Rational Protein Design of ThDP-Dependent Enzymes— Engineering Stereoselectivity

Dörte Gocke,[a] Lydia Walter,[b] Ekaterina Gauchenova,[b] Geraldine Kolter,[a] Michael Knoll,[c] Catrine L. Berthold,[d] Gunter Schneider,[d] Jürgen Pleiss,[c] Michael Müller,[b] and Martina Pohl*[a]

Benzoylformate decarboxylase (BFD) from Pseudomonas putida is an exceptional thiamin diphosphate-dependent enzyme, as it catalyzes the formation of (S)-2-hydroxy-1-phenylpropan-1-one from benzaldehyde and acetaldehyde. This is the only currently known S-selective reaction (92% ee) catalyzed by this otherwise R-selective class of enzymes. Here we describe the molecular basis of the introduction of S selectivity into ThDP-dependent decarboxylases. By shaping the active site of BFD through the use of rational protein design, structural analysis, and molecular modeling, optimal steric stabilization of the acceptor aldehyde in a structural element called the S pocket was identified as the predominant interaction for adjusting stereoselectivity. Our studies revealed Leu461 as a hot spot for stereoselectivity in BFD. Exchange to alanine and glycine resulted in variants that catalyze the S-stereoselective addition of larger acceptor aldehydes, such as propanal with benzaldehyde and its derivatives—a reaction not catalyzed by the wild-type enzyme. Crystal structure analysis of the variant BFDL461A supports the modeling studies.

Introduction

The potential of thiamin diphosphate-dependent (ThDP-dependent) enzymes to catalyze benzoin condensation-like carboligation of aldehydes to afford chiral 2-hydroxyketones with high stereoselectivity is well established.[1] Our goal is to generate a toolbox of various ThDP-dependent enzymes in order to create a platform for the production of diversely substituted and enantiocomplementary 2-hydroxyketones.

With the current set of enzymes—including benzoylformate decarboxylase (BFD), benzaldehyde lyase (BAL), branched-chain 2-ketoacid decarboxylase (KdcA), different pyruvate decarboxylases (PDCs), and their variants—the carboligation of various aliphatic and aromatic aldehydes to yield symmetrical and mixed (R)-2-hydroxyketones predominantly with high enantioselectivity is possible. However, the corresponding S products are hardly accessible by these enzymes. Exceptions are the kinetic resolution of benzoin derivatives by BAL from Pseudomonas fluorescens[2] and the carboligation of benzaldehyde derivatives and acetaldehyde to yield (S)-2-hydroxypropiophenone derivatives with the aid of BFD from Pseudomonas putida as a catalyst.[3]

A molecular explanation for this exceptional behavior of BFD was recently suggested[4] based on the crystal structure of the enzyme.[5] A potential S pocket that exactly fits with the size of the small acetaldehyde side chain when approaching the ThDP-bound aromatic donor aldehyde prior to formation of the new C—C-bond was identified (Figure 1).[4] Recent modeling studies showed that larger aldehydes do not fit into this pocket.

Here, we have verified the S pocket approach by site-directed mutagenesis of the amino acid residues that line this part of the active center and by testing the resulting variants in carboligation reactions with different aliphatic aldehydes as acyl acceptors. Further, we provide evidence that similar S pockets are also present in other ThDP-dependent decarboxylases, which opens access to a broad range of (S)-2-hydroxyketones as valuable building blocks for compounds such as the taxol side chain and 5′-methoxyhydnocarpin.[6]

Results and Discussion

The S pocket in BFD from *P. putida* is formed by the side chains of Pro24, Ala460, and predominantly Leu461 (Figure 1). These residues were replaced by smaller amino acids in order to evaluate their impact on the S pocket. All variants were produced by site-directed mutagenesis. After cloning, over-expression,

[a] D. Gocke, G. Kolter, Dr. M. Pohl
Institute of Molecular Enzyme Technology
Heinrich-Heine University Düsseldorf
52426 Jülich (Germany)
Fax: (+49) 2461-612940
E-mail: ma.pohl@fz-juelich.de

[b] L. Walter, Dr. E. Gauchenova, Prof. Dr. M. Müller
Institute of Pharmaceutical Sciences
Albert-Ludwigs University Freiburg
Albertstrasse 25, 79104 Freiburg (Germany)

[c] M. Knoll, Prof. Dr. J. Pleiss
Institute of Technical Biochemistry, University of Stuttgart
Allmandring 31, 70569 Stuttgart (Germany)

[d] C. L. Berthold, Prof. Dr. G. Schneider
Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Tomtebodavägen 6
17177 Stockholm (Sweden)
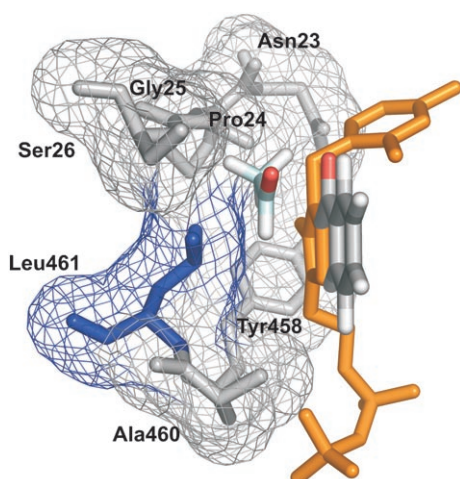
Figure 1. Active site of BFD with the cofactor ThDP (orange). The donor benzaldehyde (gray) and the acceptor acetaldehyde (light blue) were modeled inside. The side chain of acetaldehyde is bound in the *S* pocket, which is mainly defined by Leu461 (blue). The diametrically opposed orientation of the side chains of donor and acceptor results in the formation of *S* products.[6]

and purification the variants were investigated with respect to their decarboxylase and carboligase activities.

Decarboxylase activity of BFD variants

All variants were able to catalyze out the physiological function of BFD—the decarboxylation of benzoylformate (Table 1); they showed hyperbolic *v*/[S] plots like wild-type BFD (BFDwt)

Table 1. Kinetic data determined for the decarboxylation of benzoylformate catalyzed by BFDwt and several BFD variants.[a]

Enzyme	V_{max} [U mg^{-1}]	K_M [mM]
BFDwt	400 ± 7	0.37 ± 0.03
BFDP24A	367 ± 10	0.51 ± 0.07
BFDA460G	300 ± 9	1.54 ± 0.14
BFDL461V	60 ± 1	0.06 ± 0.01
BFDL461A	53 ± 2	0.25 ± 0.04
BFDL461G	49 ± 2	1.40 ± 0.16

[a] Data were measured in potassium phosphate buffer (pH 6.5, 50 mM). Kinetic parameters were calculated according to Michaelis–Menten by use of Origin 7.0 (Origin Lab Corporation, Northampton, MA, USA).

but with a decreased decarboxylase activity toward benzoylformate. The most pronounced effects were obtained with mutation at position Leu461; this yielded variants with seven- to eightfold decreased specific decarboxylase activity compared to BFDwt. It is important to note that the K_M values do not parallel the V_{max} values. Whereas BFDP24A and BFDL461A show K_M values in the same range as the wild-type enzyme, the variant BFDL461V had a six-times higher apparent affinity for the substrate, while the K_M values of the two glycine variants (Ala460Gly, Leu461Gly) were about four- to five-times higher than that of BFDwt. The substrate range of the decarboxylase reaction was not affected by the mutations, with benzoylfor-

mate being the main substrate for all variants (see the Supporting Information).

Carboligase activity of BFD variants

All variants were investigated with respect to their carboligase activity toward the self ligation of benzaldehyde to afford benzoin and acetaldehyde to give acetoin (Table 2). Mutations in

Table 2. Space–time yields and enantioselectivities of different BFD variants involving the formation of acetoin from acetaldehyde and benzoin from benzaldehyde.

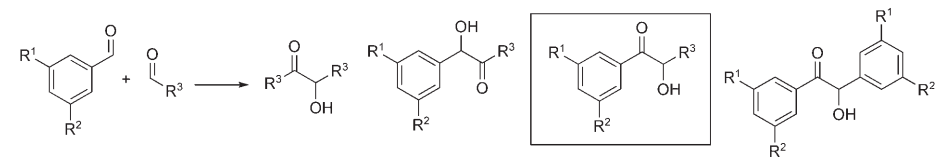| Enzyme	Acetoin [g L^{-1} d^{-1}]	*ee* [%]	Benzoin [g L^{-1} d^{-1}]	*ee* [%]
BFDwt	0.11	34 (*R*)	1.08[b]	n.d.
BFDP24A	0.19	3 (*R*)	0.58[b]	n.d.
BFDA460G	0.07	38 (*R*)	0.23[b]	n.d.
BFDL461V	0.1	33 (*R*)	1.75[a]	n.d.
BFDL461A	0.07	64 (*S*)	0.14[a]	n.d.
BFDL461G	n.d.	n.d.	n.d.	n.d.

Space–time yields were calculated within the linear ranges of [a] 3 h and [b] 2 h; n.d.: not determined.

the putative *S* pocket affected both acetoin and benzoin synthesis. The Pro24Ala and Leu461Val variants were most effective in terms of catalyzing the acetoin synthesis; the Leu461Val variant also catalyzed the benzoin synthesis 1.6-times more rapidly than BFDwt. The mutations affected the stereoselectivity of acetoin formation significantly (Table 2): while BFDwt predominantly catalyzes the formation of (*R*)-acetoin (**1**, Table 3; *ee* 34%),[7] the *S* enantiomer (*ee* 65%) was formed in excess with BFDL461A; this supports the relevance of Leu461 for the shape of the *S* pocket.

Further studies with mixed carboligations were focused on the glycine and alanine variants in position Leu461. In mixed carboligations with benzaldehyde and acetaldehyde the stereoselectivity of the (*S*)-2-hydroxy-1-phenylpropan-1-one ((*S*)-HPP, **3**) synthesis was improved by both mutations (*ee* 98%) relative to BFDwt (92%; Table 3 A).[3]

HPP is formed with benzaldehyde as the donor and acetaldehyde as the acceptor aldehyde (Figure 1). As predicted from modeling studies, the most pronounced effect of an enlarged *S* pocket should become apparent if acetaldehyde is replaced by the larger propanal; this was confirmed by analytical studies (Table 3 B). Both variants in position 461 catalyzed the synthesis of the desired (*S*)-2-hydroxy-1-phenylbutan-1-one product ((*S*)-**7**) with high stereoselectivity, and showed even higher enantioselectivities at pH 7.9 than under standard conditions (pH 7.0); variation of the substrate concentration and reaction temperature had no significant effect on the *ee* values (data not shown).

According to the predictions based on the structure of BFDL461A, a further increase in the size of the acceptor aldehyde should decrease the stereoselectivity again. This assumption was experimentally confirmed by application of butanal in

Table 3. Relative product distributions and enantiomeric excesses (*ee*) obtained in analytical scale carboligation reactions of benzaldehyde derivatives and various aliphatic aldehydes catalyzed by BFD variants.[a]

A) $R^1=R^2=H$, $R^3=CH_3$

	1	2	3	4
BFDwt	–	–	81%	4%
			92% (S)	99% (R)
BFDL461A	–	–	79%	–
			98% (S)	
BFDL461G	–	–	74.5%	–
			98% (S)	

B) $R^1=R^2=H$, $R^3=C_2H_5$

	5	6	7	4
BFDwt	8%	12%	6%	8%
	ee n.d.	98% (R)	21% (R)	99% (R)
BFDL461A	6%	–	21.5%	1.5%
	ee n.d.		88%/93%[b] (S)	*ee* n.d.
BFDL461G	–	–	23%	0.5%
			93%/97%[b] (S)	*ee* n.d.

C) $R^1=R^2=H$, $R^3=C_3H_7$

	8	9	10	4
BFDwt	34%	32%	2%	3%
	ee n.d.	99% (R)	66% (R)	99% (R)
BFDL461A	8%	3%	1%	1%
	ee n.d.	>99% (R)	63% (S)	97% (R)
BFDL461G	6%	2%	1%	–
	ee n.d.	*ee* n.d.	*ee* n.d.	

D) $R^1=R^2=H$, $R^3=OCH_3$

	11	12	13	4
BFDL461A	–	–	9.5%, 21%[c]	<1%
			93% (S)	*ee* n.d.

E) $R^1=R^2=H$, $R^3=$ cyclopropyl

	15	16	17	4
BFDL461A	–	–	5%	–
			97.5% (S)	

F) $R^1=R^2=OCH_3$, $R^3=C_2H_5$

	1	18	19	20
BFDwt	–	–	<1%	–
			ee n.d.	
BFDL461A	––	7.5%; 9%[c]	–	
		>99% (S)		
BFDL461G	–	–	31%	–
			>99% (S)	

G) $R^1=R^2=OCH_3$, $R^3=C_3H_7$

	21	22	23	20
BFDwt	19%	26%	–	–
	ee n.d.	*ee* n.d.		
BFDL461A	–	–	–	–
BFDL461G	2%	26%	4%	–
	ee n.d.	*ee* n.d.	*ee* n.d.	

H) $R^1=R^2=OCH_3$, $R^3=C_4H_{11}$

	24	25	26	20
BFDL461A	6%	4%	11%	<1%
	ee n.d.	*ee* n.d.	>90% (S)	*ee* n.d.

I) $R^1=CN$, $R^2=H$, $R^3=C_2H_5$

	1	27	28	29
BFDL461A	–	–	22.5%, 60%[c]	
			98% (S)	

[a] Relative product distributions are given in mol% (NMR); *ee* values were determined by HPLC analysis. All studies were performed with equimolar concentrations of both aldehydes (18 mM) in potassium phosphate buffer (50 mM, pH 7, 2.5 mM MgSO₄, 0.1 mM ThDP, 20 vol% DMSO), with 0.3 mg mL⁻¹ purified enzyme, at 30 °C, unless otherwise indicated. [b] Carboligations were performed at pH 7.9. [c] Carboligations were performed with a threefold excess of the aliphatic aldehyde (54 mM).

mixed carboligation reactions with benzaldehyde to yield **10** with decreased stereoselectivity (*ee* 63%, (*S*); Table 3C) and activity. As well as propanal, monomethoxyacetaldehyde also functions as an acceptor aldehyde in the presence of benzaldehyde for BFDL461A and results in a yield of about 20% in analytical biotransformations (Table 3D). In contrast, cyclopropanecarbaldehyde yielded smaller amounts of the mixed carboligation product but with high selectivity (Table 3E).

The stereocontrol in mixed carboligations with propanal was even better with 3,5-dimethoxybenzaldehyde and 3-cyanobenzaldehyde than with benzaldehyde (Table 3F and I). Very interesting results were obtained in carboligations of 3,5-dimethoxybenzaldehyde with butanal and pentanal. Whereas no product was obtained with the BFDL461A variant, BFDL461G was able to catalyze the mixed carboligation with pentanal even better than with butanal (Table 3G and H).

The carboligation of benzaldehyde and propanal was investigated in more detail on a preparative scale. After 70 h, BFDL461A had produced a 35% yield (*w/w*) and BFDL461G a 31% yield (*w/w*) of product **7**, while negligible amounts of **4**

and **5** were formed; this demonstrates very high chemoselectivity for both variants. Confirming the predictions of the modeling studies, the desired product (*S*)-2-hydroxy-1-phenylbutan-1-one ((*S*)-**7**) was formed with very high stereoselectivity of 93–97% with variants BFDL461A and BFDL461G, which is in contrast to BFDwt (yields (*R*)-**7**; *ee* 21%; Table 4). Although carboligation with BFDwt resulted in a total conversion of 37%, a mixture of **4**, **6**, and **7** in almost equal amounts was formed; this demonstrates the low chemo- and stereoselectivity of BFDwt for this reaction.

Table 4. Preparative scale carboligation of benzaldehyde and propanal catalyzed by different BFD variants. Products were isolated by flash column chromatography.

Substrate/product[a]	BFDwt (*ee*)	BFDL461A (*ee*)	BFDL461G (*ee*)
benzaldehyde	1.1%	–	–
4	35.6% (96% *R*)	–	–
6	36.7% (98% *R*)	–	–
7	26.7% (21% *R*)	>99% (93% *S*)	>99% (97% *S*)
isolated yield (**7**)	16.2 mg 37 mol% (*w/w*)	15 mg 35 mol% (*w/w*)	13.7 mg 31 mol% (*w/w*)

[a] Numbers refer to Table 3. Product compositions are given in mol% (as obtained by NMR spectroscopy); *ee* values were determined by chiral HPLC.

Structural investigation of BFDL461A

In order to verify that the site-specific mutagenesis did not alter the 3D structure of the enzyme beyond the exchanged amino acid, the crystal structure of the BFDL461A variant was solved with a resolution of 2.2 Å. Despite the desired increase in the size of the *S* pocket, no significant structural modifications were detected. In contrast to BFDwt, the *S* pocket in the BFDL461A variant offers optimal space for the ethyl group of propanal and provides higher stereoselectivity (Figure 2B and D).

Our data show that stereoselectivity is predominantly a consequence of the optimal stabilization of the acceptor aldehyde side chain in the *S* pocket. If this fit is not optimal, as is already observed for BFDwt with propanal (Figure 2A and C) and for BFDL461A with butanal, the *S* selectivity is reduced (Table 3C).

Conclusions

We have successfully engineered *S*-specific BFD variants using a structure-guided approach. By investigating the origin of *S* selectivity we have demonstrated the potential to shape this part of the active-site selectively for longer chain aliphatic acceptor aldehydes. The experimental data are very readily predictable by modeling studies, which allows the in silico design of *S*-specific biocatalysts for special requirements. In order to generalize this strategy the 3D structures of other ThDP-dependent enzymes related to BFD have been compared.
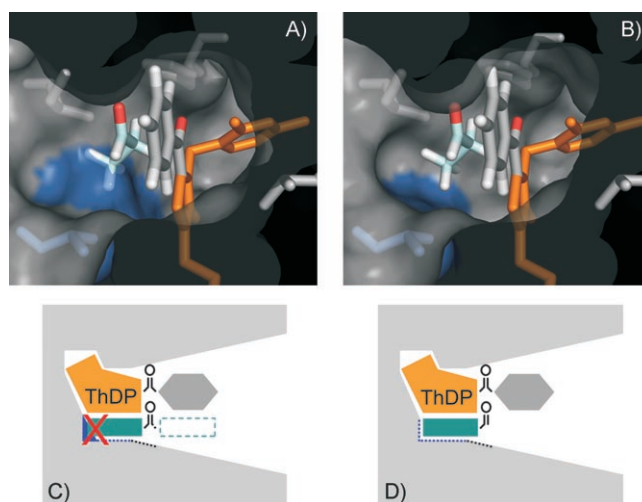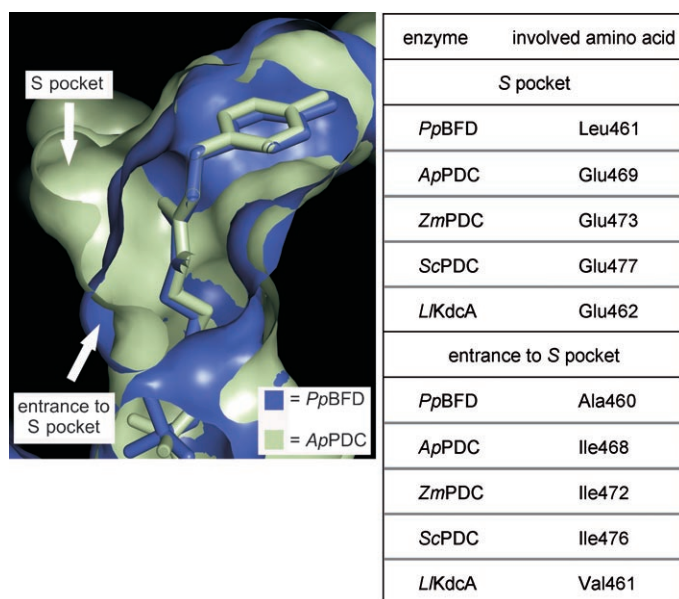

Figure 2. Crystal structures of the active sites of A) BFDwt and B) BFDL461A with benzaldehyde and propanal modeled inside. The side chain of the amino acid residue in position 461 is marked in blue. Benzaldehyde (gray) is bound to the C2 atom of the thiazolium ring (orange) and is arranged in coplanar fashion, due to steric and electronic demands. Propanal (light blue) is located in the *S* pocket. Models show a perfect stabilization of the acceptor aldehyde in the *S* pocket of the variant (B, D), while Leu461 causes steric hindrance with propanal in BFDwt (A, C). Consequently, C) in BFDwt the propanal predominantly approaches parallel to benzaldehyde (dotted square) to yield mainly the *R* enantiomer, while D) in BFDL461A the perfect fitting allows an antiparallel arrangement; this leads to an excess of the *S* product.

A superimposition of the crystal structures of BFDwt,[5] BAL from *P. fluorescens*,[8] PDCs from *Zymomonas mobilis* (*Zm*PDC)[9a] and *Saccharomyces cerevisiae* (*Sc*PDC),[9b] as well as the recently solved structures of PDC from *Acetobacter pasteurianus* (*Ap*PDC),[10] and KdcA from *Lactococcus lactis* (*Ll*KdcA),[11–13] gave profound insights. While there is no *S* pocket visible in BAL, the *S* pockets of the other enzymes increased in the series *Pp*BFD < *Ll*KdcA < *Zm*PDC/*Sc*PDC < *Ap*PDC (Figure 3).

However, the entrances to the *S* pockets in KdcA and both PDCs are restricted by bulky residues, such as isoleucine or valine, which could explain why all these enzymes are strictly *R* selective. Consequently, (*S*)-2-hydroxyketones could be formed by improving the access to this pocket. This has successfully been shown with the variant *Zm*PDCI472A, which catalyzes the formation of (*S*)-HPP (**3**; *ee* 70%), while exclusively (*R*)-phenylacetylcarbinol (**2**; *ee* > 98%) is formed with the wild-type enzyme with benzaldehyde and acetaldehyde as substrates.[14]

The predominant *R* selectivity of ThDP-dependent enzymes is therefore soundly explicable mainly from their structures. However, many of these enzymes have the latent inherent property of *S* selectivity, since such *S* pockets are visible in almost all the 3D structures mentioned above although they are not accessible in many cases. Our results pave the way for expanding the shaping strategy of the *S* pockets to a broad range of other 2-ketoacid decarboxylases. This is a powerful tool for enlarging the toolbox of enzymatically accessible 2-hydroxyketones with *S* enantiomers, and thus provides a valuable platform for chemoenzymatic synthesis.

enzyme	involved amino acid
S pocket	
*Pp*BFD	Leu461
*Ap*PDC	Glu469
*Zm*PDC	Glu473
*Sc*PDC	Glu477
*Ll*KdcA	Glu462
entrance to *S* pocket	
*Pp*BFD	Ala460
*Ap*PDC	Ile468
*Zm*PDC	Ile472
*Sc*PDC	Ile476
*Ll*KdcA	Val461

Figure 3. Superimposition of the *S* pockets of BFDwt and *Ap*PDC. In comparison with BFD, *Ap*PDC shows an enlarged *S* pocket, but the entrance is blocked by residue Ile468 (left). The amino acids mainly bordering the *S* pockets, as well as those defining the entrances to the pockets, are given for BFD and *Ap*PDC, and additionally for *Zm*PDC, *Sc*PDC, and *Ll*KdcA (right).

Experimental Section

Site-directed mutagenesis: The 1611 bp gene of benzoylformate decarboxylase (BFD, E.C. 4.1.1.7) from *Pseudomonas putida* was ligated into a pKK233-2 plasmid (Pharmacia),[3, 15] which contained the information for a C-terminal His$_6$-tag. For mini- and midipreparations, *E. coli* XL1-blue (Stratagene) was transformed with the construct by electroporation. For over-expression, *E. coli* SG13009/pRep4 (Qiagen) was used as host. Site-directed mutagenesis was performed with the aid of the QuikChange® site-directed mutagenesis kit (Stratagene). The sequences of the mutagenesis primers are given in the Supporting Information. Gene sequences were confirmed by DNA sequencing (Sequiserve).

Expression and purification: Incubation of the variants was carried out in shaking cultures (1 L LB medium, pH 7.5, 5 L flasks). Overexpression was induced by addition of IPTG (1 mM) at $OD_{600} \leq 0.45$. For biotransformations the variants were purified according to a protocol previously developed for BFDwt[3, 15] [Ni-NTA chromatography: disintegration buffer (50 mM potassium phosphate, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP), washing buffer (50 mM potassium phosphate, pH 7.0, 20 mM imidazole), elution buffer (50 mM potassium phosphate, pH 7.8, 250 mM imidazole); G25-chromatography (10 mM potassium phosphate pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP)]. After purification the enzyme variants were either freeze dried or diluted with glycerol (50%, *v/v*) and stored at −20 °C.

For enzyme crystallization of BFDL461A the same purification protocol was used but with different buffers [Ni-NTA chromatography: disintegration buffer (Mes/NaOH, 50 mM, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP), washing buffer (Mes/HCl, 50 mM, pH 7.0, 50 mM imidazole), elution buffer (Mes/HCl, 50 mM, pH 7.0, 250 mM imidazole); G25-chromatography (Mes/NaOH, 20 mM, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP)]. Concentration of enzyme solutions was performed in vivaspin 20 centrifuge columns (Sartorius, cut-off 10 kDa) up to 130 mg mL^{-1}. Superdex G200 (GE Healthcare) size-exclusion chromatography showed 97.5% purity of the tetrameric

BFDL461A (2% dimeric BFDL461A, 0.5% impurity). For storage the enzyme solution was shock frozen in liquid nitrogen and kept at −20 °C.

Decarboxylase activity assay: One unit of decarboxylase activity is defined as the amount of enzyme that catalyzes the decarboxylation of 1 μmol benzoylformate per minute under standard conditions (pH 6.5, 30 °C). Activity was measured by coupled photometric assay as previously described.[3] For determination of the substrate range, different 2-ketoacids were applied in a final concentration of 30 mM in this assay; except in the case of indole-3-pyruvate (1 mM; see the Supporting Information).

Protein concentration determination: Protein concentrations were determined as described by Bradford[16] with bovine serum albumin (BSA) as standard.

Benzoin syntheses

Reaction conditions: Benzaldehyde (20 mM), DMSO (20 vol%), BFD variant (0.3 mg mL^{-1}), potassium phosphate buffer (50 mM, pH 7.5, 2.5 mM MgSO$_4$, and 0.1 mM ThDP), were incubated at 30 °C and 100 rpm. To avoid evaporation of the aldehydes the reaction batch was divided into GC vials, each with a volume of 400 μL, after the starting sample had been taken. The reaction was stopped by addition of acetonitrile (400 μL) followed by intense vortexing and centrifugation of the precipitate. Calibration curves with benzoin were prepared in the same way. Conversions were determined by HPLC, with use of a Dionex HPLC instrument (Germering) equipped with a 250×4.6 Multohyp ODS-5 μ (CS-Chromatography) and a UV detector (mobile phase 60% (*v/v*) H$_2$O: 40% (*v/v*) acetonitrile, flow 1.1 mL min^{-1}, pressure 130 bar, 20 μL injection volume, detection $\lambda = 250$ nm), t_R (benzoin) = 32.2 min.

Acetoin syntheses

Reaction conditions: Acetaldehyde (40 mM), DMSO (20 vol%), BFD variant (0.3 mg mL^{-1}), potassium phosphate buffer (50 mM, pH 7.5, 2.5 mM MgSO$_4$, 0.1 mM ThDP) were incubated at 30 °C and 100 rpm. As described for the benzoin synthesis, the reaction batch was divided into GC vials each with a volume of 400 μL. For enzyme inactivation the vial was heated for 60 s at 90 °C followed by centrifugation of the precipitate. Conversion and enantiomeric excess were determined by chiral GC by using 6890 N Agilent GC (Palo Alto) equipped with a Cyclodex b-1/P column (50 m × 320 μm) and a FID detector (flow 3.4 mL min^{-1}, pressure 0.8 bar, split 5:1, 1 μL injection volume, temperature gradient: 50 °C for 5 min, 40 °C min^{-1} to 190 °C), t_R (*R*)-acetoin = 6.98 min, t_R (*S*)-acetoin = 7.11 min.

Mixed carboligations of benzaldehyde and different aliphatic aldehydes

Analytical scale

Reaction conditions (1.5 mL scale): Benzaldehyde (0.027 mmol, 2.9 mg) was dissolved in a mixture of DMSO (0.3 mL) and potassium phosphate buffer (50 mM, 1.2 mL, pH 7, 2.5 mM MgSO$_4$, 0.1 mM ThDP). Acetaldehyde, propanal, or butanal (0.027 mmol) was added to this solution. After addition of purified enzyme (0.45 mg) the reaction mixture was stirred slowly at 30 °C for 72 h. The reaction mixture was extracted with CDCl$_3$.

Preparative scale synthesis

Reaction conditions (15 mL scale): Benzaldehyde (29 mg, 0.27 mmol) and propanal (16 mg, 0.27 mmol) were dissolved in DMSO (3 mL). After addition of potassium phosphate buffer (50 mM, 12 mL, pH 7.9, 2.5 mM MgSO$_4$, 0.1 mM ThDP) the reaction was started with

the purified BFD variant (4.5 mg) and the mixture was stirred slowly at 30 °C. After 26.5 h, further BFDwt or variants (4.5 mg) were added. The reaction was stopped either after 70 h (BFDwt, BFDL461A) or 50 h (BFDL461G) by extracting three times with ethyl acetate (25 mL), and the organic layer was dried over Na_2SO_4. The solvent was evaporated, and the crude product was dissolved in ether (5 mL). The ether extract was washed with brine and dried over Na_2SO_4, followed by evaporation of the solvent.

Analysis of (S)-2-hydroxy-1-phenylbutan-1-one ((S)-7): HPLC: (chiral OD-H, n-hexane/propan-2-ol, 95:5, 0.5 mL min^{-1}, 40 °C): t_R (S) = 11.8 min, t_R (R) = 13.9 min; $[\alpha]_D^{21.4}$ = −11.52 (c = 0.4, CHCl$_3$); CD (acetonitrile): λ (mol CD) = 298 (−0.3250), 281 (−1.7045), 239 (6.2025), 206 nm (−7.6004); ^{1}H NMR (400 MHz, CDCl$_3$, 300 K): δ = 0.96 (t, $^3J_{H,H}$ = 7.4 Hz, 3H; CH$_3$), 1.63 (dqd, $^2J_{H,H}$ = 14.3 Hz, $^3J_{H,H}$ = 7.4 Hz, 7.3 Hz, 1H; CH$_2$), 1.98 (dqd, $^2J_{H,H}$ = 14.3 Hz, $^3J_{H,H}$ = 7.4 Hz, 3.8 Hz, 1H; CH$_2$), 3.73 (d, $^3J_{H,H}$ = 6.4 Hz, 1H; OH), 5.08 (ddd, $^3J_{H,H}$ = 7.3 Hz, 6.4 Hz, 3.8 Hz, 1H, CHOH), 7.52 (ddm, $^3J_{H,H}$ = 7.4 Hz, 2H; Ar-H), 7.64 (ddm, $^3J_{H,H}$ = 7.4 Hz, 1H; Ar-H), 7.93 ppm (dm, $^3J_{H,H}$ = 7.4 Hz, 2H; Ar-H); ^{13}C NMR (100 MHz, CDCl$_3$, 300 K): δ = 8.8 (CH$_3$), 28.8 (CH$_2$), 73.9 (CHOH), 128.4 (2 CH$_{Ar}$), 128.8 (2 CH$_{Ar}$), 133.3 (CH$_{Ar}$), 202.1 ppm (CO); GCMS t_R = 8.7 min; MS (70 eV, EI): m/z (%): 164 (0.1 %) [M]$^+$, 105 (100 %), 77 (46 %).

Crystallization: BFDL461A was crystallized by the hanging-drop vapor diffusion method. Droplets were set up for crystallization by mixing of protein solution [2 µL; containing protein (13 g mL^{-1}) in Mes/NaOH (20 mM, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP)] and the reservoir solution (2 µL). Screening and optimization revealed a reservoir solution consisting of PEG 2000 MME (18–24 %, w/v), sodium citrate (pH 5.2–5.8, 0.1 M), and (NH$_4$)$_2$SO$_4$ (100–150 mM) to be optimal. After equilibration for 3 days, diffraction-quality crystals were obtained.

Data collection and processing: For cryoprotection the crystals were quickly dipped into the well solution supplemented with ethylene glycol (25 %) before being frozen in a cryogenic nitrogen gas stream at 110 K. Data were collected to a resolution of 2.2 Å at beamline I911-3 (Max-lab, Lund, Sweden). Images were processed with the aid of MOSFLM,[17] and the unit cell parameters were determined by the autoindexing option. The data set was scaled with the program SCALA implemented in the CCP4 program suite.[18] The crystal belongs to the space group P2$_1$2$_1$2$_1$ with the cell dimensions a = 96 Å, b = 140 Å, and c = 169 Å. Four monomers were packed in one asymmetric unit. Data collection statistics are given in Table 5.

Structure solution and crystallographic refinement: The structure of BFDL461A was determined by molecular replacement by use of the program MOLREP.[18, 19] The BFDwt structure (PDB ID code: 1bfd)[5] was used as search model to place the four monomers into the asymmetric unit. Atomic positions and B factors of the model were refined by the maximum likelihood method in REFMAC5,[18, 20]

which was interspersed with rounds of manual model building in COOT.[21] Noncrystallographic symmetry restrains were initially applied but were released toward the end of refinement. Water assignment was performed in COOT, in which the model was also validated. The quality of the final structure was examined with PROCHECK.[18, 22] Statistics of the refinement and final model are given in the Supporting Information. The coordinates of BFDL461A, in addition to the structure factors, have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Databank PDB with the ID code 2v3w.

Structural analysis and substrate placement: The modeling studies were carried out with the programs PyMol[23] and Swiss-Pdb Viewer.[24] To investigate the differences in the stereoselectivities of 2-hydroxyketone formation by BFDwt and BFDL461A, and to predict the optimal substrate size of the acceptor aldehyde fitting in the S pocket, the acceptor and donor aldehydes were placed into the active sites as described previously.[6] Models of the molecules were created with the molecular builder in the program SYBYL (Tripos, St. Louis, MO, USA).

Acknowledgements

[1] M. Pohl, G. A. Sprenger, M. Müller, *Curr. Opin. Biotechnol.* **2004**, *15*, 335–342.
[2] A. S. Demir, M. Pohl, E. Janzen, M. Müller, *J. Chem. Soc. Perkin Trans. 1* **2001**, 633–635.
[3] a) H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir, M. Pohl, *Chem. Eur. J.* **2000**, *6*, 1483–1495; b) B. Lingen, D. Kolter-Jung, P. Dünkelmann, R. Feldmann, J. Grötzinger, M. Pohl, M. Müller, *ChemBioChem* **2003**, *4*, 721–726.
[4] M. Knoll, M. Müller, J. Pleiss, M. Pohl, *ChemBioChem* **2006**, *7*, 1928–1934.
[5] a) M. S. Hasson, A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A. Gerlt, G. L. Kenyon, G. A. Petsko, D. Ringe, *Biochemistry* **1998**, *37*, 9918–9930; b) E. S. Polovnikova, M. J. McLeish, E. A. Sergienko, J. T. Burgner, N. L. Anderson, A. K. Bera, F. Jordan, G. L. Kenyon, M. S. Hasson, *Biochemistry* **2003**, *42*, 1820–1830.
[6] a) K. C. Nicolaou, Z. Yang, J. J. Liu, H. Ueno, P. G. Nantermet, R. K. Guy, C. F. Claiborne, J. Renaud, E. A. Couladouros, K. Paulvannan, E. J. Sorensen, *Nature* **1994**, *367*, 630–634; b) F. R. Stermitz, P. Lorenz, J. N. Tawara, L. A. Zenewicz, K. Lewis, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1433–1437.
[7] P. Domínguez de María, M. Pohl, D. Gocke, H. Gröger, H. Trauthwein, T. Stillger, L. Walter, M. Müller, *Eur. J. Org. Chem.* **2007**, 2940–2944.
[8] T. G. Mosbacher, M. Müller, G. E. Schulz, *FEBS J.* **2005**, *272*, 6067–6076.
[9] a) D. Dobritzsch, S. König, G. Schneider, G. Lu, *J. Biol. Chem.* **1998**, *273*, 20196–20204; b) P. Arjunan, T. Umland, F. Dyda, S. Swaminathan, W. Furey, M. Sax, B. Farrenkopf, Y. Gao, D. Zhang, F. Jordan, *J. Mol. Biol.* **1996**, *256*, 590–600.
[10] D. Gocke, C. Berthold, T. Graf, H. Brosi, I. Frindi-Wosch, M. Knoll, T. Stillger, L. Walter, M. Müller, J. Pleiss, G. Schneider, M. Pohl, unpublished results.
[11] A. Yep, G. L. Kenyon, M. J. McLeish, *Bioorg. Chem.* **2006**, *34*, 325–336.
[12] D. Gocke, C. L. Nguyen, M. Pohl, T. Stillger, L. Walter, M. Müller, *Adv. Synth. Catal.* **2007**, *349*, 1425–1435.
[13] C. Berthold, D. Gocke, M. D. Wood, F. Leeper, M. Pohl, G. Schneider, *Acta Crystallogr. Sec. D Biol. Crystallogr.* **2007**, *63*, 1217–1224.

Table 5. Data collection statistics for BFDL461A. Values in parentheses are given for the highest resolution interval.

resolution [Å]	2.2 (2.32–2.2)
no. of observations	321349 (46653)
no. of unique reflections	111569 (16561)
completeness [%]	97.0 (99.2)
multiplicity [%]	2.9 (2.8)
mean {I/σ(I)}	8.8 (2.2)
Wilson B factor [Å^2]	27.8
R_{merge} [%]	12.4 (45.8)

[14] P. Siegert, M. J. McLeish, M. Baumann, H. Iding, M. M. Kneen, G. L. Kenyon, M. Pohl, *Protein Eng. Des. Sel.* **2005**, *18*, 345–357.

[15] A. Y. Tsou, S. C. Ransom, J. A. Gerlt, D. D. Buechter, P. C. Babbitt, G. L. Kenyon, *Biochemistry* **1990**, *29*, 9856–9862.

[16] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248–254.

[17] A. G. W. Leslie, Joint CCP4–ESF–EAMCB *Newsl. Protein Crystallogr.* **1992**, 26.

[18] Collaborative Computational Project, Number 4, *Acta Crystallogr. D Biol. Crystallogr.* **1994**, *50*, 760–763.

[19] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.

[20] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sec. D Biol. Crystallogr.* **1997**, *53*, 240–255.

[21] P. Emsley, K. Cowtan, *Acta Crystallogr. Sec. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

[22] R. A. Laskowski, M. W. McArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

[23] W. L. DeLano, The PyMOL Molecular Graphics System, http://www.pymol.org, San Carlo, CA, USA, **2002**.

[24] N. Guex, C. Peitsch, *Electrophoresis* **1997**, *18*, 2714–2723; http://www.expasy.org/spdbv/.