourlay, F. Di Pisa, F. Molinari, *Bioorg. Chem.* **2021**, *108*, 104644.
[4] B. González, R. Vicuña, *J. Bacteriol.* **1989**, *171*, 2401–2405.
[5] a) P. Hinrichsen, I. Gomez, R. Vicuña, *Gene* **1994**, *144*, 137–138; b) E. Janzen, M. Pohl, *"Pseudomonas fluorescens biovar I benzaldehyde lyase (bznB) gene, complete cds"*, can be found under https://www.ncbi.nlm.nih.gov/nuccore/9965497, **2000**.
[6] A. Chang, L. Jeske, S. Ulbrich, J. Hofmann, J. Koblitz, I. Schomburg, M. Neumann-Schaal, D. Jahn, D. Schomburg, *Nucleic Acids Res.* **2021**, *49*, D498-D508.
[7] P. C. F. Buchholz, C. Vogel, W. Reusch, M. Pohl, D. Rother, A. C. Spieß, J. Pleiss, *ChemBioChem* **2016**, *17*, 2093–2098.
[8] M. Beigi, E. Gauchenova, L. Walter, S. Waltzer, F. Bonina, T. Stillger, D. Rother, M. Pohl, M. Müller, *Chem. Eur. J.* **2016**, *22*, 13999–14005.
[9] a) A. Cosp, C. Dresen, M. Pohl, L. Walter, C. Röhr, M. Müller, *Adv. Synth. Catal.* **2008**, *350*, 759–771; b) K. Hernández, T. Parella, J. Joglar, J. Bujons, M. Pohl, P. Clapés, *Chem. Eur. J.* **2015**, *21*, 3335–3346; c) A. S. Demir, Ö. Şeşenoglu, P. Dünkelmann, M. Müller, *Org. Lett.* **2003**, *5*, 2047–2050; d) A. S. Demir, Ö. Şeşenoglu, E. Eren, B. Hosrik, M. Pohl, E. Janzen, D. Kolter, R. Feldmann, P. Dünkelmann, M. Müller, *Adv. Synth. Catal.* **2002**, *344*, 96; e) M. Pérez-Sánchez, C. R. Müller, P. Domínguez de María, *ChemCatChem* **2013**, *5*, 2512–2516; f) P. Ayhan, İ. Şimşek, B. Çifçi, A. S. Demir, *Org. Biomol. Chem.* **2011**, *9*, 2602–2605.
[10] P. Ayhan, A. S. Demir, *Adv. Synth. Catal.* **2011**, *353*, 624–629.
[11] K. Hernández, T. Parella, G. Petrillo, I. Usón, C. M. Wandtke, J. Joglar, J. Bujons, P. Clapés, *Angew. Chem. Int. Ed.* **2017**, *56*, 5304–5307; *Angew. Chem.* **2017**, *129*, 5388–5391.
[12] X. Chen, Z. Wang, Y. Lou, Y. Peng, Q. Zhu, J. Xu, Q. Wu, *Angew. Chem. Int. Ed.* **2021**, *60*, 9326–9329.
[13] M. M. Kneen, I. D. Pogozheva, G. L. Kenyon, M. J. McLeish, *Biochim. Biophys. Acta* **2005**, *1753*, 263–271.
[14] J. B. Siegel, A. L. Smith, S. Poust, A. J. Wargacki, A. Bar-Even, C. Louw, B. W. Shen, C. B. Eiben, H. M. Tran, E. Noor et al., *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3704–3709.
[15] E. Janzen, M. Müller, D. Kolter-Jung, M. M. Kneen, M. J. McLeish, M. Pohl, *Bioorg. Chem.* **2006**, *34*, 345–361.
[16] F. Hildebrand, S. Kühl, M. Pohl, D. Vasic-Racki, M. Müller, C. Wandrey, S. Lütz, *Biotechnol. Bioeng.* **2007**, *96*, 835–843.
[17] T. Stillger, M. Pohl, C. Wandrey, A. Liese, *Org. Process Res. Dev.* **2006**, *10*, 1172–1177.

[18] M. B. Ansorge-Schumacher, L. Greiner, F. Schroeper, S. Mirtschin, T. Hischer, *Biotechnol. J.* **2006**, *1*, 564–568.
[19] N. Kurlemann, A. Liese, *Tetrahedron: Asymmetry* **2004**, *15*, 2955–2958.
[20] S. B. Sopaci, I. Simşek, B. Tural, M. Volkan, A. S. Demir, *Org. Biomol. Chem.* **2009**, *7*, 1658–1664.
[21] J. Döbber, M. Pohl, *J. Biotechnol.* **2017**, *241*, 170–174.
[22] a) R. Kloss, T. Karmainski, V. D. Jäger, D. Hahn, A. Grünberger, M. Baumgart, U. Krauss, K.-E. Jaeger, W. Wiechert, M. Pohl, *Catal. Sci. Technol.* **2018**, *8*, 5816–5826; b) V. D. Jäger, M. Piqueray, S. Seide, M. Pohl, W. Wiechert, K.-E. Jaeger, U. Krauss, *Adv. Synth. Catal.* **2019**, *351*, 1842; c) V. D. Jäger, R. Lamm, R. Kloß, E. Kaganovitch, A. Grünberger, M. Pohl, J. Büchs, K.-E. Jaeger, U. Krauss, *ACS Synth. Biol.* **2018**, *7*, 2282–2295.
[23] G. Li, K. S. Rabe, J. Nielsen, M. K. M. Engqvist, *ACS Synth. Biol.* **2019**, *8*, 1411–1420.
[24] Y. Dehouck, B. Folch, M. Rooman, *Protein Eng. Des. Sel.* **2008**, *21*, 275–278.
[25] D. Natalia, C. Kohlmann, M. B. Ansorge-Schumacher, L. Greiner, *Biotechnol. Res. Int.* **2011**, 478925.
[26] Y. Liu, Y. Li, X. Wang, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8633–8649.
[27] A. Kwasiborski, S. Mondy, T.-M. Chong, K.-G. Chan, A. Beury-Cirou, D. Faure, *Genetica* **2015**, *143*, 253–261.
[28] G. S. Brandt, N. Nemeria, S. Chakraborty, M. J. McLeish, A. Yep, G. L. Kenyon, G. A. Petsko, F. Jordan, D. Ringe, *Biochemistry* **2008**, *47*, 7734–7743.
[29] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, *Nucleic Acids Res.* **2018**, *46*, W296-W303.
[30] M. Bertoni, F. Kiefer, M. Biasini, L. Bordoli, T. Schwede, *Sci. Rep.* **2017**, *7*, 10480.
[31] S. Bienert, A. Waterhouse, T. A. P. de Beer, G. Tauriello, G. Studer, L. Bordoli, T. Schwede, *Nucleic Acids Res.* **2017**, *45*, D313-D319.
[32] P. Benkert, M. Biasini, T. Schwede, *Bioinformatics* **2011**, *27*, 343–350.
[33] N. Guex, M. C. Peitsch, T. Schwede, *Electrophoresis* **2009**, *30*, S162–73.
[34] T. G. Mosbacher, M. Müller, G. E. Schulz, *FEBS J.* **2005**, *272*, 6067–6076.
[35] M. Knoll, M. Müller, J. Pleiss, M. Pohl, *ChemBioChem* **2006**, *7*, 1928–1934.
[36] W. Tian, C. Chen, X. Lei, J. Zhao, J. Liang, *Nucleic Acids Res.* **2018**, *46*, W363-W367.
[37] a) Q. Meng, N. Capra, C. M. Palacio, E. Lanfranchi, M. Otzen, L. Z. van Schie, H. J. Rozeboom, A.-M. W. H. Thunnissen, H. J. Wijma, D. B. Janssen, *ACS Catal.* **2020**, *10*, 2915–2928; b) V. Moore, A. Kanu, O. Byron, G. Campbell, M. J. Danson, D. W. Hough, S. J. Crennell, *Extremophiles* **2011**, *15*, 327–336; c) A. Bashir, R. N. Perham, N. S. Scrutton, A. Berry, *Biochem. J.* **1995**, *312*, 527–533.
[38] M. Kokova, M. Zavrel, K. Tittmann, A. C. Spiess, M. Pohl, *J. Mol. Catal. B* **2009**, *61*, 73–79.
[39] a) T. R. Hoye, C. S. Jeffrey, F. Shao, *Nat. Protoc.* **2007**, *2*, 2451–2458; b) J. A. Dale, H. S. Mosher, *J. Am. Chem. Soc.* **1973**, *95*, 512–519.
[40] a) J. Donnelly, C. R. Müller, L. Wiermans, C. J. Chuck, P. Domínguez de María, *Green Chem.* **2015**, *17*, 2714–2718; b) Z. Maugeri, P. Domínguez de María, *J. Mol. Catal. B* **2014**, *107*, 120–123.
[41] a) A. K. Schweiger, N. Ríos-Lombardía, C. K. Winkler, S. Schmidt, F. Morís, W. Kroutil, J. González-Sabín, R. Kourist, *ACS Sustainable Chem. Eng.* **2019**; b) G. de Gonzalo, C. Martin, M. W. Fraaije, *Catalysts* **2020**, *10*, 447; c) Y. Fredes, L. Chamorro, Z. Cabrera, *Molecules* **2019**, *24*, 792.
[42] T. Gerhards, U. Mackfeld, M. Bocola, E. von Lieres, W. Wiechert, M. Pohl, D. Rother, *Adv. Synth. Catal.* **2012**, *354*, 2805–2820.
[43] T. Schmidt, M. Zavrel, A. Spieß, M. B. Ansorge-Schumacher, *Bioorg. Chem.* **2009**, *37*, 84–89.
[44] a) P. Domínguez de María, T. Stillger, M. Pohl, M. Kiesel, A. Liese, H. Gröger, H. Trauthwein, *Adv. Synth. Catal.* **2008**, *350*, 165–173; b) M. M. C. H. van Schie, J.-D. Spöring, M. Bocola, P. Domínguez de María, D. Rother, *Green Chem.* **2021**, *23*, 3191–3206.
[45] B. Zakeri, J. O. Fierer, E. Celik, E. C. Chittock, U. Schwarz-Linek, V. T. Moy, M. Howarth, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E690–E697.
[46] a) A. H. Keeble, M. Howarth, *Chem. Sci.* **2020**, *11*, 7281–7291; b) G. Zhang, M. B. Quin, C. Schmidt-Dannert, *ACS Catal.* **2018**, *8*, 5611–5620; c) T. Peschke, P. Bitterwolf, S. Gallus, Y. Hu, C. Oelschlaeger, N. Willenbacher, K. S. Rabe, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2018**, *57*, 17028–17032; *Angew. Chem.* **2018**, *130*, 17274–17278; d) P. Bitterwolf, F. Ott, K. S. Rabe, C. M. Niemeyer, *Micromachines* **2019**, *10*, 783; e) E. Mittmann, S. Gallus, P. Bitterwolf, C. Oelschlaeger, N. Willenbacher, C. M. Niemeyer, K. S. Rabe, *Micromachines* **2019**, *10*, 795; f) Y. Lin, G. Zhang, *ACS Appl. Nano Mater.* **2020**, *3*, 44–48; g) F. Sun, W.-B. Zhang, A. Mahdavi, F. H. Arnold, D. A. Tirrell, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11269–11274; h) C. Hentrich, S.-J. Kellmann, M. Putyrski, M. Cavada, H. Hanuschka, A. Knappik, F. Ylera, *Cell Chem. Biol.* **2021**, *28*, 813–824; i) M. Romero-Fernández, F. Paradisi, *Curr. Opin. Chem. Biol.* **2020**, *55*, 1–8; j) M. Romero-Fernández, F. Paradisi in *Catalyst Immobilization: Methods and Applications* (Eds.: M. Benaglia, A. Puglisi), Wiley-VCH, Weinheim, **2020**, pp. 409–435.
[47] T. Peschke, M. Skoupi, T. Burgahn, S. Gallus, I. Ahmed, K. S. Rabe, C. M. Niemeyer, *ACS Catal.* **2017**, *7*, 7866–7872.
[48] J. K. Kristjansson, *Trends Biotechnol.* **1989**, *7*, 349–353.
[49] T. Peschke, K. S. Rabe, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2017**, *56*, 2183–2186; *Angew. Chem.* **2017**, *129*, 2215–2219.
[50] T. Burgahn, P. Pietrek, R. Dittmeyer, K. S. Rabe, C. M. Niemeyer, *ChemCatChem* **2020**, *12*, 2452–2460.
[51] P. Bitterwolf, S. Gallus, T. Peschke, E. Mittmann, C. Oelschlaeger, N. Willenbacher, K. S. Rabe, C. M. Niemeyer, *Chem. Sci.* **2019**, *10*, 9752–9757.
[52] a) W. Besenmatter, P. Kast, D. Hilvert, *Proteins* **2007**, *66*, 500–506; b) D. M. Taverna, R. A. Goldstein, *J. Mol. Biol.* **2002**, *315*, 479–484.

Manuscript received: September 2, 2021
Revised manuscript received: September 23, 2021
Accepted manuscript online: September 24, 2021
Version of record online: ■■, ■■■

FULL PAPERS

A novel benzaldehyde lyase for the production of chiral α-hydroxy ketones was identified by a machine learning model and two variants with altered substrate scopes were created by rational design. All three variants exhibit significantly higher thermostability compared to the only previously described benzaldehyde lyase. When immobilized in a packed-bed reactor, the most thermostable variant showed stable substrate turnover over days.



*M. Peng, D. L. Siebert, Dr. M. K. M. Engqvist, Prof. Dr. C. M. Niemeyer, Dr. K. S. Rabe**