Tailor-made thiamine diphosphate-dependent enzymes for *S*-selective carboligation

Complementation of the α-hydroxy ketone platform

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Robert Westphal
aus Waren (Müritz)

Aachen, Oktober 2013

aus dem Institut für Bio- und Geowissenschaften (IBG-1: Biotechnologie)
des Forschungszentrums Jülich GmbH

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Martina Pohl

Korreferent: Prof. Dr. Vlada Urlacher

Tag der mündlichen Prüfung: 28.01.2014

SELBSTSTÄNDIGKEITSERKLÄRUNG

Hiermit versichere ich an Eides statt, dass die vorgelegte Dissertation „Tailor-made thiamine diphosphate-dependent enzymes for *S*-selective carboligation“ von mir selbstständig verfasst und unter ausschließlicher Verwendung der angegebenen Literatur und Hilfsmittel gemäß der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt wurde.
Bisher habe ich keine erfolglosen Promotionsversuche unternommen.

_____ _____

Ort, Datum Robert Westphal

ABSTRACT

Thiamine diphosphate (ThDP)-dependent enzymes enable the challenging stereoselective synthesis of α-hydroxy ketones via asymmetric C–C bond formation. However, the steric and chemical properties of the enzymes' active sites often limit the product range, specifically the access to (*S*)-α-hydroxy ketones. Introduction of the *S*-pocket concept, which explains stereoselectivity by the orientation of the substrates prior to carboligation, paved the way for the design of *S*-selective ThDP-dependent enzymes, however, often only with moderate stereoselectivity.

This thesis aimed for new concepts to design tailor-made *S*-selective variants to expand the toolbox of ThDP-dependent enzymes. After confirmation of the general validity of the *S*-pocket concept by transfer to the new toolbox enzymes acetohydroxyacid synthase and 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD), two novel strategies were developed to tailor ThDP-dependent enzymes with high *S*-selectivity: (i) suppression of the "*R*-pathway", and (ii) creation of "thiamine enzyme hybrids". The combination of the basic *S*-pocket concept with targeted destabilization of the *R*-pathway improved the *S*-selectivity of pyruvate decarboxylase (PDC) as well as new MenD variants enabling carboligations with up to > 99 % *ee*. The design of a hybrid substrate-binding site from an *S*-selective PDC variant with a large *S*-pocket and a benzaldehyde lyase with a large donor binding site combined the different selectivity-determining modules and thus solved the long-standing problem of *S*-selective benzoin synthesis starting from readily available benzaldehydes. The results demonstrate the high quality of current structure-function relationships as well as the robustness of ThDP-dependent enzymes towards active site mutations. *S*-Selectivity can be adjusted now by different strategies. Moreover, the hybridization approach might pave the way for a tailor-made design of ThDP-dependent enzymes with desired and novel carboligation activities.

This work complemented the toolbox of ThDP-dependent enzymes by four wild-type enzymes and about 35 variants with characterized carboligation activity, which broadened the enzymatically accessible α-hydroxy ketone platform by novel functionalized mixed araliphatic α-hydroxy ketones as well as (*S*)-benzoins with excellent *ee*s.

KURZFASSUNG

Thiamindiphosphat (ThDP)-abhängige Enzyme ermöglichen die anspruchsvolle stereoselektive Synthese von α-Hydroxyketonen durch asymmetrische C–C Knüpfungsreaktionen. Die sterischen und chemischen Eigenschaften der aktiven Zentren begrenzen jedoch häufig den Zugang zu den Produkten, speziell zu (*S*)-α-Hydroxyketonen. Die Einführung des *S*-Taschen Konzepts, das die Stereoselektivität durch die Anordnung der Substrate vor der Carboligation erklärt, ebnete den Weg für das Design *S*-selektiver ThDP-abhängiger Enzyme. Diese zeigen allerdings oftmals nur moderate Stereoselektivität.

Die Doktorarbeit zielte auf die Entwicklung neuer Konzepte zum Design maßgeschneiderter *S*-selektiver Enzymvarianten zur Erweiterung der Toolbox ThDP-abhängiger Enzyme. Zunächst wurde die generelle Gültigkeit des *S*-Taschen Konzepts durch den Transfer auf die neuen Toolbox-Enzyme Acetohydroxysäuresynthase und 2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadien-1-carboxylat Synthase (MenD) bestätigt. Anschließend wurden zwei neue Strategien für das Design hoch *S*-selektiver ThDP-abhängiger Enzyme entwickelt: (i) Unterdrückung des „*R*-Wegs“ und (ii) Erzeugung von „Thiamin-Enzymhybriden“. Die Kombination des *S*-Taschen Konzepts mit der gezielten Destabilisierung des *R*-Wegs verbesserte die *S*-Selektivität der Pyruvatdecarboxylase (PDC) und neuer MenD Varianten, wodurch Carboligationen mit bis zu > 99 % *ee* ermöglicht wurden. Die Kombination unterschiedlicher Selektivität-bestimmender Module einer *S*-selektiven PDC und einer Benzaldehydlyase löste das seit langem bestehende Problem der *S*-selektiven Benzoinsynthese ausgehend von leicht verfügbaren Benzaldehyden. Die Ergebnisse belegen die hohe Qualität gegenwärtiger Struktur-Funktionsbeziehungen sowie die Robustheit ThDP-abhängiger Enzyme gegenüber Mutationen im aktiven Zentrum. *S*-Selektivität kann nun mithilfe verschiedener Strategien reguliert werden. Darüber hinaus könnte die Hybrid-Strategie den Weg zu maßgeschneiderten ThDP-abhängigen Enzymen mit neuen Carboligationsaktivitäten ebnen.

In dieser Arbeit wurde die Toolbox ThDP-abhängiger Enzyme um vier Wildtyp-Enzyme und etwa 35 charakterisierten Enzymvarianten mit Carboligationsaktivität vergrößert, wodurch die enzymatisch zugängliche α-Hydroxyketon-Plattform um neue funktionalisierte α-Hydroxyketone sowie (*S*)-Benzoine mit exzellenten *ee*s erweitert werden konnte.

LIST OF PUBLICATIONS

1. Sehl, T.; Hailes, H. C.; Ward, J. M.; Wardenga, R.; von Lieres, E.; Offermann, H.; Westphal, R.; Pohl, M. & Rother, D. (2013). Two Steps in One Pot: Enzyme Cascade for the Synthesis of Nor(pseudo)ephedrine from Inexpensive Starting Materials. *Angew. Chem. Int. Ed.*, 52, 6772–6775.

2. Westphal, R.; Waltzer, S.; Mackfeld, U.; Widmann, M.; Pleiss, J.; Beigi, M.; Müller, M.; Rother, D. & Pohl, M. (2013). (*S*)-Selective MenD variants from *Escherichia coli* provide access to new functionalized chiral α-hydroxy ketones. *Chem. Commun.*, 49, 2061–2063.

3. Westphal, R.; Hahn, D.; Mackfeld, U.; Waltzer, S.; Beigi, M.; Widmann, M.; Vogel, C.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2013). Tailoring the *S*-selectivity of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) from *Escherichia coli*. *ChemCatChem*, 5, 3587–3594.

4. Westphal, R.; Jansen, S.; Vogel, C.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2013). MenD from *Bacillus subtilis*: A Potent Catalyst for the Enantiocomplementary Asymmetric Synthesis of Functionalized α-Hydroxy Ketones. *ChemCatChem*, DOI: 10.1002/cctc.201300690.

5. Westphal, R.; Vogel, C.; Schmitz, C.; Pleiss, J.; Müller, M.; Pohl, M. & Rother, D. (2014). Tailor-made Pyruvate Decarboxylase for the Direct Asymmetric Synthesis of (*S*)-Benzoins. *to be submitted*

6. Hailes, H. C.; Rother, D.; Müller, M.; Westphal, R.; Ward, J. M.; Pleiss, J.; Vogel, C. & Pohl, M. (2013). Engineering Stereoselectivity of ThDP-dependent Enzymes. *FEBS J.*, 280, 6374–6394.

7. Sehl, T.; Kulig, J.; Westphal, R. & Rother, D. (2013). Synthetic enzyme cascades for valuable diols and amino alcohols – smart composition and optimisation strategies. In *Industrial Biocatalysis* (Ed.: Grunwald, P.), Pan Stanford Publishing, Singapore, *accepted for publication*.

LIST OF ORAL PRESENTATIONS

1. Westphal, R. (*presenting author*); Waltzer, S.; Widmann, M.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2012). Transfer of the *S*-pocket concept: Engineering stereoselectivity by structure-based design of MenD from *Escherichia coli*. *Trends in Enzymology*, Göttingen.

2. Westphal, R. (*presenting author*); Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2012) Integrated engineering of ThDP-dependent enzymes: Access to new chiral (*S*)-α-hydroxy ketones. *Zing Conference on Biocatalysis*, Xcaret (Mexico).

3. Sehl, T.; Kulig, J.; Westphal, R.; Baraibar, A.; Jakoblinnert, A.; Wachtmeister, J.; Pohl, M. & Rother, D. (*presenting author*) (2013). Modular synthetic enzyme cascades for the production of chiral diols and amino alcohols. *9th European Congress of Chemical Engineering and 2nd European Congress of Applied Biotechnology*, The Hague (The Netherlands).

4. Pohl, M. (*presenting author*); Westphal, R.; Rother, D.; Müller, M.; Pleiss, J. (2013). Tailoring stereoselectivity of ThDP-dependent enzymes. *11th Biotrans*, Manchester (England).

LIST OF POSTER PRESENTATIONS

1. Westphal, R.; Waltzer, S.; Widmann, M.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2011). Stereochemical promiscuity of thiamine enzymes: Engineering the stereoselectivity of MenD from *Escherichia coli*. *2nd DECHEMA Summer School Biotransformations*. Bad Herrenalb.

2. Westphal, R.; Waltzer, S.; Widmann, M.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2011). Stereochemical promiscuity of thiamine enzymes: Engineering the stereoselectivity of MenD from *Escherichia coli*. *10th Biotrans*. Giardini Naxos (Italy).

3. Westphal, R.; Waltzer, S.; Widmann, M.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2012).Transfer of the *S*-pocket concept: Engineering stereoselectivity by structure-based design of MenD from *Escherichia coli*. *Trends in Enzymology*, Göttingen.

4. Westphal, R.; Waltzer, S.; Widmann, M.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M. (2011). Engineering stereoselectivity of MenD from *Escherichia coli*: Access to new functionalized chiral (*S*)-α-hydroxy ketones. *6th International Congress on Biocatalysis*, Hamburg-Harburg.

POSTER AWARDS

1. Biocat Poster Award 2012. *6th International Congress on Biocatalysis*, Hamburg.

2. Trends in Enzymology Poster Award 2012. *Trends in Enzymology*, Göttingen.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all people who made this thesis possible, who became part of my PhD project, and those people who have accompanied me through the time of my doctorate.

I would like to express my sincere gratitude and deepest appreciation to my supervisors Prof. Dr. Martina Pohl and Dr. Dörte Rother for the opportunity to work on a great and challenging project, for all of the many inspiring scientific discussions, their helpful suggestions, and the perfect working conditions. Especially, I would like to thank you for your trust in my person and work as well as for the freedoms I had to try all my (crazy) ideas.

Furthermore, I would like to thank Prof. Dr. Vlada Urlacher from the Institute of Biochemistry (Heinrich-Heine-University Düsseldorf) for the kind effort to be co-referee of this thesis.

I would like to thank Prof. Dr. Wolfgang Wiechert for the opportunity to perform my doctoral thesis at the IBG-1 in an inspiring and pleasant environment. In this context, I would also like to thank Marianne Hess, Maria Lauer, and Marion Dreßen-Combach who made my life a lot easier in the “German Jungle of bureaucracy”.

Sincere thanks go to all people of the DFG Research unit 1296 “Diversity of Asymmetric Thiamine Catalysis”, especially Prof. Dr. Michael Müller as the speaker of this research group, for the inspiring, fruitful, constructive, but very familiar meetings. Furthermore, I would like to thank Alexander Fries, Sabrina Loschonsky, Maryam Beigi, Shiromi Beier, Cindy Wechsler, Michael Widmann, Thomas Broja, Kai Mahnken, Simon Waltzer, and Constantin Vogel for wonderful PhD meetings, fruitful discussions, and constructive collaborations.

I would also like to thank the graduate school “BioNoCo” (GK1166) that enabled me to visit interesting conferences and wonderful places all over the world, and to take part in several inspiring workshops. Especially, I would like to thank Zaira Maugeri (ITMC, RWTH Aachen) for the nice collaboration regarding deep eutectic solvents.

A successful project always lives from scientific discussion and cooperation. In this context, I would like to express my sincere acknowledgements to Prof. Dr. Michael Müller from the

Institute of Pharmaceutical Science (Albert-Ludwigs-University Freiburg) and his technician Simon Waltzer, as well as to Prof. Dr. Jürgen Pleiss from the Institute of Technical Biochemistry (University of Stuttgart) and his Postdoc Michael Widmann and doctoral student Constantin Vogel. Special thanks go to Simon and his effort and skillful assistance with respect to the analysis of various MenD-catalyzed carboligations. I would also like to express my sincere gratitude to my friend Constantin for his effort, help, ideas, and motivation during the last two years.

Furthermore, I would like to thank further cooperation partners: Dr. Jan Marienhagen (IBG-1) for his help with respect to the generation of mutant libraries; Dr. Marco Bocola (Computational Biology, RWTH Aachen) for all of the fruitful discussions regarding proteins and his help to get an insight into the complex field of protein modeling; Dr. Doreen Dobritzsch (Department of Chemistry, Uppsala University, Sweden) for her (still) intensive efforts to crystalize *Ec*MenD; and Dr. Friedrich Käß (IBG-1) for his skillful technical assistance concerning the BioLector® system.

This work would not have been possible without the help of several people. In this context, I would like to thank my two bachelor students Carlo Schmitz and Sascha Jansen for their high motivation and their tremendous effort in the lab. Thank you for the excellent work in the characterization of *Ec*AHAS, *Bs*MenD, and *Ap*PDC. Furthermore, special thanks go to Doris Hahn, Ursula Mackfeld, Lilia Arnold, and Heike Offermann for their helping hands, their high motivation, and brilliant work in many different projects. Especially, I would like to thank Doris and Ursula for their effort in the development of a robot-assisted screening setup as well as their patience in the preparation, screening, and characterization of a variety of new variants. Thank you Heike, for the purification of various *S*-benzoins.

Additionally, I would like to thank all of my (former) colleagues of the Biocatalysis and Biosensors group for this great working atmosphere, the support in the lab, the fruitful discussion, and of course the fun we had during the last three years.

Finally, I would like to appreciate the support of my family and especially the understanding my wife Anke has given me during my work, particularly during all the nights and weekends of work.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviations for enzymes

*Ap*PDC	Pyruvate decarboxylase from *Acetobacter pasteurianus*
*Bs*MenD	2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase from *Bacillus subtilis*
CDH	Cyclohexane-1,2-dione hydrolase from *Azoarcus* sp. 22Lin
*Ec*AHAS	Acetohydroxyacid synthase from *Escherichia coli*
*Ec*MenD	2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase from *Escherichia coli*
*Ll*KdcA	Branched-chain keto acid decarboxylase from *Lactococcus lactis*
*Pf*BAL	Benzaldehyde lyase from *Pseudomonas fluorescens*
*Pp*BFD	Benzoylformate decarboxylase from *Pseudomonas putida*
SucA	E1 component of the α-ketoglutarate dehydrogenase subunit from *Escherichia coli* K12

Further abbreviations

2-HPP	2-Hydroxypropiophenone
α-KG	α-Ketoglutarate
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
CD	Circular dichroism
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EC	Enzyme commission number
ee	Enantiomeric excess
GC	Gas chromatography
HPLC	High performance liquid chromatography

IPTG	Isopropyl-β-D-thiogalactopyranoside
LB medium	Lysogeny broth medium
MD	Molecular dynamics
MTBE	Methyl *tert*-butyl ether
NADH	Nicotinamide adenine dinucleotide
NMR	Nuclear magnetic resonance
PAC	Phenylacetylcarbinol
PCR	Polymerase chain reaction
pdb	Protein Data Bank
PPC	Phenylpropionylcarbinol
PP-domain	Pyrophosphate domain
PYR-domain	Pyrimidine domain
SEPHCHC	2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate
SIBC	Swiss Industrial Biocatalysis Consortium
ThDP	Thiamine diphosphate
TEED	Thiamine diphosphate-dependent Enzyme Engineering Database
TTC	Triphenyl tetrazolium chloride
U	Units

1 INTRODUCTION

1.1 Organic compounds – The carbon-based life

Carbon represents the chemical basis of all earthly life. Owing to its four valence electrons, carbon is able to form bonds with all other vital elements such as oxygen, hydrogen, nitrogen, phosphor, and sulfur. These bonds can be formed, broken, and rearranged at temperatures commonly encountered on earth. Moreover, carbon readily recombines with other carbon atoms representing unusual polymer-forming characteristics. These exceptional abilities and the abundance of carbon on earth are the basis for the formation of a tremendous variety of structurally different macromolecules, so-called organic compounds. Free of doubt, the most important groups of organic macromolecules in living organisms include nucleic acids, proteins, carbohydrates, and lipids.

In addition to the relevance for living organisms, organic compounds possess an outstanding role for the human life, or rather the human "carbon-based" life. Although current social and ecological developments have to be critically scrutinized, organic compounds essentially contribute to the wealth, health as well as technical progress of the present society. Thereby, their range of application is enormous, encompassing petrochemicals, pharmaceuticals, agrochemicals, and polymers (plastics), among others. In this context, the synthesis of organic compounds has become one of the most important branches of organic chemistry.

1.1.1 Organic synthesis

Synthetic organic chemistry, or organic synthesis, is a special discipline of synthetic chemistry, which addresses the construction of organic compounds using chemical reactions. In 1828, Friedrich Wöhler described the first targeted synthesis of the natural organic product urea (Figure 1), starting from inorganic ammonium cyanate (Wöhler, 1828). This event is regarded as the beginning of organic synthesis (Nicolaou *et al.*, 2000). Since then, this area of chemistry steadily evolved. One main focus was directed towards the design of complex organic molecules, starting from simple and commercially available precursors. Probably one of the most famous examples of organic syntheses is the total synthesis of the alkaloid strychnine (Figure 1) by the father of modern organic synthesis Robert B. Woodward in 1954 (Woodward *et al*., 1954; Nicolaou *et al*., 2000). His achievements in organic synthesis were honored with the Nobel Prize for Chemistry in 1965. A vast variety of other impressive

examples highlights the progress in synthetic organic chemistry, such as the total synthesis of the anti-cancer drug taxol (Figure 1) (Holton *et al*., 1994a; Holton *et al*., 1994b; Nicolaou *et al*., 1994).

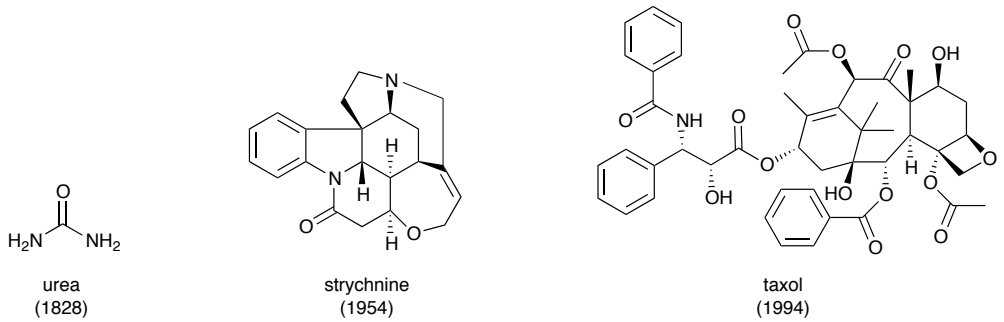


Figure 1: Selected landmarks in synthetic organic chemistry.

The impact of organic synthesis on other scientific disciplines, like biology and medicine, and on society has been undoubtedly enormous (Nicolaou *et al.*, 2000). From an industrial point of view, synthetic organic chemistry constitutes the key technology for the production of fine chemicals.

1.2 Fine chemicals

There is no precise definition of the term "fine chemical". The general characteristics of fine chemicals are (Pollak, 2011):

- Single, pure, often complex and poly-functionalized chemical substances
- Production via complex (multi-step) synthesis reactions
- High degree of purity (exact specifications)
- Limited production volumes at relatively high prices (low volume / high added value).

The range of application of fine chemicals is versatile. *Inter alia*, they are used in research for chemical syntheses or analyses, as food additives (*e.g.* vitamins), or ingredients in cosmetics and washing powder. Moreover, they find use as starting material for the production of specialty chemicals, particularly pharmaceuticals and agrochemicals (Pollak, 2011).
Figure 2 A demonstrates the importance of fine chemicals for the chemical industry exemplarily for Germany. Fine and specialty chemicals, and pharmaceuticals contributed together about €80 billion (43 %) to the total turnover of €184 billion of the German chemical industry in 2011 (Source: VCI, 2012).

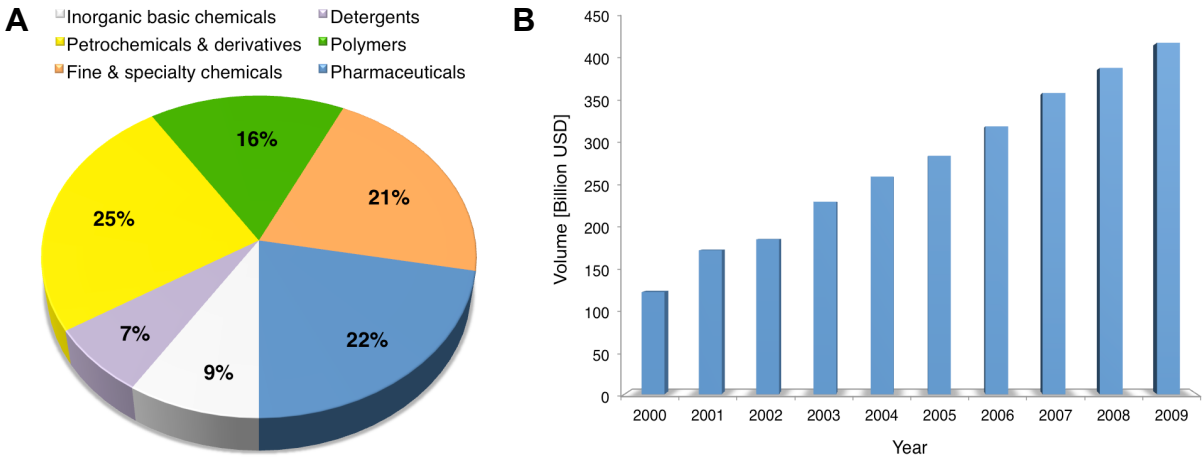


Figure 2: Market segmentation of the German chemical industry in 2011 (**A**) (Source: VCI, 2012) and the worldwide sales of chiral pharmaceuticals during the period of 2000–2009 (**B**) (Chir TU-Berlin, 2011).

Owing to their biological activity and their benefit for the society, agrochemicals and pharmaceuticals play major roles in the fine chemistry sector. Especially the pharmaceutical industry continuously grew with an increased turnover of 36 % since 1999 (Source: VCI, 2012), accompanied by new developments in drug discovery, synthetic organic chemistry, as well as biotechnology (chapter 1.4). The demands on pharmaceuticals are very high and still challenging. Thereby, the major concern of the pharmaceutical industry is stereoselectivity.

1.2.1 Stereochemically pure pharmaceuticals

Chirality is tightly connected with the carbon-based life because it is most often found in the presence of an asymmetric carbon atom. The most prominent examples of chiral molecules in living organisms are amino acids and sugars. Owing to their inherently chiral character, living systems often react (biochemically) differently with respect to the various stereoisomers of a given compound. There are numerous examples in the case of pharmaceutically active molecules, where one stereoisomer (also called eutomer) is biologically active or shows the desired effect, whereas the other stereoisomer (also called distomer) is either inactive or exhibits negative (toxic) effects (Crossley, 1992; Smith, 2009; Faber, 2011; Mitra & Chopra, 2011). The experience and the knowledge about dangerous effects on health of certain stereoisomers led to the development of more structured drug regulations and controls. In 1992, the US Food and Drugs Administration passed a new law that obliged drug producers to test both single isomers as well as the racemic mixture of all racemic drugs before they enter the market (Smith, 2009; Faber, 2011). At this point, the pharmaceutical (chemical) industry has undergone a shift in paradigm from the production of racemic drugs to the production of new chiral products in optically pure form as a consequence of the costly and time consuming

test procedure; racemates virtually disappeared (Faber & Patel, 2000; Agranat *et al*., 2002; Smith, 2009). As a result, stereochemistry (and stereoselective synthesis) experienced a real renaissance in the pharmaceutical industry and the commercial importance of stereochemically pure fine chemicals steadily increased (Figure 2 B).

1.3 Production of stereochemically pure fine chemicals

Most complex natural products and pharmaceutical drugs with biological activity are chiral. There are principally three possible routes for stereoselective synthesis in organic synthesis, each one afflicted with different drawbacks (Faber, 2011):

- **Separation of racemates:** total synthesis of racemic mixtures, which are then separated into single stereoisomers
  Drawback: usually the separation of racemates is limited to 50 % yield
- **Chiral pool:** isolation of stereochemically pure natural compounds, which provides the desired chiral center and might be used for further catalytic steps
  Drawback: only a limited number of compounds is available at an affordable prize
- **Asymmetric synthesis:** formation of only one stereoisomer starting with prochiral compounds
  Drawback: often limited stereoselectivity.

From an economical and ecological point of view, asymmetric synthesis is superior to resolution techniques and isolation of natural compounds, because chiral compounds can be formed stereoselectively from usually cheap prochiral (achiral) precursors with a theoretical yield of > 99 % (Gawley & Aubé, 2012). Therefore, asymmetric synthesis evolved to a key technology in modern organic synthesis and many excellent examples of asymmetric synthesis reactions are known so far, such as the Sharpless asymmetric dihydroxylation, which describes the chemical reaction of an alkene with osmium tetroxide in the presence of chiral quinine ligands to form vicinal diols (Jacobsen *et al.*, 1988). K. Barry Sharpless was awarded for this reaction a share of the Nobel Prize in Chemistry in 2001. Nonetheless, the use of classical organic syntheses in the fine chemistry sector often implies several limitations and drawbacks particularly with respect to ecological issues (*e.g.* toxicity of the catalyst) and also relevant reaction parameters (*e.g.* selectivity). Therefore, biotechnology became an attractive, promising, and fast growing alternative for the production of fine chemicals, particularly in the pharmaceutical industry (Liese *et al.*, 2006), which is going to be discussed in the following chapter.

1.4 Enzyme catalysis in organic synthesis

1.4.1 Industrial biotechnology

Industrial biotechnology, in Europe mainly known as white biotechnology, is the application of biological systems for industrial purposes, such as the production of fine chemicals. In recent years, industrial biotechnology became a strongly developing cutting-edge technology with highest potential to address current important issues of society, such as environmental pollution or dependency on fossil fuels (Sijbesma & Schepens, 2003). In this context, the vision of a bio-based economy arose, accompanied by the concept of *Green Chemistry*, which refers to the prevention of waste production, the use of renewable materials, toxicity of the products, and safety, among others (Anastas & Warner, 1998; Tang *et al*., 2005).

1.4.2 Enzymes in biocatalysis

The attractiveness of biotechnological processes with respect to a sustainable future is due to the application and potential of natural catalysts: enzymes or whole cell biocatalysts. The use of biocatalysis in chemical transformations on organic compounds offers a reasonable and sustainable alternative to classical organic synthesis. The well known advantages of biocatalysts are briefly summarized in the following (Patel *et al*., 2003; Faber, 2011; Gröger & Asano, 2012):

- **Versatility:** Nature provides a vast variety of versatile enzymes that are able to carry out almost any type of organic reaction. There are even enzyme-mediated reactions, which cannot be accomplished by the means of organic chemistry so far (*e.g.* selective hydroxylation of non-activated C–H bonds).
- **Mild reaction conditions:** Enzymes usually work under mild pH (5–8), ambient temperatures (20–40 °C), and atmospheric pressure. Harsh conditions often used in chemical reactions, like high pressures and temperatures, can be avoided, which minimize side reactions like isomerization, or racemization.
- **Efficacy and waste reduction:** Enzymatic reactions can proceed up to 10^{19}-fold faster than non-catalyzed reactions (Wolfenden & Snider, 2001), which considerably exceeds the turnover numbers achieved by chemical catalysts. As a consequence, enzyme-mediated reactions can be performed with much less catalyst load compared to chemical reactions. Moreover, the use of enzymes can significantly reduces the number of reactions steps compared to chemical syntheses, which in turn simplifies the process and lowers production costs and waste production.

- **Environmental compatibility:** Provided by nature, enzymes are non-toxic and completely biodegradable in contrast to most chemical catalysts (*e.g.* heavy metal, or hazardous catalysts).

The aforementioned features of enzymes fulfill many criteria of the concept of *Green Chemistry* and a future bio-based economy. The access to sustainable production processes starting from renewable (fossil fuel-independent) natural resources, including simultaneous reduction of energy consumption, waste production, and total costs, has allowed biocatalysis to become a key technology of the 21th century (BASF pressrelease, 2008). However, it is not only the ecological and economical aspect, which makes biocatalysis an attractive technology. With respect to the production of chiral fine chemicals and pharmaceuticals, one of the most pivotal arguments for the implementation of biocatalysis in organic synthesis is the selectivity of enzymes. There are three major types of selectivity (Faber, 2011):

- **Chemoselectivity:** is the specific reaction of a single type of functional group in the presence of other groups. In contrast to enzymatic reactions, chemical reactions generally tend to be more unselective, which often results in protection/deprotection steps, or the laborious removal of impurities.
- **Regioselectivity:** describes the ability to distinguish between identical functional groups that are located at different positions. Enzymes are most often highly regioselective.
- **Stereoselectivity:** is the preferential formation of one stereoisomer over the other. In contrast to chemical reactions, enzymes are mostly highly stereoselective. Thus, stereoselective synthesis avoids the production of the undesired stereoisomer and therewith the production of waste as well as additional purification steps.

Obviously, there is a huge potential of biocatalysis including many advantages for the chemical industry, and more important for our society and environment. Indeed, enzymatic catalysis is increasingly recognized by organic chemists and biocatalysis has been more frequently used especially in the area of chiral fine chemicals and pharmaceuticals (Gröger & Asano, 2012; Meyer *et al*., 2013a). Furthermore, biocatalysis already reached an industrially proven level (Bornscheuer *et al.*, 2012). There are several examples of successfully applied biotransformations in organic syntheses, such as in the production of the blockbuster drug atorvastatin (Lipitor®) for the treatment of dyslipidemia and the prevention of cardiovascular disease (global sales of almost US$ 11 billion in 2010) (Bornscheuer *et al.*, 2012). However, the potential for sustainable manufacturing is far from being fully realized and the use of enzymes is still below expectation in organic chemistry (Meyer *et al*., 2013a).

1.4.3 Limitations of biocatalysis

Biocatalysis is still often only the second choice technology for the manufacturing of chiral fine chemicals (Meyer *et al*., 2013a). The reasons are versatile and various excellent reviews cover this topic already. In the following, three main aspects are briefly summarized based on the review of Wenda *et al*. (Wenda *et al.*, 2011):

- **Water as solvent:** Organic compounds often reveal poor water solubility, which may reduce the substrate load. In addition, the products often have to be extracted with organic solvents, which in turn diminishes ecological advantages of using water as well as increases the costs of the process.
- **Productivity:** Biotransformations often result in low volumetric productivity due to low activity and/or stability of the catalyst, low substrate solubility, or substrate or/and product inhibition. Low reactant concentrations in turn result in respective huge efforts for downstream processing.
- **Time frame:** Usually, a new biocatalyst has to be developed or at least modified for a certain reaction (enzyme engineering). In general, the time frame with an average of ten years (from laboratory scale to industrial scale) is not competitive with chemical methods (Geibler *et al.*, 2005).

For industry, mainly economic aspects combined with time-to-market pressure determine the choice of the process, whereas ecological aspects are less important (Wenda *et al.*, 2011). In this context, chemists often see biocatalysis as the last alternative approach, if chemical syntheses failed (Thayer, 2006). Therefore, continuous improvement of the technology applied for biocatalysis is required to overcome the existing limitations and obstacles (Wenda *et al.*, 2011). Indeed, there is constant progress in various areas of biotechnology, *e.g.* new ways to apply enzymes in organic solvents (Hernáiz *et al.*, 2010), or new concepts in protein engineering in order to enhance productivity (chapter 1.5). Nonetheless, limitations are still present. In 2004, the Swiss Industrial Biocatalysis Consortium (SIBC) was founded to help overcoming the limitations of biocatalysis in industrial applications. They suggested several approaches to realize the full potential of biocatalysis, *e.g.* ready to use and stable enzymes from all enzyme classes, new technologies for rapid scale up, or new models of collaboration/networking between academia and industry. Particularly, they stressed the need for enzyme toolboxes to make biocatalysis attractive for industrial use (Meyer *et al*., 2013a).

1.5 The toolbox concept

An enzyme toolbox encompasses a group of biocatalysts, which catalyze the same reaction type but with different substrate spectra and selectivity, and thus provide access to respective product platforms (exemplarily shown for the toolbox of thiamine diphosphate (ThDP)-dependent enzymes, Figure 3). The establishment of those platforms may enable a faster reaction to the customers' demands and a more efficient reaction to market fluctuations (Gocke, 2007). The implementation of enzyme toolboxes can be expensive. Nonetheless, variable application and particularly the high selectivity of enzymes make them economically favorable, especially in the pharmaceutical sector, where stereochemically pure drugs are required (Thayer, 2006; Gocke, 2007). Consequently, biotechnology and pharmaceutical companies built up various enzyme toolboxes to provide rapid access to complex chiral compounds (*e.g.* the ChiPros® platform of BASF for the production of chiral amines, alcohols and epoxides) (Thayer, 2006).

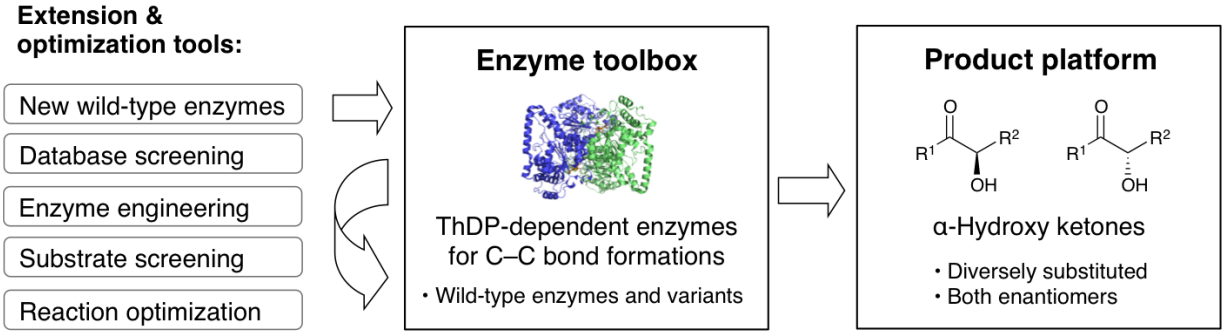


Figure 3: The toolbox strategy exemplarily shown for ThDP-dependent enzymes (according to Gocke, 2007; Sehl *et al*., 2013). New enzymes and variants thereof should continuously enlarge the toolbox in order to provide access to a broad platform of α-hydroxy ketones.

To fully exploit the whole potential of biocatalysis, it is of critical importance to continuously expand the toolbox of biocatalysts with new and improved enzymes in order to meet timeline compressions in the development of *e.g.* pharmaceuticals. Furthermore, enzyme variants should be available to catalyze both enantiocomplementary routes to provide access to both isomers (Meyer *et al*., 2013a). Nowadays, various methods are available to extent and improve the platform of suitable biocatalysts (Figure 3), such as several protein engineering strategies (Bornscheuer & Kazlauskas, 2011). The scientific and technological advances were recently reviewed in detail by Bornscheuer *et al*. (Bornscheuer *et al.*, 2012) and comprise:

- **DNA technology:** Parallel sequence analysis on large scale at low cost by next-generation DNA sequencing (high throughput sequencing) gives access to entire genomes of organisms from different environments including uncultivable organisms (metagenome approach), which in turn facilitates the search for novel biocatalysts (Lorenz & Eck, 2005).
- **Bioinformatics:** Multiple sequence alignments across enzyme families as well as homology searches allow the construction of comprehensive databases (*e.g.* the Thiamine diphosphate-dependent Enzyme Engineering Database (Widmann *et al.*, 2010)), which enable the identification of yet unidentified protein-coding genes and therewith potential biocatalysts (Höhne *et al.*, 2010).
- **Protein engineering:** The rational structure-guided protein design benefits from the rapidly growing RSCB Protein Data Bank (http://www.pdb.org). Currently, almost 94.000 structures are available. The progress in directed evolution mainly contributes to smarter library design (reduction of library size) and improvement of the screening methods in order to increase efficiency and reduce required time. In this context, the application of robotic systems enormously facilitates the laboratory effort.
- **Catalytic promiscuity:** In addition to the actual physiological reaction, an enzyme often catalyzes further transformations (Kazlauskas, 2005). Disabling the natural activity using protein engineering can enhance the desired activity and result in potent biocatalysts, as shown for the ThDP-dependent pyruvate decarboxylase (PDC) (Meyer *et al.*, 2011).

In recent years, the group *Biocatalysis and Biosensors* (IBG-1, Forschungszentrum Jülich GmbH) of Prof. Dr. Martina Pohl followed a similar toolbox strategy by establishing a toolbox of ThDP-dependent enzymes (Gocke, 2007; Gocke *et al*., 2010).

1.5.1 Toolbox of ThDP-dependent enzymes

ThDP-dependent enzymes (chapter 1.7) catalyze asymmetric C–C bond formations, also called carboligations. Carboligations are important reactions in organic synthesis, as they allow the assembly of complex molecules from simple starting molecules. Although C–C bond formations are key steps in most syntheses, they are still not accessible to chemical syntheses at large-scale. As a consequence, the SIBC emphasized that the class of C–C bond-forming enzymes “deserves major investigative efforts” (Meyer *et al.*, 2013a).

The toolbox of ThDP-dependent enzymes was developed for the asymmetric synthesis of chiral α-hydroxy ketones (Figure 3). Owing to the continuously extension of the toolbox with new wild-type enzymes and variants thereof, the resulting overlapping substrate ranges and different selectivities provided access to a broad product platform including various substituted aliphatic, aromatic, and araliphatic α-hydroxy ketones (chapter 1.7.3 and 1.7.5) (Gocke *et al*., 2010; Pohl *et al*., 2012).

1.6 α-Hydroxy ketones – Valuable chiral building blocks

As already described in chapter 1.1.1, organic synthesis addresses the synthesis of complex molecules starting from simple, usually commercially available and cheap material. In general, there are two different strategies: the linear and the convergent synthesis. In the linear synthesis the target molecule is built step by step, whereas in the convergent synthesis individual key intermediates, so-called building blocks or synthons, are combined to the desired product. For the syntheses of highly complex molecules, the latter approach is often preferable to improve the efficiency and yields of the synthesis (Nicolaou & Sorensen, 1996).

α-Hydroxy ketones are versatile and highly valuable building blocks for the pharmaceutical industry (Ward & Singh, 2000; Hoyos *et al*., 2010). Due to their bifunctional nature and the presence of a stereocenter next to a prochiral carbonyl group (Figure 4 A), they represent excellent precursors for further synthetic transformations (Sukumaran & Hanefeld, 2005), such as stereoselective reduction using alcohol dehydrogenases (Kulig *et al.*, 2012), or reductive amination by transaminases (Sehl *et al.*, 2012). The resulting products, 1,2-diols and amino alcohols (Figure 4 A), respectively, are already pharmaceutically active compounds

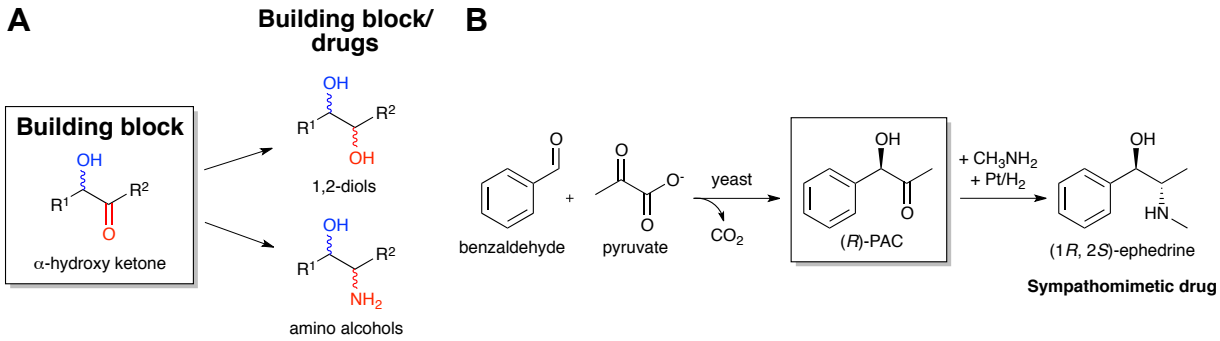


Figure 4: α-Hydroxy ketones as building blocks for synthetic transformations. **A:** α-Hydroxy ketones can be used as precursors for the further reduction to 1,2-diols and amino alcohols, themselves valuable building blocks or pharmaceutically active compounds, respectively. **B:** Chemoenzymatic synthesis of ephedrine by yeast fermentation in the presence of benzaldehyde. The resulting (*R*)-phenylacetylcarbinol (PAC) is chemically reduced in the presence of methylamine to the desired ephedrine (Klawehn & Hilderbrandt, 1934), which is commonly used as *e.g.* stimulant, appetite suppressant, or decongestant.

(*e.g.* pseudoephedrine) or represent valuable synthons (*e.g.* Bergmeier, 2000; Putman *et al.*, 2007). The probably most famous α-hydroxy ketone is phenylacetylcarbinol (PAC), which was used for the production of L-ephedrine already in the early 1930's (Figure 4 B) (Neuberg & Hirsch, 1921). Nowadays, (*R*)-PAC is still a key intermediate for the synthesis of several drugs, such as norephedrine, adrenaline, or amphetamine (Shukla & Kulkarni, 2000).

In virtue of this importance, the development of efficient synthesis strategies for enantioenriched α-hydroxy ketones is a major research focus in the pharmaceutical industry (Hoyos *et al.*, 2010).

1.6.1 Synthesis of α-hydroxy ketones

Several chemical routes to α-hydroxy ketones have been described so far (Figure 5). Among them, the probably most employed strategies are the α-hydroxylation of ketones, the Sharpless asymmetric dihydroxylation of the silylenol ether of the respective ketone (Zhou *et al.*, 2001; Hoyos *et al.*, 2010), and the traditional benzoin condensation-like reaction using chiral thiazolium and triazolium salts (Enders & Kallfass, 2002; O'Toole *et al.*, 2011). However, in many cases these approaches require a significant number of chemical steps, which results in lower overall yields and increased waste production, thus being incompatible with the concept of *Green chemistry* (chapter 1.4.1). Furthermore, although these methods are nowadays often highly chemoselective (O'Toole *et al.*, 2011), they often lack the required high stereoselectivity in pharmaceutical industry (Hoyos *et al.*, 2010) (chapter 1.2.1). At this point, chemical strategies are not competitive with existing biocatalytic methods.

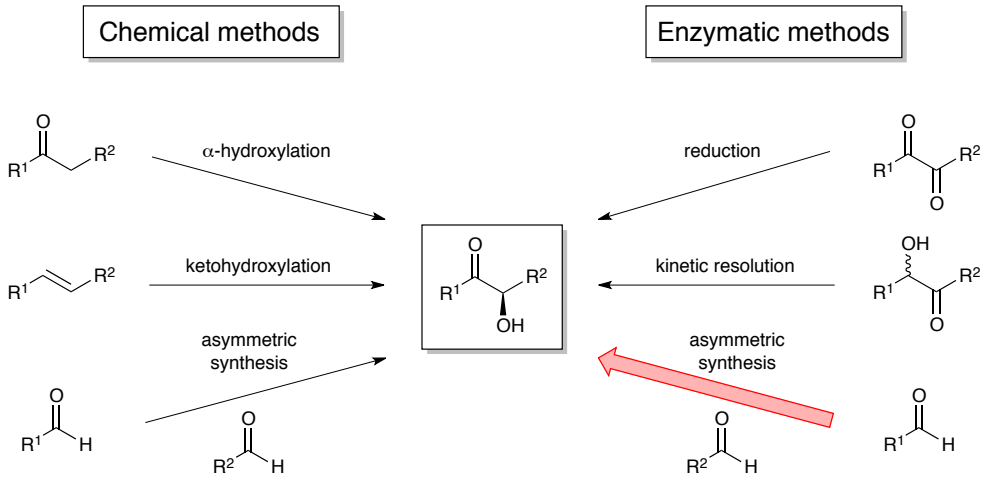


Figure 5: Production of chiral α-hydroxy ketones using the most common chemical and enzymatic methods. Among them, the enzymatic asymmetric synthesis (red arrow) constitutes a highly selective and sustainable method.

Biocatalysis enables the synthesis of α-hydroxy ketones with high chemo-, regio-, and stereoselectivity. The high enantiomeric excesses (*ee*) with up to 99.9 % cannot be reached so far by the means of classic organic asymmetric synthesis. Moreover, such high selectivities are compatible with the concept of *Green chemistry* due to usually economic and environmentally friendly conditions (Patel, 2008, 2011; Hoyos *et al*., 2010). Currently, there are several biocatalytic routes available towards α-hydroxy ketones (Figure 5), including for example the reduction of 1,2-diketones by whole cells (*ee* up to 98%) (*e.g.* Demir *et al*., 2008), or the kinetic resolution of racemic α-hydroxy ketones using benzaldehyde lyase (BAL) (*ee* > 99 %) (*e.g.* Demir *et al*., 2001). In addition, asymmetric synthesis using ThDP-dependent enzymes represents the probably most promising and sustainable method to access chiral α-hydroxy ketones. These reactions can be carried out highly selective (chapter 1.7.3) in a single step starting from readily available and cheap prochiral substrates (*e.g.* aldehydes and α-keto acids) with theoretical yields of > 99 % (Gocke, 2007; Pohl *et al*., 2009; Müller *et al*., 2013). The following chapter describes the properties, the potential, as well as existing limitations of ThDP-dependent enzymes.

1.7 ThDP-dependent enzymes

Thiamine diphosphate, the biologically active derivative of the water-soluble vitamin B_1 (thiamine) (Figure 6), is an essential organic cofactor that exhibits key functions in several central biosynthetic pathways of probably all living organisms, such as in the tricarboxylic acid cycle, in the pentose phosphate pathway, and in the energy-yielding metabolism of carbohydrates of muscle and brain. In addition, it is involved in the biosynthesis of amino acids and neurotransmitters, among others (Frank *et al.*, 2007).

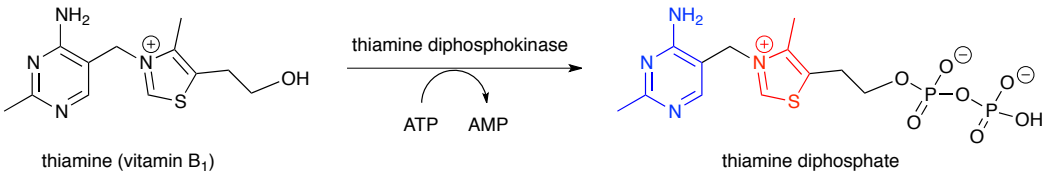


Figure 6: The chemical structure of ThDP, which is synthesized by thiamine diphosphokinase (Leuthardt & Nielsen, 1952). ThDP consists of a six-membered aminopyrimidine ring (blue) linked via a methylene bridge to a five-membered thiazolium ring (red) with a diphosphate-terminated side chain. AMP = adenosine mono-phosphate; ATP = adenosine triphosphate.

Although all organisms use thiamine, only bacteria, fungi, and plants are able to synthesize it. Consequently, animals must ingest it through their diet in order to produce the essential cofactor ThDP. A nutritional deficit in thiamine results in serious neurological disorders,

known as the Wernicke-Korsakoff syndrome and beriberi (Pschyrembel, 2013).

Owing to its chemical properties, ThDP is able to catalyze the formation and cleavage between C–O, C–S, C–N, and most importantly the chemically challenging formation and cleavage of C–H and C–C bonds (Frank *et al.*, 2007). The diversity of thiamine-catalyzed reactions directed the attention and the focus of research on the multifunctionality of ThDP-dependent enzymes and their use as biocatalysts.

1.7.1 General reaction mechanism

Although ThDP-dependent enzymes catalyze diverse reactions, the reaction mechanism is similar in all cases, which is explained in more detail for C–H and C–C bond cleavage and formation reactions in the following. The catalytic activity of ThDP results from the acidic C2 atom of the ThDP-thiazolium ring, which must be first activated by deprotonation to form a potent nucleophilic ylide* (Figure 7 A) (Breslow, 1957).

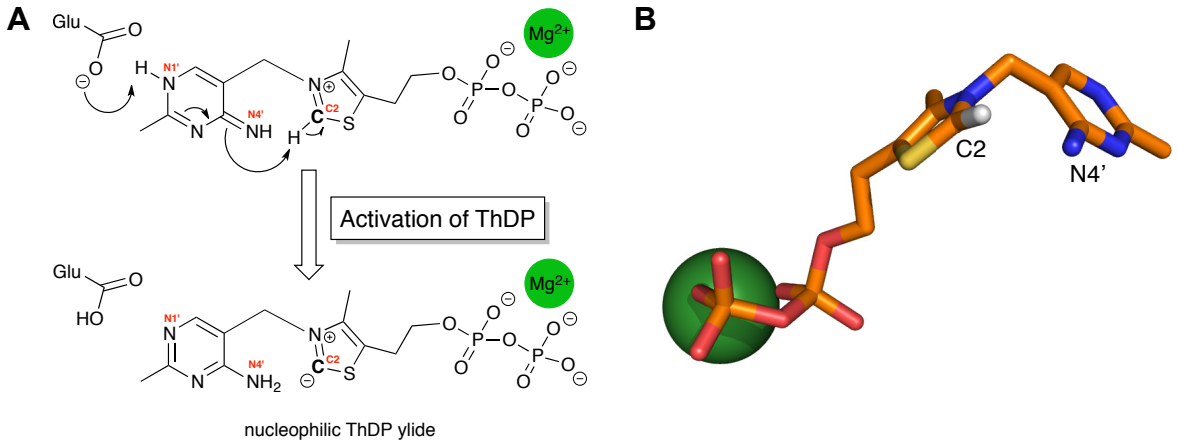


Figure 7: Initial activation of the cofactor ThDP in its conserved "V-conformation". **A:** Activation of ThDP is started with the deprotonation of the N1' by a conserved glutamate. The resulting nucleophilic ThDP ylide is able to attack electrophilic groups (*e.g.* aldehydes) (according to Frank *et al.*, 2007). Extraction of the C2 proton is facilitated by the conserved "V-conformation" of ThDP, placing the 4'-imino group of the pyrimidine ring next to C2 of the thiazolium ring. **B:** The typical "V-conformation" of ThDP shown as PyMOL stick model. ThDP is anchored in the active site via a divalent magnesium ion (green ball). Orange: C; white: H; red: P, O; blue: N; yellow: S.

ThDP itself shows already very poor catalytic activity under physiological conditions. This activity is drastically enhanced in the global protein environment of ThDP-dependent enzymes, where ThDP is anchored in the active site via a divalent metal ion (usually Mg^{2+}) and aligned in the so-called "V-conformation" (Figure 7 B, see also chapter 1.7.2)

*An ylide is a reactive dipolar compound, in which a formally negatively charged atom (usually a carbanion) is attached directly to a formally positively charged heteroatom (usually nitrogen, phosphorus, or sulfur) (IUPAC, 1997).

(Dobritzsch *et al.*, 1998; Frank *et al.*, 2007). This special alignment, in which both aromatic rings of the cofactor are linked via the methylene bridge in a V-shaped conformation, allows the thiazolium C2 to be in reactive distance to the 4-imino group of the pyrimidine ring that is responsible for abstracting the C2 proton (Figure 7 A) (Jordan, 2003; Frank *et al.*, 2007). The "V-conformation" is highly conserved and almost identical in all of the structurally characterized ThDP-dependent enzymes. After the initial deprotonation step, which is supported exclusively by a glutamate residue in ThDP-dependent enzymes (Kern *et al.*, 1997; Shaanan & Chipman, 2009), the nucleophilic ThDP-thiazolium ylide is able to attack electrophilic groups, such as carbonyls (*e.g.* aldehydes, α-keto acids) (Figure 8).

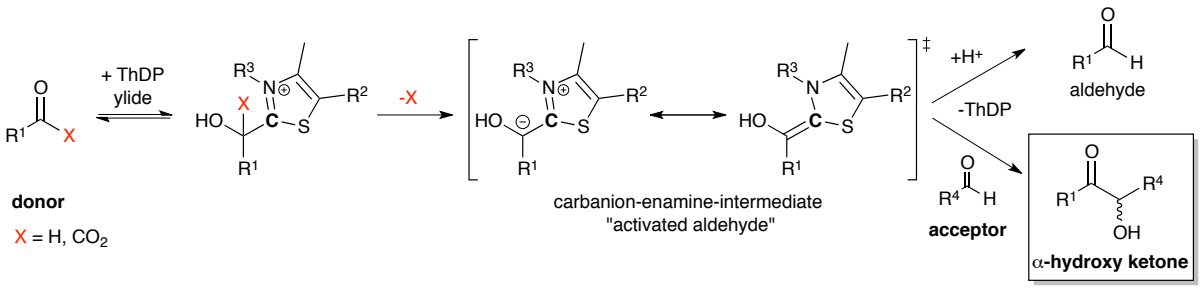


Figure 8: The general reaction mechanism of ThDP-dependent enzymes (according to Schellenberger, 1998). The substrate that binds first to the ThDP-ylide is called donor, whereas the second substrate, which is attacked by the "activated aldehyde", is called acceptor. R_3: pyrimidine ring of ThDP; R_4: diphosphate rest of ThDP.

The so-called donor substrate binds covalently to C2 of the ThDP-thiazolium, which involves the deprotonation (for aldehydes) or decarboxylation (for the corresponding α-keto acid) of the donor (Figure 8). As a result, the same metastable carbanion/enamine intermediate, also called "activated aldehyde" or "Breslow intermediate", is formed, which now has nucleophilic reactivity at the carbon atom that previously had been part of the carbonyl group in the donor substrate (Kluger & Tittmann, 2008). This principle is called "*Umpolung*". Subsequently, the ThDP-bound "activated aldehyde" is either protonated to release the corresponding aldehyde, or another specific electrophilic second substrate (the so-called acceptor) binds and reacts (ligation) leading to diverse products, such as α-hydroxy ketones. Finally, the resulting ThDP-thiazolium ylide* is regenerated via protonation catalyzed by specific proton relay systems in the active site (Frank *et al.*, 2007).

The enzymatic mechanism of ThDP-dependent enzymes, which enables a diverse range of reactions, seems to be highly conserved as a consequence of unique structural features. The

*Meanwhile, the first stable carbene at the active site of ThDP-dependent enzymes was observed (Meyer *et al.*, 2013b). However, it is not possible to define whether a carbene or carbanion acts as nucleophile due to the intrinsic nucleophilicity of both species.

common feature of all ThDP-dependent enzymes is the binding mechanism of ThDP to the protein.

1.7.2 Structure of ThDP-dependent enzymes

Although ThDP-dependent enzymes usually exhibit sequence identities of less than 20 %, most of the tertiary structures (folds), which were found in structures from eukaryotic, prokaryotic and archeal sources, are very similar (Pohl *et al.*, 2009). This discrepancy between sequence and structure conservation is not unusual within an enzyme family and indicates a divergent evolution from a common ancestor (Frank *et al.*, 2007).

Secondary structure

Related to the binding and activation of ThDP, two structurally conserved protein domains are necessary: the pyrimidine (PYR)-domain and the pyrophosphate (PP)-domain (Duggleby, 2006). The PYR-domain, as the name already indicates, interacts with the pyrimidine ring of ThDP. In addition, it contains the conserved glutamate for cofactor activation. The PP-domain is characterized by a common structural motif, the $GDG(X)_{25\text{-}30}N$ motif (Hawkins *et al.*, 1989), and is responsible for the binding of the divalent cation (*e.g.* Mg^{2+}), which in turn serves as anchor for the diphosphate group of ThDP (Dobritzsch *et al.*, 1998).

In addition to the highly conserved PYR- and PP-domains, ThDP-dependent enzymes often possess other less conserved domains, such as the TH3-domain in the superfamily of decarboxylases or the TKC-domain in the transketolase superfamily. These domains are usually not involved in the binding of ThDP and in many cases their function in the catalytic process is still unclear (Duggleby, 2006; Costelloe *et al*., 2008).

Quaternary structure

ThDP-dependent enzymes all function as dimer of active sites (Frank *et al.*, 2007) with two ThDPs bound at the interface of two monomers (subunits). Thereby, the PYR-domain of ThDP interacts with one subunit, whereas the diphosphate residue interacts with the PP-domain of the neighboring subunit (Dobritzsch *et al*., 1998; Hasson *et al*., 1998). Thus, both domains are responsible for the correct alignment of ThDP in its “V-conformation” (Muller *et al.*, 1993). Moreover, a large conserved hydrophobic residue next to the thiazolium ring supports this conformation (Lindqvist *et al*., 1992; Guo *et al*., 1998), although, it is not necessarily required to maintain the “V-conformation” (Andrews *et al.*, 2013). Consequently, a dimer with two active sites represents the minimal functional unit (Muller *et al.*, 1993). Figure 9 exemplarily demonstrates the domain arrangement in the quaternary structure of the

ThDP-dependent enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) from *Bacillus subtilis* (*Bs*MenD) (Dawson *et al.*, 2010). ThDP-dependent enzymes are mostly active either as homodimers, *e.g.* MenD from *Escherichia coli* (Dawson *et al.*, 2008), or as homotetramers, *e.g.* pyruvate decarboxylase (PDC) from *Acetobacter pasteurianus* (*Ap*PDC) (Rother *et al.*, 2011). The tetrameric structure can be best described as dimer of dimers (Duggleby, 2006). Therefore, often both association states can be found to a certain extent in respective enzyme crystals (*e.g.* Dawson *et al.*, 2008). Moreover, even higher-order assemblies such as octamers are not unusual in aqueous solutions of *e.g.* PDCs (Pohl *et al.*, 1994; König, 1998).

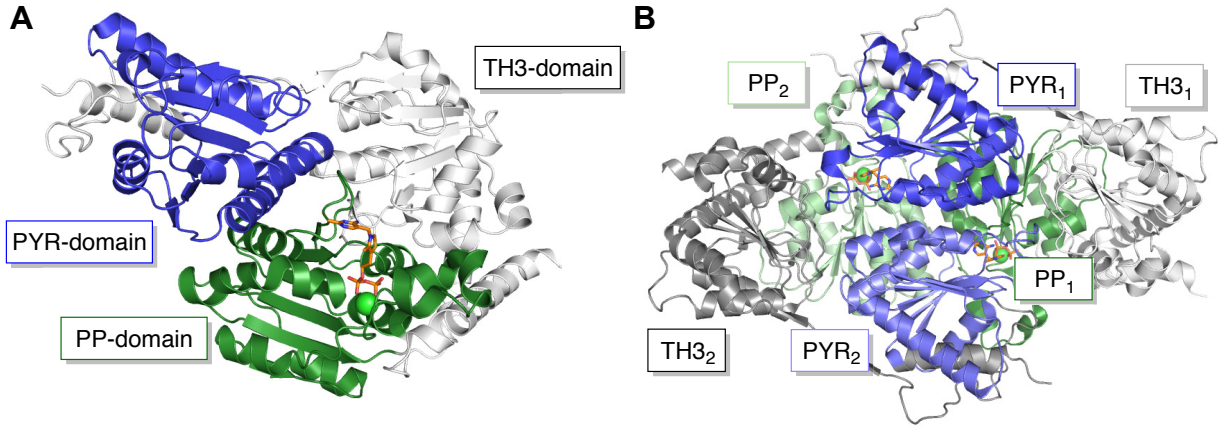
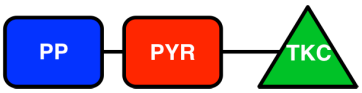


Figure 9: Domain structure of ThDP-dependent enzymes exemplarily shown for *Bs*MenD (pdb: 2X7J). **A:** The pyrimidine (PYR)-, pyrophosphate (PP)-, and TH3-domain form one monomer (subunit). ThDP (orange) is anchored in the PP-domain via Mn^{2+} (green ball). **B:** Two subunits form the catalytically active dimer containing two active sites. The ThDP is located at the interface of both subunits and interacts with the PYR-domain of one subunit and the PP-domain of the other subunit (*e.g.* PYR_1 and PP_2).

Classification

ThDP-dependent enzymes can be classified by different approaches, based on their biological functions, sequences, or their structures. Depending on the arrangement of the conserved PYR- and PP-domains and the existence of additional domains, the currently known ThDP-dependent enzymes can be grouped into eight structural superfamilies (Duggleby, 2006; Costelloe *et al.*, 2008; Widmann *et al.*, 2010). Among them, the decarboxylase (DC) superfamily and the transketolase (TK) superfamily have been most extensively characterized, due to their high catalytic potential in C–C bond formations. The characteristics of both families are briefly summarized in Table 1 according to the recently published ThDP-dependent Enzyme Engineering Database (TEED) (Widmann *et al.*, 2010).

Table 1: Characteristics of the DC and TK superfamily in the class of ThDP-dependent enzymes[1].

Superfamily	Domain arrangement[2]	Number of proteins with sequence information	Number of proteins with structure information
DC		3008	223
TK		1972	32

[1]according to Widmann *et al*., 2010
[2]according to Costelloe *et al*., 2008

In this thesis, enzymes of the DC superfamily were further investigated with respect to structure-function relationships and their biocatalytic potential to produce valuable α-hydroxy ketones. In the following, all ThDP-dependent enzymes relevant for this work are briefly introduced.

1.7.3 ThDP-dependent enzymes of the DC superfamily

With few exceptions, ThDP-dependent enzymes of the DC superfamily belong to the enzyme classes of transferases (EC 2) and lyases (EC 4). According to the *International Union of Biochemistry and Molecular Biology* transferases catalyze the transfer of functional groups (*e.g.* methyl, acyl, or amino groups) from one substrate (the so-called donor) to another (the so-called acceptor), whereas lyases catalyze the non-hydrolytic formation or cleavage of bonds, such as C–C, C–S, or C–N bonds (Enzyme Nomenclature, 1992). Table 2 gives an overview about all relevant enzymes for this thesis and their physiological roles in the respective organism. Although the described ThDP-dependent enzymes show broad physiological activities, all of them catalyze the 1,2-addition* of two aldehydes, yielding the desired α-hydroxy ketones (Figure 8). Owing to their usually high chemo- and stereoselectivity in these carboligation reactions, they became attractive tools in organic synthesis because competitive selective chemical methods are still missing (chapter 1.6.1) (Brovetto *et al*., 2011; Pohl *et al*., 2012). In the following, the state of the art is briefly summarized with respect to the wild-type enzymes and their carboligation potential in the 1,2-addition. Progress or developments concerning reaction engineering and operation modes

*A 1,2-addition is a nucleophilic addition reaction, where the addition occurs at two adjacent sites (IUPAC, 1997). In this case the carbonyl oxygen (1) is attached by the electrophile (H^+) and the carbonyl carbon (2) attaches to the nucleophile (the "activated aldehyde").

(*e.g.* membrane reactor, two-phase systems, immobilized enzymes) to further improve C–C bond formations catalyzed by ThDP-dependent enzyme are not addressed at this point.

Table 2: ThDP-dependent enzymes of the DC superfamily important in this work.

Enzyme / EC number	Organism	Physiological activity	Reference
AHAS 2.2.1.6	*Escherichia coli*	Biosynthesis of branched-chain amino acids	Umbarger & Brown, 1958
BAL 4.1.2.38	*Pseudomonas fluorescens*	Unknown	González & Vicuña, 1989
BFD 4.1.1.7	*Pseudomonas putida*	Mandelate catabolism	Hegeman, 1970
KdcA 4.1.1.72	*Lactococcus lactis*	Leucine metabolism	Smit *et al*., 2005
MenD 2.2.1.9	*Bacillus subtilis* *Escherichia coli*	Biosynthesis of menaquinone	Jiang *et al*., 2007
PDC 4.1.1.1	*Acetobacter pasteurianus*	Oxidative lactic acid metabolism	Raj *et al*., 2001

AHAS = acetohydroxyacid synthase
BFD = benzoylformate decarboxylase

Acetohydroxyacid synthase from Escherichia coli

Acetohydroxyacid synthases (AHAS) have been found in plants, fungi, and bacteria, but not in mammals so far, which makes them interesting targets for specific inhibitors, *e.g.* herbicides or antibiotics (Umbarger & Brown, 1958; Umbarger, 1978; Chipman *et al*., 1998). AHAS is a key enzyme in the biosynthesis of the branched-chain amino acids valine, leucine, and isoleucine, catalyzing the decarboxylation of pyruvate and the subsequent addition to either a second pyruvate or to ketobutyrate yielding (*S*)-acetolactate and (*S*)-acetohydroxy butyrate, respectively (Vinogradov *et al.*, 2006). Three AHAS isozymes are encoded in *E. coli* (*Ec*AHAS-I-III) that differ with respect to their biochemical properties (*e.g.* substrate specificity and sensitivity towards inhibition) (Vinogradov *et al.*, 2006). Furthermore, they exhibit three major differences in comparison to most of the other ThDP-dependent enzymes of the DC superfamily:

- **Structure:** *Ec*AHAS is composed of two catalytic subunits, which is similar to other ThDP-dependent enzymes. However, each catalytic subunit is associated with an additional regulatory subunit, which is required for activation and regulation of feedback inhibition (Schloss *et al*., 1985; Weinstock *et al*., 1992; Vyazmensky *et al*., 2009).

- **Cofactor:** *Ec*AHAS requires flavin adenine dinucleotide (FAD) as additional cofactor, which serves as structural element and is not involved in catalysis (Steinmetz, 2010).
- **Catalytic mechanism:** The active site of *Ec*AHAS suppresses the protonation of the ThDP-bound donor ("activated aldehyde") formed from pyruvate, thus preventing the release of the respective aldehyde (acetaldehyde) (Chipman *et al.*, 1998).

The *Ec*AHAS isozymes I-III have been recently characterized concerning their carboligation potential with non-physiological acceptor substrates. All three enzymes catalyze the highly selective formation of the valuable precursor (*R*)-PAC (*ee* > 97 %), originating from pyruvate as the donor and benzaldehyde as the acceptor (Figure 10), however, with different efficiency (Engel *et al.*, 2003).

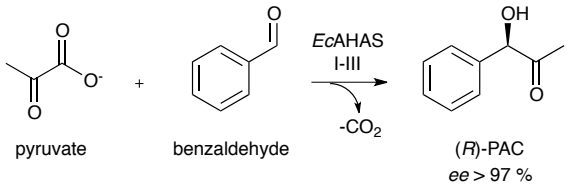


Figure 10: Mixed carboligation of pyruvate and benzaldehyde to (*R*)-PAC catalyzed by *Ec*AHAS-I-III.

*Ec*AHAS-I catalyzes the formation of (*S*)-acetolactate with higher initial rate compared to the formation of (*R*)-PAC in presence of pyruvate and benzaldehyde. However, (*S*)-acetolactate is cleaved again during the reaction. The resulting "activated aldehyde" and pyruvate can be used for the carboligation with benzaldehyde, which enables the formation of (*R*)-PAC with high conversions (Engel *et al.*, 2003). In contrast, *Ec*AHAS-II is not able to catalyze the cleavage of the formed acetolactate again, resulting in accumulated acetolactate and lower conversions to (*R*)-PAC relative to *Ec*AHAS-I. Furthermore, *Ec*AHAS-II is more sensitive towards product inhibition compared to *Ec*AHAS-I (Engel *et al.*, 2003). Unlike *Ec*AHAS-I and II, *Ec*AHAS-III exhibits a poor ability to produce PAC resulting in much lower conversions (Engel *et al.*, 2003).

*Ec*AHAS-I, the most effective *Ec*AHAS isozymes for the chiral synthesis of (*R*)-PAC, also catalyzes the carboligation of pyruvate and monosubstituted benzaldehydes or heterocyclic and heteroaromatic aldehydes, respectively, to the corresponding (*R*)-arylacetylcarbinols with high conversions (Engel *et al.*, 2004a). However, other donor substrates than pyruvate have not been tested so far.

Benzaldehyde lyase from Pseudomonas fluorescens

BAL from *P. fluorescens* (*Pf*BAL) was first described in 1989 (González & Vicuña, 1989), cloned and characterized later on in 1994 (Hinrichsen *et al.*, 1994). Up to now, it is the only biochemically characterized BAL with reasonable activity. The physiological function of this enzyme is still unknown. However, the gram-negative bacterium is able to grow on lignin-derived anisoin (4,4'-dimethoxybenzoin) and benzoin as the sole carbon and energy sources. *Pf*BAL can cleave these hydroxy ketones to the respective aldehydes, which in turn may be metabolized in the β-ketoadipate pathway (Stanier & Ornston, 1973).

Initially, the cleavage reaction (lyase activity) of (*R*)-benzoin catalyzed by *Pf*BAL was assumed to be irreversible (González & Vicuña, 1989). However, it also catalyzes the formation of (*R*)-benzoin (Figure 11 A) with remarkably high specific activity of 320 U mg^{-1} (Janzen *et al.*, 2006) and excellent stereoselectivity (*ee* > 99 %) (Demir *et al*., 2001). For this benzoin formation the substrate range of *Pf*BAL is very broad including various *ortho*-, *meta*-, and *para*-substituted benzaldehydes (Demir *et al*., 2002). Furthermore, *Pf*BAL catalyzes the selective cross-coupling reaction of aromatic aldehydes giving access to enantioenriched mixed benzoins (Dünkelmann *et al.*, 2002), which is not feasible by classical organic chemistry.

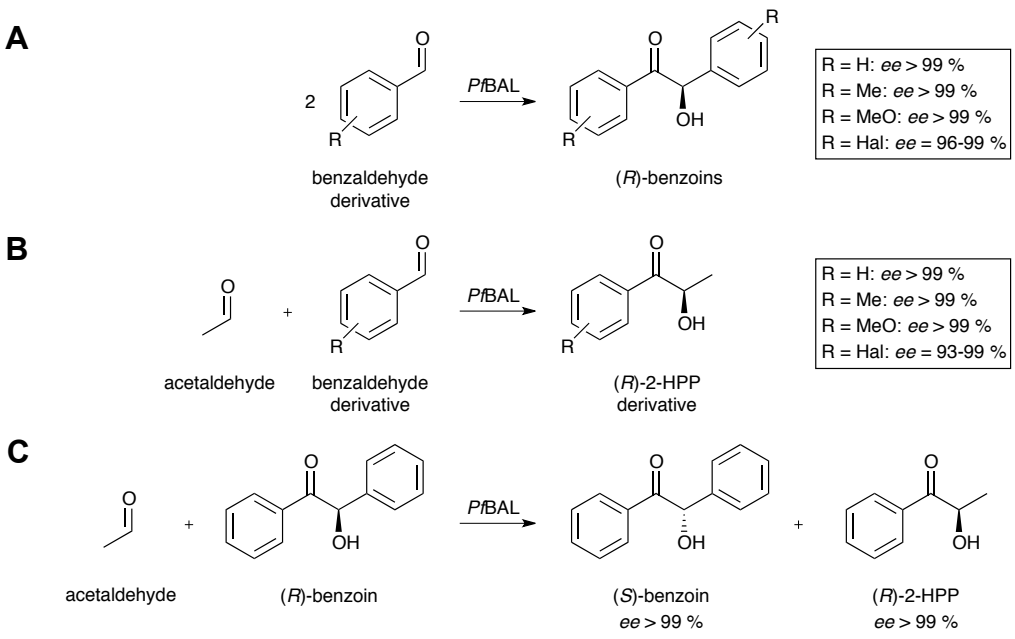


Figure 11: Examples of carboligation reactions catalyzed by *Pf*BAL. **A:** Homocoupling of benzaldehyde derivatives to (*R*)-benzoins (Demir *et al*., 2002). **B:** Mixed carboligation of acetaldehyde and benzaldehyde derivatives yielding (*R*)-2-hydroxypropiophenones (2-HPP) (Demir *et al*., 2002). **C:** Kinetic resolution of racemic benzoin in the presence of acetaldehyde yielding (*R*)-2-HPP and (*S*)-benzoin (Demir *et al*., 2001).

In addition to the ligation of two benzaldehyde molecules, *Pf*BAL also catalyzes the mixed carboligation of benzaldehyde derivatives as donor and different aliphatic acceptor aldehydes, such as acetaldehyde (Figure 11 B), yielding (*R*)-2-hydroxypropiophenone (2-HPP) derivatives with excellent *ee*s (Demir *et al*., 2001, 2002, 2003). Furthermore, owing to its high (*R*)-selectivity and the ability to cleave acyloin linkages, which has not yet been observed with other ThDP-dependent enzymes (Müller *et al.*, 2009), *Pf*BAL can be used for kinetic resolution of racemic benzoin (Figure 11 C) yielding (*S*)-benzoin and (*R*)-2-HPP (*ee* > 99 %) (Demir *et al*., 2001).

Benzoylformate decarboxylase from Pseudomonas putida

Benzoylformate decarboxylases (BFD) have been found in several organisms, such as *Acinetobacter calcoaceticus* (Barrowman & Fewson, 1985) or *Bradyrhizobium japonicum* (Wendorff, 2006). BFD from *P. putida* (*Pp*BFD), which was first described in 1966 (Hegeman, 1966), is by far the best-characterized BFD. In addition to its physiological activity of phenylglyoxylic acid (benzoylformate) decarboxylation to benzaldehyde and CO_2 (Hegeman, 1966, 1970), *Pp*BFD is able to catalyze C–C bond formations (Wilcocks & Ward, 1992; Wilcocks *et al*., 1992) with the same chemoselectivity and a similar broad substrate spectrum as exhibited by *Pf*BAL. The substrate range encompasses various aromatic, heteroaromatic, aliphatic, cyclic aliphatic as well as olefinic aldehydes providing access to 2-HPP analogues with high conversions and high *ee*s (Dünnwald *et al*., 2000; Iding *et al*., 2000). In the case of acetaldehyde as the acceptor, *Pp*BFD catalyzes the formation of the respective 2-HPP derivative highly *S*-selective (Figure 12) (Wilcocks & Ward, 1992; Iding *et al*., 2000) in contrast to the strictly *R*-selective *Pf*BAL.

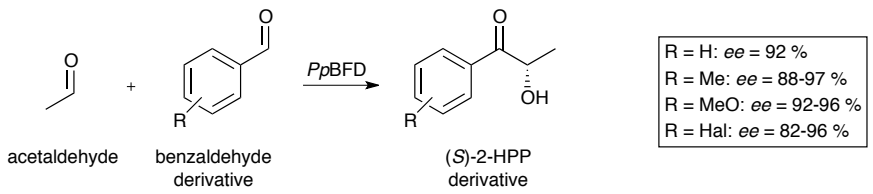


Figure 12: *Pp*BFD-catalyzed mixed carboligation reactions of acetaldehyde and benzaldehyde derivatives yielding (*S*)-2-HPP analogues (Wilcocks *et al*., 1992; Dünnwald *et al*., 2000; Iding *et al*., 2000).

Branched-chain keto acid decarboxylase from Lactococcus lactis

Branched-chain keto acid decarboxylases (KdcA) have been described from different organisms, such as *B. subtilis* (Oku & Kaneda, 1988) or *L. lactis* (Amárita *et al.*, 2001), and represent putatively key enzymes in the metabolism of leucine (Smit *et al.*, 2005). The KdcA from *L. lactis* (*Ll*KdcA) was earlier characterized with respect to its carboligation potential in

the asymmetric synthesis (Gocke *et al*., 2007). In addition to the exceptionally broad substrate range for the decarboxylation reaction (Smit *et al*., 2005; Yep *et al*., 2006), *Ll*KdcA exhibits a broad substrate spectrum for the carboligation reaction. The substrate range includes different aromatic, linear and branched-chain aliphatic, as well as CH-acidic aldehydes (*e.g.* phenylacetaldehyde), giving access to different α-hydroxy ketones with high *ee*s, such as new (*R*)-PAC derivatives (Figure 13) (Gocke *et al*., 2007).

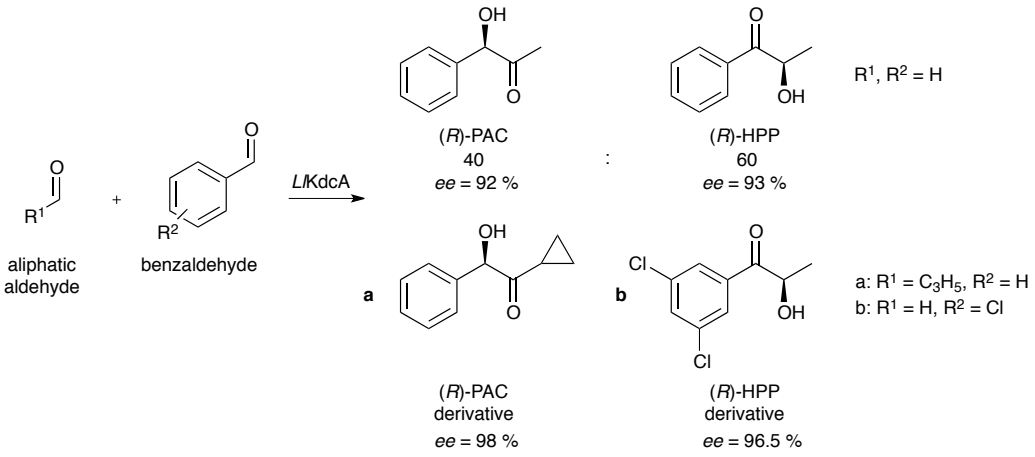


Figure 13: Examples of mixed carboligation reactions catalyzed by *Ll*KdcA. *Ll*KdcA catalyzes highly *R*-selectively the formation of PAC derivatives such as 2-cyclopropyl-1-hydroxy-1-phenylethanone (**a**). The chemoselectivity of *Ll*KdcA depends on the choice of donor and acceptor. Whereas the addition of acetaldehyde to benzaldehyde leads to a product mixture of (*R*)-PAC and (*R*)-2-HPP, the use of larger aliphatic substrates or substituted benzaldehydes results in the formation of solely the (*R*)-PAC or the (*R*)-2-HPP derivative, respectively (Gocke *et al*., 2007).

Interestingly, *Ll*KdcA does not show the usually high chemoselectivity of ThDP-dependent enzymes in the mixed carboligation reaction (chapter 1.7.4). The addition of acetaldehyde and benzaldehyde leads to almost equal amounts of (*R*)-PAC and (*R*)-2-HPP. However, the chemoselectivity can be shifted depending on the choice of donor and acceptor substrate (Figure 13) (Gocke *et al*., 2007).

MenD from Bacillus subtilis and Escherichia coli

The 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate (SEPHCHC) synthase (MenD*) is a relatively new member of the thiamine enzyme family. MenD catalyzes the decarboxylation of α-ketoglutarate (α-KG) and the subsequent addition to the physiological acceptor isochorismate yielding SEPHCHC, known as the second step in menaquinone (vitamin K) biosynthesis (Jiang *et al.*, 2007). MenD is one of the few ThDP-dependent enzymes that are able to use α-KG as a natural donor (Beigi *et al.*, 2013). This ability is of

*The name MenD refers to the name of the gene *menD* that is part of the menaquinone operon.

particular interest for biocatalytic applications, because the use of α-KG in 1,2-addition reactions enables a selective C4 chain elongation with a terminal carboxyl group (Figure 14). The resulting δ-hydroxy-γ-keto acids represent valuable functionalized α-hydroxy ketones and can be used as precursors of γ- and δ-lactones (Cavinato *et al.*, 1994), themselves valuable intermediates in organic synthesis (Yadav & Mandal, 2011). The carboligation potential of MenD from *E. coli* (*Ec*MenD) has recently been characterized with respect to the 1,2-addition of α-KG and a broad range of non-physiological acceptor substrates that include aliphatic, aromatic, and α,β-unsaturated aldehydes. Thereby, the carboligation of α-KG as donor and various benzaldehyde derivatives as acceptors (Figure 14) is characterized by an excellent chemo- and *R*-selectivity yielding exclusively 5-hydroxy-4-oxo-5-arylpentanoates (PAC derivatives) with high *ee*s. The self-ligation of two molecules of α-KG could not be observed (Kurutsch *et al*., 2009; Beigi *et al*., 2013).

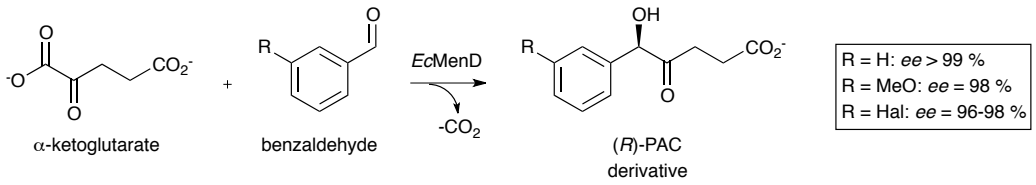


Figure 14: The carboligation potential of *Ec*MenD exemplarily shown for the 1,2-addition of α-ketoglutarate to meta-substituted benzaldehydes yielding (*R*)-5-hydroxy-4-oxo-5-arylpentanoates (PAC derivatives) (Kurutsch *et al*., 2009; Beigi *et al*., 2013).

Recently, a non-physiological reductive C–N bond formation between α-KG and nitroarenes has been reported for *Bs*MenD (Chen *et al.*, 2012). However, the non-physiological C–C bond formation between α-KG and aromatic aldehydes has not yet been described.

Pyruvate decarboxylase from Acetobacter pasteurianus

PDCs are central enzymes in the anaerobic fermentation of glucose to ethanol, where they catalyze the non-oxidative decarboxylation of pyruvate to acetaldehyde. PDCs have been already described in several organisms, such as in plants, fungi, yeasts, and bacteria (*e.g.* Holzer *et al*., 1956; Alvarez *et al*., 1993; Hossain *et al*., 1996; König, 1998). The carboligation activity of PDC from *Saccharomyces cerevisiae* has been used for almost 100 years for the production of the famous chiral building block (*R*)-PAC starting from pyruvate as donor and benzaldehyde as acceptor (chapter 1.6, Figure 4 B) (Neuberg & Hirsch, 1921). Since then, several PDCs have been thoroughly characterized, such as the PDC from *Zymomonas mobilis* (*e.g.* Bringer-Meyer & Sahm, 1988; Bornemann *et al*., 1996). The PDC from *A. pasteurianus* (*Ap*PDC) is a relatively new member of the PDC family. In contrast to other PDCs, *Ap*PDC is a key enzyme in the oxidative metabolism (Raj *et al.*, 2001). The

carboligation potential of *Ap*PDC has recently been investigated. *Ap*PDC catalyzes the mixed carboligation of acetaldehyde and benzaldehyde yielding (*R*)-PAC as the main product (Figure 15) and acetoin as a side product, originating from the homocoupling of acetaldehyde (Rother *et al.*, 2011). Moreover, *Ap*PDC is able to use larger aliphatic aldehydes as donor substrate (Rother *et al.*, 2011) as well as the respective α-keto acids, *e.g.* pyruvate or ketobutyrate (Gómez Baraibar *et al.*, 2013).

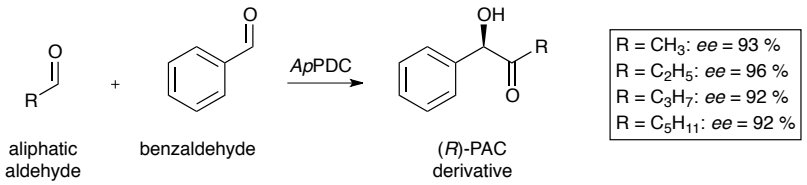


Figure 15: The mixed carboligation of linear aliphatic aldehydes and benzaldehyde catalyzed by *Ap*PDC yielding (*R*)-PAC derivatives with high *ee*s (Rother *et al.*, 2011).

1.7.4 Principles of chemo- and stereoselectivity – Structure-function relationships

Although the general reaction mechanism is highly conserved (chapter 1.7.1), ThDP-dependent enzymes reveal considerable differences in the chemo- and stereoselectivity of mixed carboligation reactions, which leads to a diverse product range (chapter 1.7.3). Taking the mixed carboligation of acetaldehyde and benzaldehyde as example, which is catalyzed by several ThDP-dependent enzymes, four different products can be principally formed, each of them in both stereoisomers (Figure 16).

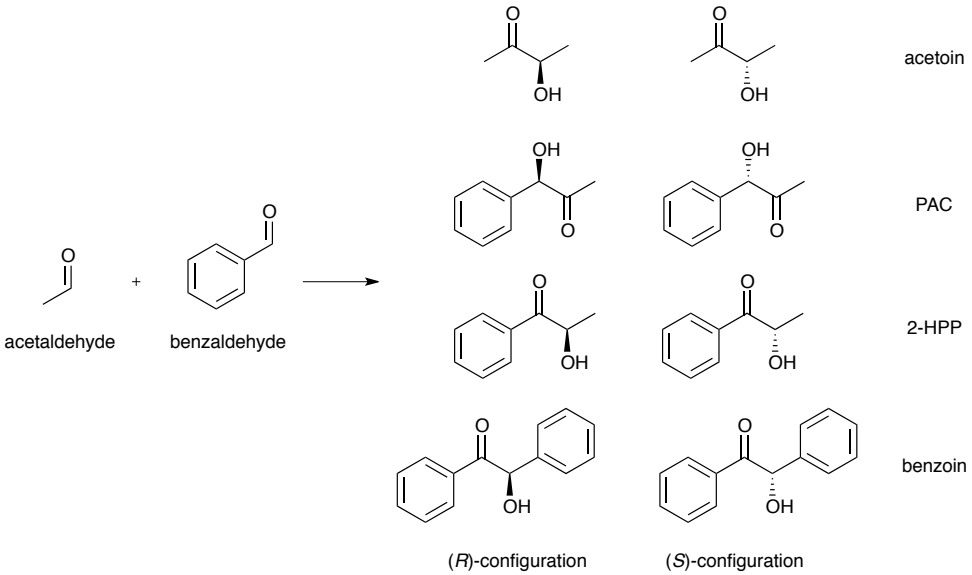


Figure 16: Possible products of the mixed carboligation of acetaldehyde and benzaldehyde catalyzed by ThDP-dependent enzymes (according to Gocke, 2007).

ThDP-dependent enzymes of the DC superfamily are usually highly selective for one or two of these possible products (chapter 1.7.3) due to a high variation of the active site residues that results in different active site architectures (Gocke, 2007; Pohl *et al.*, 2009). In the following, the principles of chemo- and stereoselectivity are explained in more detail.

Chemoselectivity

Chemoselectivity describes the donor and acceptor selectivity and is determined by the binding order of the substrates. The self-ligation products acetoin and benzoin are derived from the coupling of two molecules acetaldehyde or benzaldehyde, respectively. In contrast, the mixed product PAC is obtained from acetaldehyde as the donor and benzaldehyde as the acceptor, whereas 2-HPP is derived from benzaldehyde as the donor and acetaldehyde as the acceptor (Figure 17). Chemoselectivity depends almost exclusively on the available space for the donor and acceptor in the respective substrate-binding pocket (Pohl *et al.*, 2009). For example, *Pf*BAL has a large donor-binding site, which allows the binding of large aromatic donor substrates (Figure 17 A). In contrast, *Ap*PDC is characterized by a small donor-binding site, which allows preferentially the binding of small aliphatic donor substrates (Figure 17 B).

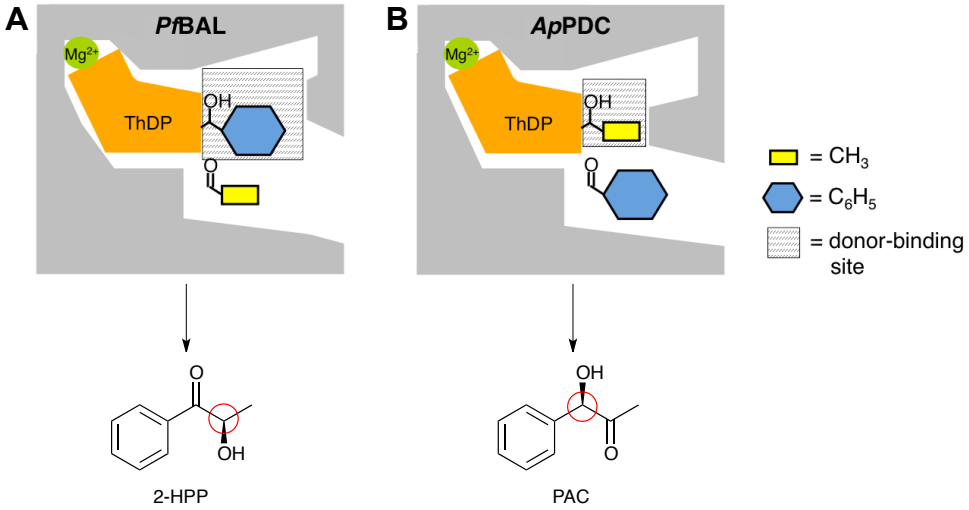


Figure 17: Principles of chemoselective carboligation exemplarily shown as schematic representation for the formation of 2-HPP and PAC catalyzed by *Pf*BAL and *Ap*PDC, respectively (Gocke, 2007; Pohl *et al.*, 2009). Chemoselectivity depends on the binding order of the substrates. **A:** The large donor-binding site of *Pf*BAL is optimal to stabilize benzaldehyde as donor. Acetaldehyde acts as acceptor yielding 2-HPP. **B:** The donor-binding site of *Ap*PDC is restricted to small donor aldehydes, such as acetaldehyde. In this case benzaldehyde functions as acceptor yielding the respective PAC product.

In addition to the steric properties of the active site, optimal stabilization of the donor in the binding pocket is another important criterion for high chemoselectivity (Pohl *et al.*, 2009). For example, although there are no steric restrictions for acetaldehyde to bind in the donor-binding site of *Pf*BAL, neither acetoin nor the PAC product were observed in the mixed

carboligation (Gerhards *et al*., 2012), due to the missing stabilization of acetaldehyde and/or the better stabilization of benzaldehyde as donor. In the case of *Ec*MenD, the donor-binding site is tailor-made for α-KG as donor, owing to a very polar environment with particularly two highly conserved arginine residues that are able to stabilize the terminal carboxylate of α-KG (Dawson *et al*., 2008). As a consequence, *Ec*MenD is highly specific for α-KG, which exclusively leads to the formation of the respective PAC derivative (chapter 1.7.3, Figure 14). Furthermore, other donor substrates (*e.g*. pyruvate) are only poorly accepted by *Ec*MenD (Kurutsch *et al.*, 2009).

Consequently, ThDP-dependent enzymes of the DC superfamily discriminate in most cases highly specifically between the donor and acceptor substrate in the mixed carboligation as long as both substrates are sufficiently different, thereby allowing the selective formation of predominantly one product (Pohl *et al.*, 2009).

Stereoselectivity

With the exception of *Pp*BFD, which catalyzes the formation of (*S*)-2-HPP (*ee* = 92 %), all ThDP-dependent enzymes from the DC superfamily are highly *R*-selective in the carboligation reaction of solely aromatic or aromatic and aliphatic aldehydes, respectively (chapter 1.7.3). Interestingly, *Pp*BFD is only *S*-selective in the presence of acetaldehyde as acceptor. Larger acceptor aldehydes are transformed *R*-selectively again (Wilcocks *et al*., 1992; Dünnwald *et al*., 2000; Iding *et al*., 2000; Gocke *et al*., 2008).

The underlying principles of stereoselectivity could be elucidated on a molecular level based on the comparison of the active sites of *Pp*BFD (Hasson *et al.*, 1998) and *Pf*BAL (Mosbacher *et al.*, 2005). These studies revealed that the stereocontrol of carboligation reactions in ThDP-dependent enzymes is determined by the relative orientation of donor and acceptor prior to carboligation (Knoll *et al.*, 2006). A parallel arrangement of both side chains results in the formation of (*R*)-products, whereas (*S*)-products are derived from an antiparallel arrangement (Figure 18). This model is based on the assumption that the hydroxyl group of the ThDP-bound donor and the carbonyl oxygen of the acceptor substrate point in the same direction towards the amino acid residues, which are essential for the proton transfer during the catalytic cycle (chapter 1.7.1). Those residues are rare in the active site of ThDP-dependent enzymes of the DC superfamily, making the orientation of both oxygen functions in the same direction very likely (Knoll *et al*., 2006; Hailes *et al*., 2013).

In *Pp*BFD a structural element called "*S*-pocket" was found that enables the antiparallel arrangement of acetaldehyde relative to the ThDP-bound donor yielding (*S*)-2-HPP (Figure 18 A). The parallel orientation of acetaldehyde is still possible indicated by the *ee* of 92 % (*S*),

however, due to the better stabilization of the methyl side chain in the *S*-pocket the antiparallel orientation is preferred. Furthermore, the *S*-pocket of *Pp*BFD is just large enough to harbor the methyl side chain of acetaldehyde. Larger aldehydes, such as benzaldehyde, are arranged in parallel orientation again leading to the formation of the respective (*R*)-product (in this case benzoin with > 98 % *ee*) (Knoll *et al*., 2006; Gocke *et al*., 2008). In *Pf*BAL such an *S*-pocket is missing explaining its high *R*-selectivity for the formation of 2-HPP (*ee* > 99 %) (Figure 18 B) (Demir *et al*., 2001; Knoll *et al*., 2006).

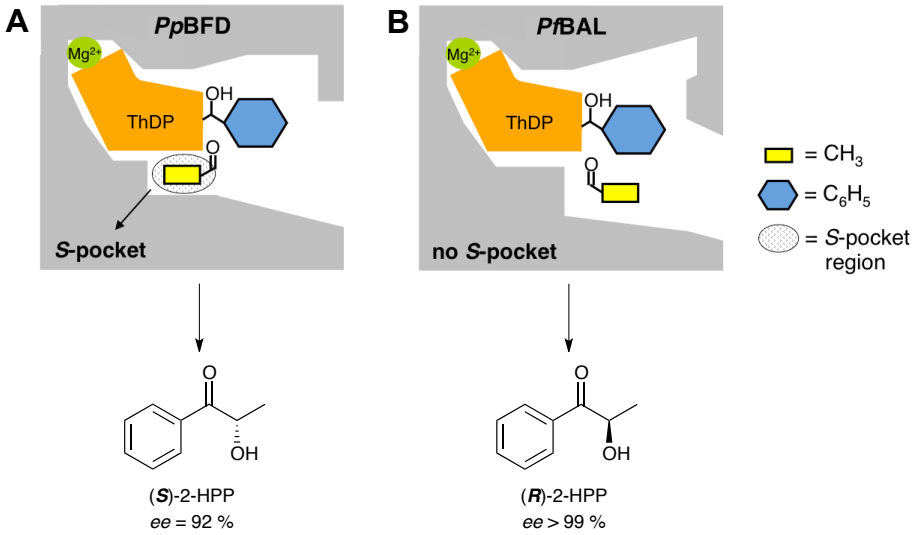


Figure 18: Principles of stereoselective carboligation in ThDP-dependent enzymes shown as schematic representation for the formation of 2-HPP (according to Knoll *et al*., 2006). Stereoselectivity depends on the relative orientation of donor and acceptor prior to carboligation. **A:** The *S*-pocket of *Pp*BFD allows the antiparallel orientation of acetaldehyde to the donor benzaldehyde yielding (*S*)-2-HPP. **B:** In *Pf*BAL only a parallel arrangement of donor and acceptor is possible due to a missing *S*-pocket that results in the formation of (*R*)-2-HPP.

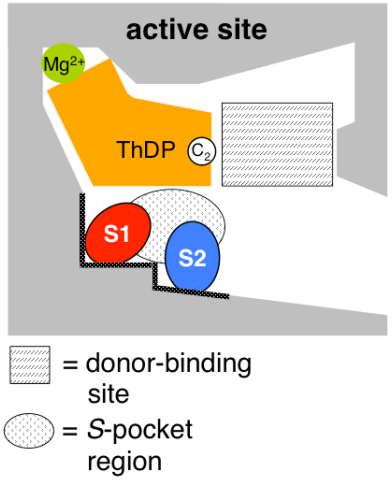
In ThDP-dependent enzymes of the DC superfamily the stereocontrol is only strict in presence of at least one aromatic substrate. In the case of two aliphatic substrates only moderate *ee*s could be obtained (Dominguez de Maria *et al*., 2007; Gocke *et al*., 2009; Gerhards *et al*., 2012), which is consistent with the fact that optimal stabilization of donor and acceptor is required for high stereoselectivity (Pohl *et al*., 2009; Hailes *et al*., 2013).

1.7.5 Engineering stereoselectivity – Access to (*S*)-α-hydroxy ketones

Enzyme engineering

The development of the *S*-pocket concept paved the way for the rational design of new ThDP-dependent enzyme variants for *S*-selective C–C bond formations and thus, the expansion of the α-hydroxy ketone platform. A comprehensive analysis on a structural level revealed potential *S*-pockets in many *R*-selective ThDP-dependent enzymes of the DC superfamily,

such as *Ll*KdcA and *Ap*PDC. However, bulky amino acid side chains block these pockets making them inaccessible for the respective acceptor aldehydes (Gocke, 2007; Rother *et al.*, 2011). Thereby, particularly the protein backbone (which cannot be mutated) and mainly two positions/residues in the *S*-pocket region define the *S*-pocket size (Figure 19).



Enzyme	Involved amino acids
	S-pocket S1
*Pp*BFD	L461
*Ap*PDC	E469
*Ll*KdcA	E462
	S-pocket entrance S2
*Pp*BFD	A460
*Ap*PDC	I468
*Ll*KdcA	V461

Figure 19: The typical *S*-pocket architecture in ThDP-dependent enzymes of the DC superfamily shown as schematic representation (according to Gocke, 2007; Gocke *et al.*, 2008). In addition to the protein backbone (black line), mainly two amino acid residues define the *S*-pocket size: S1 (red) in the pocket and S2 (blue) at the *S*-pocket entrance. These residues are given for *Pp*BFD, *Ap*PDC, and *Ll*KdcA. *Pf*BAL has no *S*-pocket due to the protein backbone close to the cofactor ThDP.

This structural information was used to modify the *S*-pockets of *Pp*BFD and *Ap*PDC by protein engineering and to support the developed *S*-pocket concept.

*Pp*BFD-L461A/G

The small *S*-pocket of *Pp*BFD only accepts the methyl side chain of acetaldehyde yielding (*S*)-2-HPP with 92 % *ee*. This pocket was enlarged by mutation of the bulky leucine in position 461 (position S1 in Figure 19) to the smallest amino acids alanine and glycine, respectively, in order increase the *S*-pocket for larger aliphatic aldehydes. As a result, the *ee* for 2-HPP was improved to 98 % *ee* (*S*) by both variants L461A and L461G, respectively. Moreover, the enlarged *S*-pockets were also able to properly stabilize the side chains of propanal and methoxyacetaldehyde yielding the respective (*S*)-2-HPP derivatives (*ee* = 93–97 %) for the first time. Even pentanal was accepted by *Pp*BFD-L461A in the mixed carboligation with 3,5-dimethoxybenzaldehyde as donor giving access to the corresponding (*S*)-2-HPP derivative with 90 % *ee* (Gocke *et al.*, 2008). However, both variants still converted benzaldehyde highly *R*-selectively to benzoin (*ee* > 97 %) due to the restriction of the *S*-pocket size by the protein backbone, which does not offer sufficient space for the phenyl ring of benzaldehyde to bind in antiparallel orientation.

*Ap*PDC-E469G

With respect to the protein backbone, *Ap*PDC provides the largest potential *S*-pocket among yet structurally investigated ThDP-dependent enzymes of the DC superfamily (Rother *et al.*, 2011). The reason for its high *R*-selectivity in the mixed carboligation of acetaldehyde and benzaldehyde is the preference for acetaldehyde as the donor (chapter 1.7.4). The large benzaldehyde acceptor is not able to enter the *S*-pocket (Figure 20 A) due to sterical hindrance with a bulky glutamate in position 469 (position S1 in Figure 19). In contrast, acetaldehyde as acceptor is able to bind in antiparallel orientation indicated by the formation of (*S*)-acetoin (*ee* = 28 %) in the self-ligation of acetaldehyde (Gerhards *et al*., 2012).

Mutation of E469 to glycine opened the *S*-pocket of *Ap*PDC yielding sufficient space for good stabilization of the antiparallel-oriented acceptor benzaldehyde (Figure 20 B). As a consequence, the longstanding problem of *S*-selective PAC formation could be solved paving the way to (*S*)-PAC and derivatives thereof with 85-89 % *ee*. In addition, the formation of (*S*)-acetoin (*ee* = 85 %) was catalyzed with improved stereoselectivity (Rother *et al*., 2011; Gerhards *et al*., 2012).

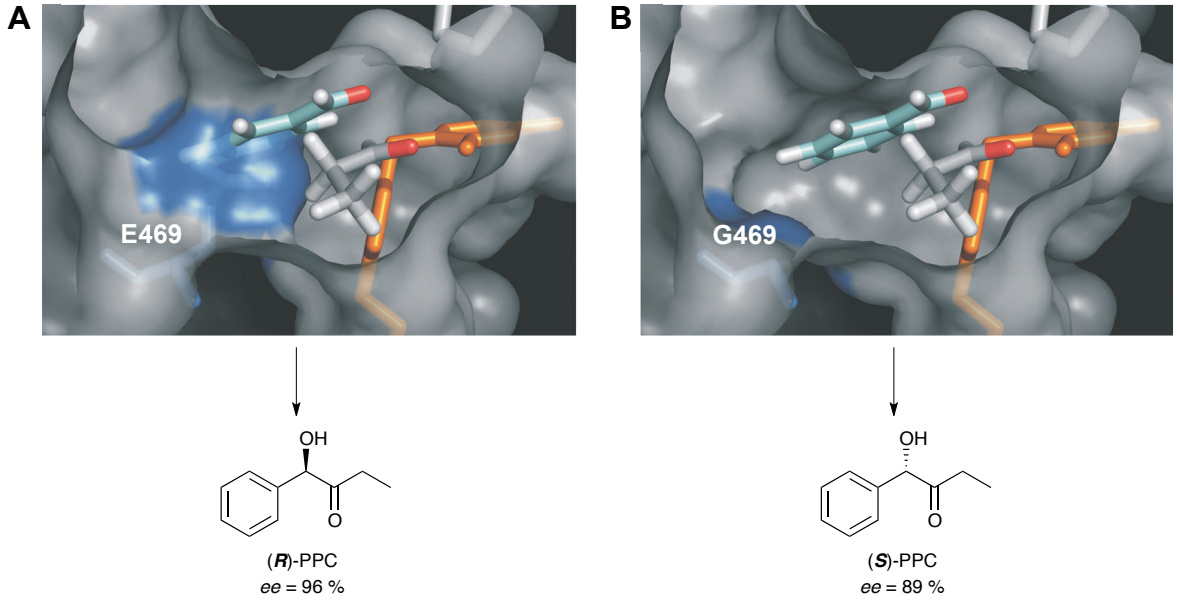


Figure 20: The mixed carboligation of propanal (gray) and benzaldehyde (cyan) to phenylpropionylcarbinol (PPC) catalyzed by *Ap*PDC (**A**) and *Ap*PDC-E469G (**B**) (Rother *et al.*, 2011). **A:** The *S*-pocket of wild-type *Ap*PDC is blocked by glutamate (blue) in position 469, which prevents the antiparallel orientation of benzaldehyde yielding the respective (*R*)-PPC with high *ee*. **B:** By mutation of E469 to glycine the *S*-pocket was opened allowing the stabilization of antiparallel-oriented benzaldehyde and thus the formation of (*S*)-PPC for the first time. ThDP is colored orange.

*Ap*PDC-W388A – A possible alternative *S*-pathway

Apart from the *S*-pocket, there is, at least theoretically, a second binding site for the acceptor in order to obtain (*S*)-α-hydroxy ketones in the mixed carboligation. A backside attack of a

parallel-oriented acceptor on the ThDP-bound donor prior to carboligation would also result in the formation of (*S*)-products (Figure 21). This alternative *S*-pathway was recently proposed for *Ap*PDC (Rother *et al.*, 2011). A bulky tryptophan residue in position 488 was exchanged to alanine in order to improve carboligation activity for the production of (*R*)-PAC and derivatives thereof. Indeed, *Ap*PDC-W388A revealed enhanced catalytic activity. However, this mutation also impaired the stereoselectivity of the carboligation, although the *S*-pocket was closed (Figure 21 B). *R*-Selectivity significantly dropped to 84 % *ee* and 50 % *ee* in the case of (*R*)-PAC and (*R*)-phenylpropionylcarbinol (PPC) formation, respectively. Stereoselectivity was even inverted using butanal as donor yielding the respective (*S*)-PAC derivative with 29 % *ee* (Rother *et al.*, 2011). The exchange of W488 to alanine possibly opened up the alternative acceptor-binding site, which might allow the acceptor benzaldehyde to approach the ThDP-bound donor from the backside. The investigation of this possible alternative *S*-pathway was part of this thesis.

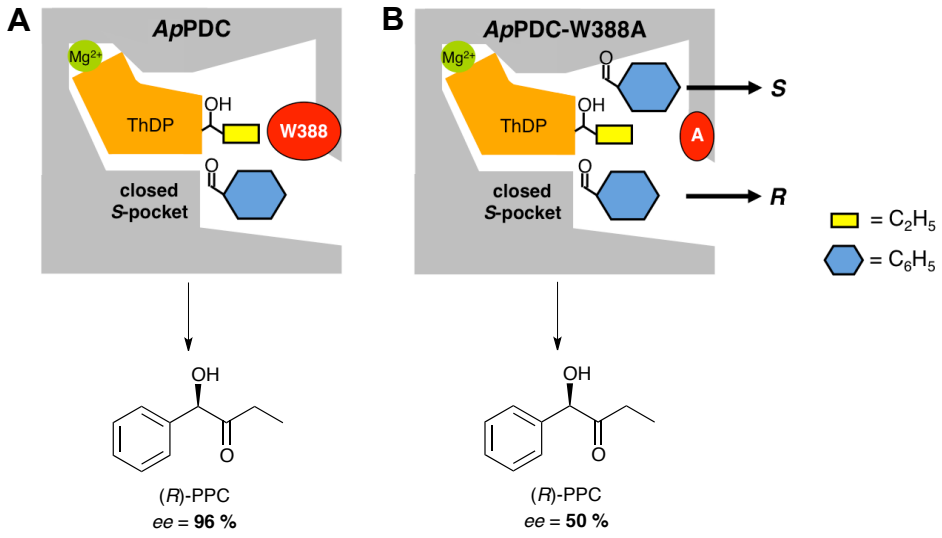


Figure 21: The hypothesis of an alternative *S*-pathway in *Ap*PDC shown as schematic representation (according to Rother *et al.*, 2011). **A:** The carboligation of propanal and benzaldehyde is catalyzed highly *R*-selective by wild-type *Ap*PDC due to a closed *S*-pocket. **B:** Although the *S*-pocket is still closed in *Ap*PDC-W388A, PPC is formed with low *R*-selectivity. The exchange of W388 to alanine opened possibly an alternative *S*-pathway that allows the parallel orientation of benzaldehyde to the ThDP-bound donor, which in turn results in the formation of the respective (*S*)-product.

Solvent engineering

Apart from protein engineering, stereoselectivity of ThDP-dependent enzymes can also be influenced by reaction conditions, *e.g.* the addition of cosolvents. Cosolvents are often used as additives in buffered systems in order to enhance the substrate solubility of aromatic compounds and thus increase the substrate load. A prominent cosolvent for ThDP-dependent enzyme-catalyzed carboligations is dimethylsulfoxide (DMSO). DMSO was shown to

improve substrate solubility on one side, but also positively affects the stability of *Pf*BAL during biotransformations (Demir *et al*., 2001). However, recent studies revealed that the addition of cosolvents is not always advantageous (Gerhards *et al*., 2012). For example, the influence of organic cosolvents on the stereoselectivity of the *S*-selective *Ap*PDC-E469G was immense, resulting in impaired *S*-selectivity in the presence of 13 different cosolvents. Whereas, (*S*)-PAC was formed with 87 % *ee* in aqueous buffer (50 mM triethanolamine), the addition of *e.g.* 20 vol% DMSO or 0.5 vol% trichloromethane yielded (*S*)-PAC with a reduced *ee* of 65 % and even (*R*)-PAC with 49 % *ee*, respectively. The results revealed that the smaller the solvent-excluded volume of the organic solvent is, the higher was the impact on stereoselectivity. In addition, a correlation of the solvent's logP[*]-value and stereoselectivity was observed. It is assumed that the cosolvent is able to directly interact with the active site, especially with the *S*-pocket. As a consequence, the *S*-pocket might be blocked by the cosolvent, which in turn reduces the *S*-selectivity of *Ap*PDC-E469G (Gerhards, 2012; Gerhards *et al*., 2012).

These results demonstrate that solvent engineering is an additional tool to alter stereoselectivity of ThDP-dependent enzymes; they further demonstrate that solvent addition has to be carefully considered and evaluated especially for *S*-selective carboligations.

Reaction engineering

In addition to cosolvents, a variety of other reaction parameter usually affects activity, stability and selectivity of biocatalysts, *e.g*. temperature, pH, and substrate concentration. These parameters are often interconnected leading to time- and material-consuming optimization processes (*e.g.* design of experiment-supported multi-parameter optimization). Such process optimization was recently applied on *Ap*PDC-E469G for the production of (*S*)-PPC (Gómez Baraibar *et al.*, 2013). Here, the initially moderate *ee* of 89 % could be improved to 98 % *ee* accompanied by a 61-fold increase of the specific space-time-yields compared to the starting point. Thereby, the highest impact on stereoselectivity as well as on productivity was obtained by altering the donor substrate from propanal to the corresponding α-keto acid ketobutyrate. Consequently, reaction engineering complements the tools for the design of highly *S*-selective ThDP-dependent enzymes for the formation of α-hydroxy ketones.

[*]logP describes the ratio of concentrations of a compound in an immiscible water/octanol mixture at equilibrium.

1.8 Scope and objectives

As part of the DFG research unit FOR1296 “Diversity of Asymmetric Thiamine Catalysis” the general objective of this thesis was a comprehensive elucidation of principles that control chemo- and especially stereoselectivity in ThDP-dependent enzymes of the DC superfamily. A thorough understanding of the structure-function relationships should allow the targeted modulation of substrate specificity and stereoselectivity of ThDP-dependent enzymes, which in turn could pave the way for new asymmetric syntheses that are currently not yet feasible by classical approaches. Here, a special focus was on yet hardly accessible (*S*)-α-hydroxy ketones.

One goal of this thesis was the further development and refinement of structure-function relationships in ThDP-dependent enzymes in order to improve the prediction quality with respect to mixed carboligation reactions, to overcome the limited *S*-selectivity, and to accelerate the development of ThDP-dependent enzymes with novel catalytic activities. The following three main aspects should be addressed:

- **Transfer and validation of the existing *S*-pocket concept**
  The transfer of the principles of stereoselectivity to *Ec*AHAS and MenD should prove the general validity of the previously developed *S*-pocket concept for ThDP-dependent enzymes of the DC superfamily. In contrast to *Pp*BFD and *Ap*PDC, which belong to the group of lyases (EC 4), AHAS and MenD are transferases (EC 2). Moreover, both enzymes require *in situ* decarboxylation of an α-keto acid donor prior to carboligation with an aldehyde acceptor, which is different to most other decarboxylases. The reason is not yet known.

- **Refinement and extension of the *S*-selectivity concept**
  Up to now, only one specific position in the *S*-pocket was targeted by site-directed mutagenesis resulting in *S*-selective variants of *Pp*BFD and *Ap*PDC. These variants catalyze the mixed carboligation of aromatic and aliphatic aldehydes in most cases only with moderate stereoselectivity (*ee* < 90 %). To meet the needs of the fine chemical market, the stereoselectivity must be improved. Therefore, the active site should be screened for further “hot spots” that directly influence stereoselectivity. Here, the site-directed mutagenesis approach to generate single or double variants should be combined with random mutagenesis of selected hot spots. In this regard, a yet not available fast and reliable screening process for ThDP-dependent enzymes had to be established.

- **Investigation of a potential alternative *S*-pathway**
  Earlier studies with *Ap*PDC indicated a potential alternative *S*-pathway that might enable access to novel (*S*)-α-hydroxy ketones. This hypothesis should be tested with *Pf*BAL, the only known ThDP-dependent enzyme lacking an *S*-pocket. Thus, (*S*)-products may only originate via the alternative *S*-pathway.

Further, one main goal of this thesis was the extension of the ThDP-dependent enzyme toolbox in order to expand the α-hydroxy ketone platform. In this context, special focus was on:

- **Functionalized (*S*)-α-hydroxy ketones:** Mixed araliphatic products with functionalized aliphatic side chains, such as 5-hydroxy-4-oxo-5-arylpentanoates, are not accessible so far and might be available by new *S*-selective MenD variants, which catalyze the carboligation of α-ketoglutarate and benzaldehyde derivatives.
- **(*S*)-Benzoins:** There was yet no direct access to (*S*)-benzoins via enzymatic asymmetric synthesis starting from benzaldehyde (derivatives). If successful, new BAL variants originated from the studies of the alternative *S*-pathway could close this gap, which in turn would complement the existing α-hydroxy ketone platform enabling the enantiocomplementary synthesis of all possible α-hydroxy ketone isomers starting from cheap acetaldehyde and benzaldehyde (derivatives) (chapter 1.7.4, Figure 16).

2 PUBLICATIONS

PUBLICATION I

Two Steps in One Pot: Enzyme Cascade for the Synthesis of Nor(pseudo)ephedrine from Inexpensive Starting Materials

Sehl, T.; Hailes, H. C.; Ward, J. M.; Wardenga, R.; von Lieres, E.;
Offermann, H.; Westphal, R.; Pohl, M. & Rother, D.

Angewandte Chemie International Edition 2013,
52, 6772–6775
DOI: 10.1002/anie.201300718
(http://onlinelibrary.wiley.com/doi/10.1002/anie.201300718/full)

Angewandte Communications

Biocatalysis

DOI: 10.1002/anie.201300718

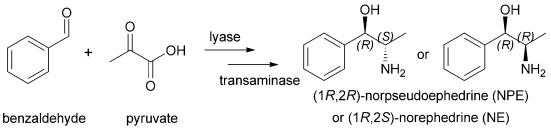
Two Steps in One Pot: Enzyme Cascade for the Synthesis of Nor(pseudo)ephedrine from Inexpensive Starting Materials**

*Torsten Sehl, Helen C. Hailes, John M. Ward, Rainer Wardenga, Eric von Lieres, Heike Offermann, Robert Westphal, Martina Pohl, and Dörte Rother**

A challenging task for chemical researchers in the next decade is the development of cleaner and more environmentally friendly reactions.[1] The traditional chemical syntheses of enantiomerically pure compounds often require multistep protocols with protection–deprotection steps as well as the isolation of potentially unstable intermediates, lowering the yields and sustainability of the overall process.[2]

Phenylpropanolamines, members of the amphetamine family of ephedra alkaloids, are compounds with multifunctional applications but challenging syntheses routes. The stereoisomers norpseudoephedrine (NPE) and norephedrine (NE) are used as building blocks for the preparation of ligands and chiral auxiliaries in organic syntheses[3] and also have direct applications as pharmaceutically active molecules.[4] Reported synthetic approaches to these compounds have disadvantages such as relatively expensive reagents, multistep preparative routes, and only moderate enantio- and diastereoselectivity.[5] Recently, a novel highly stereoselective method was described for the synthesis of all phenylpropanolamine isomers with *ee* and *de* values exceeding 99%.[6] Norephedrine isomers were accessible in four steps (40% yield) and norpseudoephedrine in seven steps (35% yield) starting from 2-phenyl-2-trimethylsilyloxyacetonitrile.

Synthetic enzyme cascades are valuable alternative routes for the stereoselective production of fine chemicals. Since the chemo- and stereoselectivities are typically high, the isolation of by-products and reaction intermediates can be circumvented[7] and thus the eco-efficiency increased.[8] Here we present an enzymatic one-pot two-step reaction for the synthesis of stereomerically pure (1*R*,2*S*)-NE and (1*R*,2*R*)-NPE from benzaldehyde and pyruvate (Scheme 1). A number of different ways to perform enzyme cascade reactions have already been described (for more details see Chapter 1 in the Supporting Information).[1b,2a,8a–c,9] Our one-pot two-step reaction combines many advantages of known synthesis strategies like high stereoselectivities, inexpensive starting materials, high step economy (only two steps), and an equilibrium shift without addition of further enzymes or cosubstrates.



Scheme 1. One-pot two-step reaction for the synthesis of norpseudoephedrine (NPE) and norephedrine (NE).

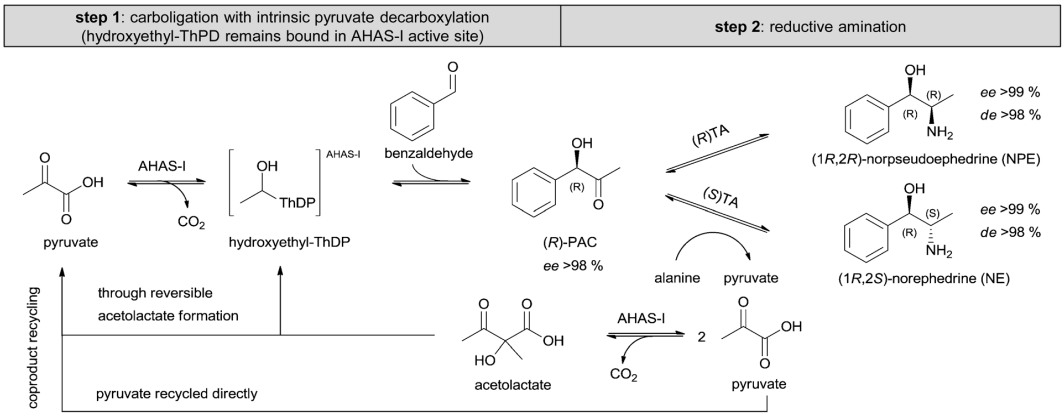
In the first step pyruvate is decarboxylated and subsequently ligated to benzaldehyde yielding (*R*)-phenylacetylcarbinol ((*R*)-PAC). The reaction is catalyzed by the thiamine diphosphate (ThDP)-dependent acetohydroxyacid synthase I (AHAS-I) from *E. coli* which performs the decarboxylation of pyruvate and the subsequent carboligation without releasing the hydroxyethyl-ThDP (see Scheme 2).[10] (*R*)-PAC is obtained with high stereoselectivity ($ee > 98\%$) and can be converted directly to the desired (1*R*,2*S*)-NE and to (1*R*,2*R*)-NPE in the second step of the cascade (reductive amination) by selectively using (*S*)- and (*R*)-selective ω-transaminases (TAs), respectively. In our previous work a set of 18 different (*S*)-selective wild-type (*S*)TAs had been screened for the conversion of 2-hydroxy ketones.[11] For the reductive amination of (*R*)-PAC, the *Cv*-(*S*)TA from *Chromobacterium violaceum* gave the most promising results. To gain access to (1*R*,2*R*)-NPE, seven different (*R*)-selective (*R*)TAs from Enzymicals AG (see Chapter 2 in the Supporting Information) were tested.

The enzymatic reductive amination requires an amine donor as a cosubstrate. Through the clever combination of cosubstrates (here: alanine) and enzymes, the resulting by-product (here: pyruvate) of the second reaction step can serve as the substrate for the first step. This novel type of cascade design is referred to as a "recycling cascade" (Scheme 2 and Chapter 1 in the Supporting Information). We determined the

[*] M. Sc. T. Sehl, Dr. E. von Lieres, H. Offermann, Dipl.-Biotechnol. R. Westphal, Prof. Dr. M. Pohl, Dr. D. Rother
Institute of Bio- and Geosciences
IBG-1: Biotechnology, Forschungszentrum Jülich GmbH
Leo-Brandt-Strasse 1, 52425 Jülich, (Germany)
E-mail: do.rother@fz-juelich.de
Prof. H. C. Hailes
Department of Chemistry, University College London
20 Gordon Street, London WC1H OAJ (UK)
Prof. J. M. Ward
The Advanced Centre for Biochemical Engineering
Department of Biochemical Engineering, University College London
Torrington Place, London, WC1E 7JE (UK)
Dr. R. Wardenga
Enzymicals AG
Walther-Rathenau-Strasse 49a, 17489 Greifswald (Germany)

[**] This work was supported by the CLIB Graduate Cluster Industrial Biotechnology of the Heinrich-Heine-University Düsseldorf and financed by the DFG (German Research Foundation) in the frame of Research Group FOR 1296.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201300718.



Scheme 2. One-pot two-step reaction for the synthesis of nor(pseudo)ephedrine performed as a recycling cascade combining acetohydroxyacid synthase I (AHAS-I) and a (*S*)- or (*R*)-selective ω-transaminases ((*R*)TA, (*S*)TA).

thermodynamic equilibrium constant for the reductive amination of PAC with alanine as the amine donor to be 2.31×10^{-3} (Chapter 4 in the Supporting Information). Consequently, when equimolar concentrations of alanine and PAC are used and the by-product pyruvate is not removed, a theoretical conversion of less than 5% is obtained (see Chapter 4.4 in the Supporting Information). In our reaction setup, pyruvate can be removed by two different carboligation reactions mediated by AHAS-I: 1) the carboligation with benzaldehyde yielding PAC or 2) a carboligation with another pyruvate molecule yielding acetolactate. The reversible reaction giving acetolactate is kinetically favored, whereas the reaction equilibrium lies on the side of PAC formation.[10] Thus, acetolacetate is a suitable substrate for the carboligation of (*R*)-PAC by the cleavage reaction to pyruvate and hydroxyethyl-ThDP.

A challenge in this one-pot two-step cascade is the fact that the starting material benzaldehyde might serve as a substrate for AHAS-I as well as for the ω-transaminases. As a consequence of the higher chemical reactivity of aldehydes relative to ketones and steric constraints in the active site of ω-TAs, it was not possible to find an enzyme among the 25 screened ω-TAs for which the reductive amination of PAC was kinetically favored over the reductive amination of benzaldehyde. The most promising (*S*)-selective transaminase *Cv*-(*S*)TA has a roughly 17-fold higher initial rate in the reaction with benzaldehyde than with PAC (Figure 1 A). As a consequence, in a one-pot two-step cascade reaction where AHAS-I and *Cv*-(*S*)TA were added simultaneously, 98% of the benzaldehyde was converted to benzylamine (Figure 2 A). However, in the case of the (*R*)-selective ω-TAs we could surprisingly identify enzymes for which the simultaneous one-pot two-step reaction provided (1*R*,2*R*)-NPE with conversions of up to 85% (Figure 2 A). The initial rate activities of the ω-TA from *Aspergillus terreus* (*At*-(*R*)TA) for reactions with PAC and benzaldehyde were on the same order of magnitude (Figure 1 B), but roughly ten times lower than the initial rates for the (*R*)-PAC formation

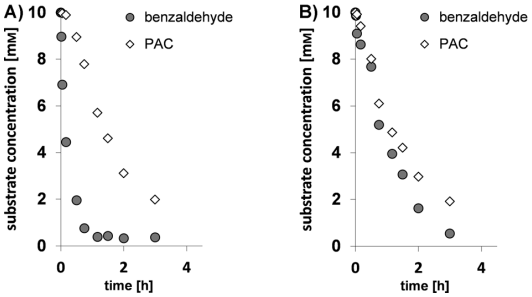


Figure 1. Conversion curves for the reductive amination of benzaldehyde (10 mM) and PAC (10 mM) by A) ω-TA *Cv*-(*S*)TA (1 mg mL^{-1}) and by B) *At*-(*R*)TA (1 mg mL^{-1}), respectively. The reaction was carried out in 100 mM HEPES (pH 7.5 with 200 μM pyridoxal-5′-phosphate (PLP), 50 μM flavine adenine dinucleotide (FAD), 100 μM ThDP, 5 mM $MgCl_2$) containing (*S*)- or (*R*)-α-methylbenzylamine (10 mM) as amine donor.

catalyzed by AHAS-I. These differences suffice to reduce the amount of formed by-product (benzylamine) to merely 10%.

In line with these experimental data, the NE/benzylamine ratio is low when the one-pot two-step reaction is performed as a simultaneous cascade including a recycling step. Here, both enzymes were added simultaneously to a mixture of 20 mM benzaldehyde, 10 mM pyruvate, and 50 mM alanine. Since no further pyruvate was added, product concentrations higher than 10 mM (NE or NPE) are only possible by the successful recycling of pyruvate that is generated by deamination of alanine. Remarkably, with *At*-(*R*)TA about 14 mM (1*R*,2*R*)-NPE and only 5.5 mM benzylamine were formed in this simultaneous recycling mode, while in case of the *Cv*-(*S*)TA the major product is benzylamine.

Although benzylamine can be separated from NPE and NE by column chromatography (mobile phase EtOAc/MeOH/NH_3 = 85:10:5), it is more advantageous to reduce

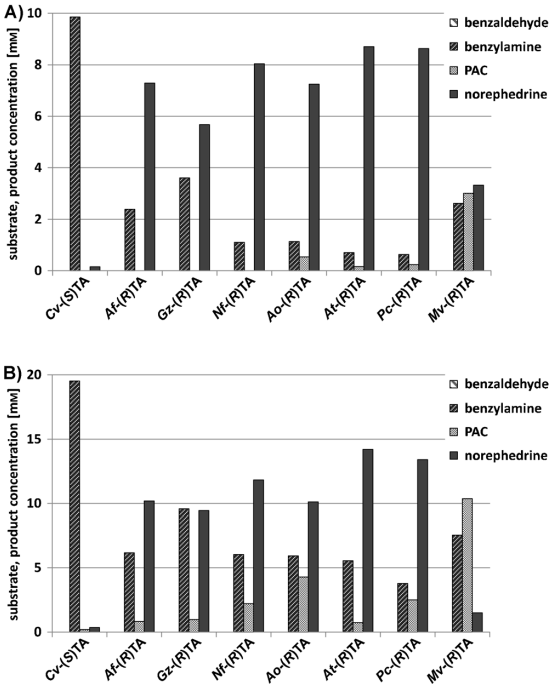


Figure 2. A) One-pot two-step simultaneous cascade and B) one-pot two-step simultaneously recycling cascade with *Cv*-(*S*)TA and seven (*R*)-selective TAs. For the one-pot two-step reaction 10 mM benzaldehyde, 10 mM pyruvate, and 50 mM D- or L-alanine were dissolved in 100 mM HEPES (pH 7.5 with 200 μM PLP, 50 μM FAD, 100 μM ThDP, 5 mM $MgCl_2$) and the two enzymes (AHAS-I and ω-TA) were added simultaneously. The recycling cascade reactions (B) were performed analogously with 20 mM benzaldehyde instead of 10 mM.

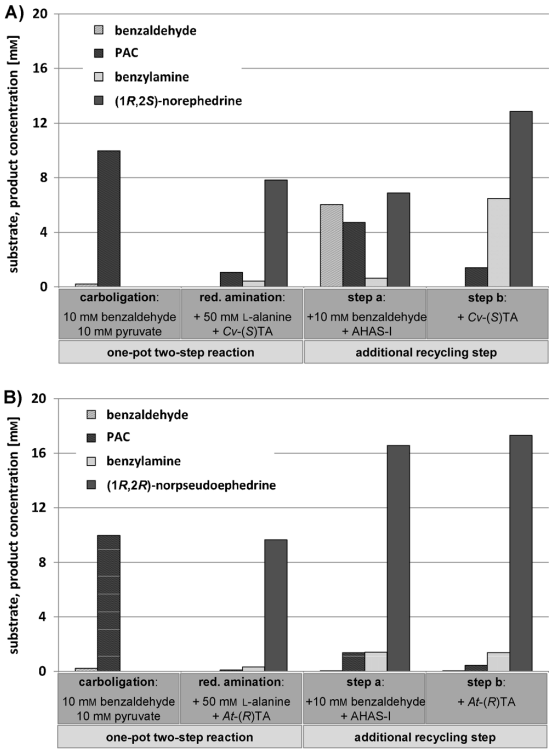


Figure 3. Synthesis of A) (1*R*,2*S*)-NE and B) (1*R*,2*R*)-NPE performed as a sequential one-pot two-step reaction with an additional recycling step (time-dependent reaction curve can be found in Chapter 6 in the Supporting Information). Reaction conditions: 100 mM HEPES (pH 7.5 with 200 μM PLP, 50 μM FAD, 100 μM ThDP, 5 mM $MgCl_2$), 25 °C, 100 rpm. One-pot two-step reaction: Carboligation (1 h): 10 mM benzaldehyde, 10 mM pyruvate, 0.5 mg mL^{-1} AHAS; reductive amination (12 h): +50 mM alanine, +1 mg mL^{-1} TA. Recycling step:step a: +10 mM benzaldehyde, +0.5 mg mL^{-1} AHAS-I (A: 1.5 h, B: 5 h); step b: +1 mg mL^{-1} TA (A: *Cv*-(*S*)TA, 12 h, B: *At*-(*R*)TA, 5 h).

the formation of byproducts by appropriate process engineering in order to increase process efficiency. There are two general ways to perform a cascade reaction: one is the already described simultaneous mode, the other one the so-called sequential mode, where the catalysts are added consecutively.[9] In our sequential synthetic enzyme cascade, the limiting step is the reductive amination. In order to circumvent this bottleneck, we optimized the reaction parameters of the reductive amination step regarding pH, temperature, the concentrations of transaminase and AHAS-I, and the amine donor/PAC ratio (see Chapter 5 in the Supporting Information). For the enzyme combination *Cv*-(*S*)TA/AHAS-I conversions exceeding 80% could be achieved under optimized cascade conditions (pH 7.5, 25 °C, 1 mg mL^{-1} *Cv*-(*S*)TA, 0.5 mg mL^{-1} AHAS-I, alanine/PAC = 5:1).

These optimized conditions were applied in the one-pot two-step sequential mode. Here, the transaminase was added after the benzaldehyde had been completely consumed in the AHAS-I-catalyzed carboligation step (after 1 h 100% conversion was achieved, Figure 3A). This increased the conversion of (1*R*,2*S*)-NE from 2% (Figure 2A: one-pot two-step simultaneous cascade) to 78% (7.8 mM, Figure 3A) with the combination AHAS-I/*Cv*-(*S*)TA. Under these conditions the undesired by-product benzylamine amounted to less than 0.5 mM (Figure 3A). Upon subsequent addition of further 10 mM of benzaldehyde and fresh AHAS-I, PAC was formed in a second carboligation step. Since no further pyruvate was added, this result demonstrates that the recycling of pyruvate, generated by deamination of alanine, was successful in this sequential enzyme recycling cascade mode. However, the reaction resulted in the formation of only 4.7 mM PAC (47% conversion), which is most likely due to the instability of acetolactate. If the latter is chemically decarboxylated to acetoin, it is no longer available for PAC formation. Moreover, acetoin (and probably also acetolactate) can act as substrates for *Cv*-(*S*)TA as described previously.[11] Further, neither the NE nor the benzylamine concentration increased significantly, which suggests almost complete inactivation of *Cv*-(*S*)TA (Figure 3A). Addition of fresh *Cv*-(*S*)TA started the reaction again yielding 12.9 mM (1*R*,2*S*)-NE ($de > 98\%$, $ee > 99\%$). This corresponds to roughly 65% of the possible

Angewandte International Edition Chemie

product concentration (related to a total benzaldehyde concentration of 20 mM, Figure 3A).

In the case of *At*-(*R*)TA the cascade reaction is even more efficient (Figure 3B): (1*R*,2*R*)-NPE is accessible in the one-pot two-step sequential cascade with conversions greater than 96% and very high stereomeric purity ($de > 98\%$, $ee > 99\%$). After addition of another 10 mM benzaldehyde and fresh AHAS-I, the *At*-(*R*)TA was still active. Thus, without addition of further transaminase, 16.6 mM (1*R*,2*R*)-NPE (83% conversion) was obtained in 5 h by the complete one-pot two-step recycling cascade. Further addition of *At*-(*R*)TA did not considerably increase the final product concentration.

In summary, we have developed a strategy for the synthesis of (1*R*,2*S*)-NE and (1*R*,2*R*)-NPE. Both compounds are accessible in a biocatalytic one-pot two-step reaction in high stereoisomeric purity ($de > 98\%$, $ee > 99\%$) from inexpensive starting materials without isolation of the intermediate product. Additionally, these cascade reactions can be performed in a novel "recycling mode" in which the coproduct of the second step is removed without addition of further catalysts or cosubstrates and recycled as a substrate for the first cascade step.

By combining reaction and process optimization, the sequential cascade consisting of AHAS-I and *Cv*-(*S*)TA provided (1*R*,2*S*)-NE with a conversion of 80% (8.0 mM). For the production of (1*R*,2*R*)-NPE we could identify (*R*)-selective ω-TAs catalyzing the one-pot two-step cascade in a simultaneous mode with conversions up to 85% (8.5 mM). Moreover, in the sequential mode, formation of the side product (benzylamine) is reduced and (1*R*,2*R*)-NPE was obtained with a conversion exceeding 96% within 13 h. In the recycling step (addition of another 10 mM benzaldehyde, but no pyruvate) a total concentration of 16.6 mM (1*R*,2*R*)-NPE (83% conversion) was observed after further 5 h reaction time.

The recycling mode can be applied to any set of reactions for which a clever combination of cosubstrates and coproducts is possible, such that the coproducts of one reaction can be reused as substrates for the other. This approach optimizes the atom economy of the reaction by reducing the waste production.

Experimental Section

L-alanine (Merck), D-alanine (Sigma Aldrich), and pyruvate (Sigma–Aldrich) were of >99% purity. Benzaldehyde (Sigma–Aldrich) was freshly distilled before use. The preparation of the catalysts *Cv*-(*S*)TA and AHAS-I is described in the supporting information. (*R*)-selective TAs are commercially available from Enzymicals AG (Germany) as lyophilized crude cell extracts. Descriptions of the reaction details, reaction analysis, cascade optimizations, and equilibrium determination can be found in the Supporting Information.

Received: January 27, 2013
Published online: May 9, 2013

Keywords: asymmetric synthesis · biocatalysis · enzyme cascades · phenylpropanolamine · ω-transaminase

[1] a) F. Cavani, *J. Chem. Technol. Biotechnol.* **2010**, *85*, 1175–1183; b) E. Ricca, B. Brucher, J. H. Schrittwieser, *Adv. Synth. Catal.* **2011**, *353*, 2239–2262.
[2] a) S. Shanmuganathan, D. Natalia, L. Greiner, P. Dominguez de Maria, *Green Chem.* **2012**, *14*, 94–97; b) M. J. Climent, A. Corma, S. Iborra, M. Mifsud, A. Velty, *Green Chem.* **2010**, *12*, 99–107.
[3] a) M. Sherman, *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*, Wiley, Hoboken, **2010**, ISBN: 9780470699584; b) G. Blay, I. Fernández, B. Monje, J. R. Pedro, *Tetrahedron Lett.* **2004**, *45*, 8039–8042; c) F. K. Cheung, A. M. Hayes, J. Hannedouche, A. S. Y. Yim, M. Wills, *J. Org. Chem.* **2005**, *70*, 3188–3197; d) D. A. Evans, M. D. Ennis, T. Le, N. Mandel, G. Mandel, *J. Am. Chem. Soc.* **1984**, *106*, 1154–1156; e) J. L. García Ruano, C. Alemparte, M. T. Aranda, M. M. Zarzuelo, *Org. Lett.* **2003**, *5*, 75–78; f) R. Guo, A. J. Lough, R. H. Morris, D. Song, *Organometallics* **2004**, *23*, 5524–5529; g) S. R. Hitchcock, D. M. Casper, J. F. Vaughn, J. M. Finefield, G. M. Ferrence, J. M. Esken, *J. Org. Chem.* **2004**, *69*, 714–718; h) C. Unaleroglu, A. E. Aydin, A. S. Demir, *Tetrahedron: Asymmetry* **2006**, *17*, 742–749.
[4] a) N. A. Flavahan, *J. Pharmacol. Exp. Ther.* **2005**, *313*, 432–439; b) R. B. Rothman, N. Vu, J. S. Partilla, B. L. Roth, S. J. Hufeisen, B. A. Compton-Toth, J. Birkes, R. Young, R. A. Glennon, *J. Pharmacol. Exp. Ther.* **2003**, *307*, 138–145; c) R. Krizevski, N. Dudai, E. Bar, E. Lewinsohn, *J. Ethnopharmacol.* **2007**, *114*, 432–438.
[5] J. A. Groeper, S. R. Hitchcock, G. M. Ferrence, *Tetrahedron: Asymmetry* **2006**, *17*, 2884–2889.
[6] H. K. Lee, S. Kang, E. B. Choi, *J Org Chem.* **2012**, *77*, 5454–5460.
[7] F. Lopez-Gallego, C. Schmidt-Dannert, *Curr. Opin. Chem. Biol.* **2010**, *14*, 174–183.
[8] a) A. Bruggink, R. Schoevaart, T. Kieboom, *Org. Process Res. Dev.* **2003**, *7*, 622–640; b) S. F. Mayer, W. Kroutil, K. Faber, *Chem. Soc. Rev.* **2001**, *30*, 332–339; c) J. H. Schrittwieser, J. Sattler, V. Resch, F. G. Mutti, W. Kroutil, *Curr. Opin. Chem. Biol.* **2011**, *15*, 249–256; d) R. A. Sheldon, *Green Chem.* **2007**, *9*, 1273–1283; e) R. A. Sheldon, *Chem. Commun.* **2008**, 3352–3365.
[9] C. V. Voss, C. C. Gruber, W. Kroutil, *Angew. Chem.* **2008**, *120*, 753–757; *Angew. Chem. Int. Ed.* **2008**, *47*, 741–745.
[10] S. Engel, M. Vyazmensky, S. Geresh, Z. Barak, D. M. Chipman, *Biotechnol. Bioeng.* **2003**, *83*, 833–840.
[11] T. Sehl, R. C. Simon, H. C. Hailes, J. M. Ward, U. Schell, M. Pohl, D. Rother, *J. Biotechnol.* **2012**, *159*, 188–194.

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Two Steps in One Pot: Enzyme Cascade for the Synthesis of Nor(pseudo)ephedrine from Inexpensive Starting Materials**

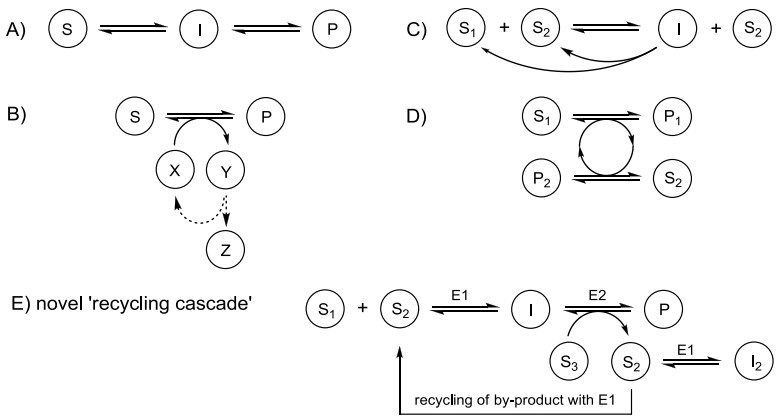
*Torsten Sehl, Helen C. Hailes, John M. Ward, Rainer Wardenga, Eric von Lieres, Heike Offermann, Robert Westphal, Martina Pohl, and Dörte Rother**

anie_201300718_sm_miscellaneous_information.pdf

Table of contents:

1.) Definition and explanation of different cascade reactions types

A number of different multi-enzyme cascade reactions have already been described.[1] In general, they can be subdivided into four types of cascade designs (Scheme 1, A-D).[1b] In the 'linear cascade' a single product is synthesized via one or more intermediates from a single substrate (Scheme 1A). In more complex enzyme cascades like the 'orthogonal cascade', product formation is coupled with a cofactor/co-substrate regeneration step or by-products removal reactions (Scheme 1B). Whereas in 'cyclic cascades' a mixture of substrates is transformed into an intermediate, which is subsequently converted back into starting material, so that one substrate is subsequently enriched (Scheme 1C). A typical example is the dynamic kinetic resolution of racemic compounds. In 'parallel cascade' reactions formation of two products by two biocatalytic reactions is coupled via their cofactors/co-substrates (Scheme 1D). Except for the linear cascade, all cascade modes require the introduction of further enzymes and/or co-substrates in order to allow cofactor regeneration or by-products removal to shift the reaction equilibrium. These additional components have to be separated from the desired products, which can lower the economy of the process.



Scheme 1: Four previously described enzymatic cascade designs:[1b] A) linear cascade, B) orthogonal cascade, C) cyclic cascade, D) parallel cascade and E) the novel recycling cascade (adapted from ref. [1a])

Here, we present a novel type of cascade reaction, the 'recycling cascade' (Scheme 1E), which was developed for a two-step synthesis of chiral 2-amino alcohols. In this synthetic enzyme cascade acetaldehyde (upon decarboxylation of pyruvate) is ligated to an aromatic aldehyde by a thiamine diphosphate (ThDP)-dependent enzyme (E1) yielding a 2-hydroxy ketone which is subsequently reduced to a 2-amino alcohol by an ω-transaminase (E2). The enzymatic reductive amination requires an amine donor as co-substrate. By a smart combination of this co-substrate and enzyme E1 the reaction runs because E1 has two functions in this 1-pot cascade: beside formation of the 2-hydroxy ketone (scheme 1, intermediate I) it is further used to shift the equilibrium of the transaminase reaction, which is hampered by a low equilibrium constant. This equilibrium shift is possible because the co-product, formed upon deamination of the co-substrate, can be recycled as a substrate for the carboligation step. This removal shifts the equilibrium towards the desired product.

2.) Catalyst preparation

2.1) Cloning, expression and purification of ω-transaminase *Cv*-(*S*)TA

Preparation of ω-TA: Cloning[2], expression[3], and preparation of the cell-free extract[4] of the N-terminal His-tagged ω-TA *Cv*-(*S*)TA were performed as previously described. Protein purification (column material: Ni-NTA superflow, Qiagen, equipment: ÄKTApurifier, GE Healthcare) was performed in 100 mM HEPES buffer (pH 7.5) containing 200 µM pyridoxal-5'-phosphate (PLP). After one washing step (buffer + 25 mM imidazole) the ω-TA was eluted (buffer + 250 mM imidazole) and the buffer exchanged to 10 mM HEPES (pH 7.5) and 200 µM PLP using a Sephadex G-25 column (GE-Healthcare) before lyophilization. The transaminase was obtained in a purity >90 % as a lyophilized powder with a protein content of 70 % (w/w).

2.2) Cloning, expression and purification of acetohydroxyacid synthase I (AHAS-I)

Preparation of AHAS-I: The His-tagged AHAS-I was cloned[5] and expressed[6] as described in literature. Cell disruption was performed analogously to the process for the ω-TA (see 2.1), except the buffer ('AHAS-buffer'), which contained 100 mM potassium phosphate (KPi, pH 7.6), 500 mM KCl and 10 µM flavine adenine dinucleotide (FAD). Enzyme purification was also performed as described for ω-TA (see chapter 1.1), but with AHAS-buffer: after an equilibration step (AHAS-buffer + 20 mM imidazole) and one washing step (AHAS-buffer + 50 mM imidazole), the elution of AHAS was performed with AHAS-buffer + 250 mM imidazole. Before lyophilization the buffer was exchanged to 25 mM KPi (pH 7.5) and 10 µM FAD by using a Sephadex G-25 column (GE-Healthcare). The purity of AHAS-I lyophilized powder was >90 % and had a protein content of 15 % (w/w).

2.3) (*R*)-selective ω-transaminases

A commercially available toolbox of seven different (*R*)-selective ω-transaminases was provided by Enzymicals AG as lyophilized crude cell extract.[7]

Table 1: (*R*)-selective ω-transaminases provided by Enzymicals AG

(R)-selective TA	*Catalog no.*	*organism origin*
Af-(*R*)TA:	ECS-ATA01	*Aspergillus fumigatus*
Gz-(*R*)TA:	ECS-ATA02	*Gibberella zeae*
Nf-(*R*)TA:	ECS-ATA03	*Neosartorya fischeri*
Ao-(*R*)TA:	ECS-ATA04	*Aspergillus oryzae*
At-(*R*)TA:	ECS-ATA05	*Aspergillus terreus*
Pc-(*R*)TA:	ECS-ATA06	*Penicillium chrysogenum*
Mv-(*R*)TA:	ECS-ATA07	*Mycobacterium vanbaalenii*

3.) Reaction analytics

3.1) Quantification of benzaldehyde and phenylacetylcarbinol (PAC)

The extraction and quantification of PAC (R_t = 8.7 min) and benzaldehyde (R_t = 17.1 min) were performed by achiral HPLC-analysis as described before.[4] A typical chromatogram is presented in diagram 1.

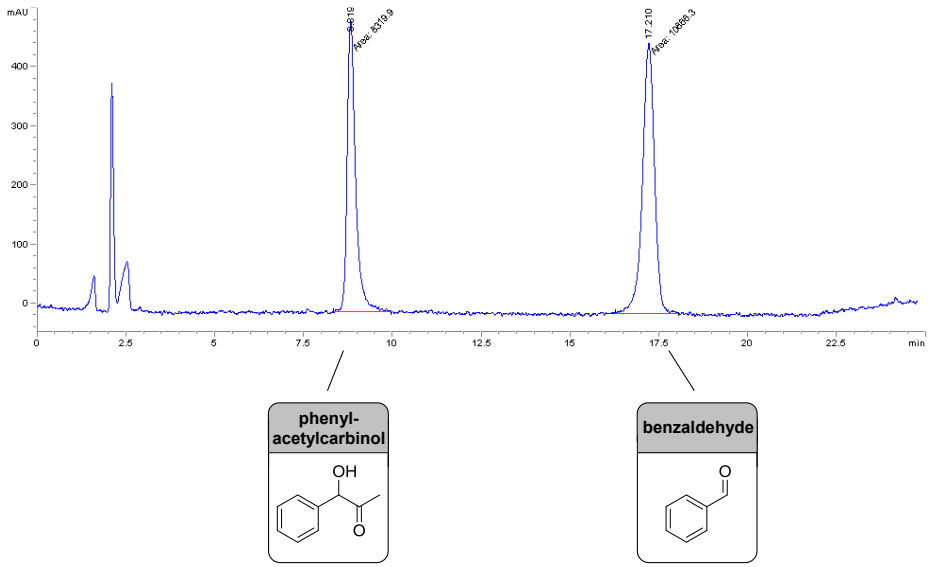


Diagram 1: Chromatogram of 10 mM benzaldehyde and 10 mM PAC on achiral HPLC analytics.[4]

3.2) Synthesis of reference compound: (1*S*,2*S*)-norpseudoephedrine

The enzymatic reductive amination of phenylacetylcarbinol (PAC) can theoretically result in four different stereoisomers of the product nor(pseudo)ephedrine: (1*S*,2*S*)-norpseudoephedrine, (1*R*,2*R*)-norpseudoephedrine, (1*S*,2*R*)-norephedrine, (1*R*,2*S*)-norephedrine. All isomers, except (1*S*,2*S*)-norpseudoephedrine were bought from Sigma Aldrich in a purity >99.5 %.

(1*S*,2*S*)-Norpseudoephedrine was synthesized starting from the (*S*)-PAC of a PAC mixture (with (*R*)-PAC having an *ee* = 84 %), which was provided by BASF AG as a yellow oil with a purity of >98 % (NMR, GC).[4] 10 mM (*R*)-PAC (141.1 mg, 0.939 mmol) and 10 mM (*S*)-α-methylbenzylamine were incubated with 1 mg/ml ω-TA (*Cv*-(*S*)TA) in 100 mM HEPES (pH 7.5), 200 µM PLP at room temperature under moderate stirring in 200 ml glass vials. After 16 h the conversion, based on the consumption of PAC, was >98 %. The solution was adjusted to pH 10 with triethylamine and extracted with ethyl acetate (5x100 ml). The combined organic layers were dried with $MgSO_4$ and evaporated. The crude product was purified by flash chromatography (eluent: $CHCl_3$/MeOH = 50:50) to afford a mixture of two nor(pseudo)ephedrine diastereomers (*de* = 84 %, according to ^{1}H-NMR and chiral GC) as a pale yellow oil in 85 % yield (121 mg, 0.801 mmol). The absolute configuration was determined by chiral GC comparison and co-injection of the available (1*R*,2*R*)-, (1*S*,2*R*)-, and (1*R*,2*S*)- isomers (see diagram 4 and 5).

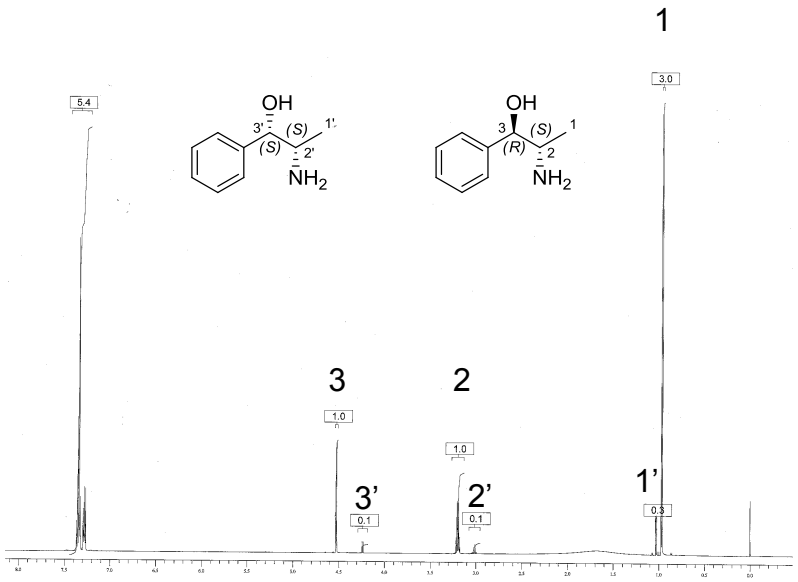


Diagram 2: ^{1}H-NMR (600 MHz; $CDCl_3$) of the diastereomer mixture obtained, (1*S*,2*S*)-norpseudoephedrine and (1*R*,2*S*)-norephedrine

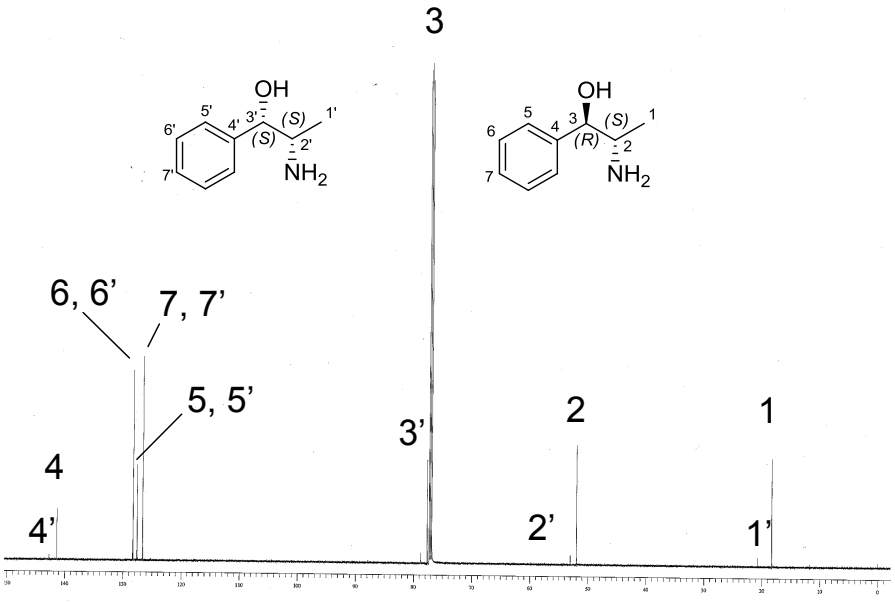


Diagram 3: ^{13}C-NMR (600 MHz; $CDCl_3$) of the diastereomer mixture obtained, (1*S*,2*S*)-norpseudoephedrine and (1*R*,2*S*)-norephedrine

(*R*)-phenylacetylcarbinol (starting material, provided by BASF[4])

R_f (PE/EtOAc 90:10) = 0.27; HPLC (Chiralcel OD-H: 250 mm × 4.6 mm × 5 µm, *n*-hexane/2-propanol = 90:10, 0.7 ml/min, 25 °C): R_t [(S)-5] = 16.2 min, R_t [(*R*)-5] = 17.4 min, *ee* = 84%; ^{1}H-NMR (600 MHz; $CDCl_3$): δ = 2.08 (s, 3 H, 3-H), 4.30 (brs, 1 H, OH), 5.09 (brs, 1 H, 1-H), 7.31–7.40 ppm (m, 5 H, arom.-H); ^{13}C-NMR (151 MHz; $CDCl_3$): 25.3 (C-1), 80.2 (C-2), 127.4 (arom.-CH), 128.8 (arom.-CH), 129.1 (arom.-CH), 138.1 ppm (arom.-C_{ipso}), 207.2 (C-3).

(1*S*,2*S*)- and (1*R*,2*S*)-nor(pseudo)ephedrine diastereomer mixture (obtained from the synthesis described above)

R_f ($CHCl_3$/MeOH = 50:50) = 0.25; chiral GC (Chirasil-DEX CB, 25 m x 0.25 mm x 0.25 µm; constant temperature of 150 °C): (1*S*,2*S*)-norpseudoephedrine = 25.4 min, (1*R*,2*S*)-norephedrine = 30.5 min; *de* = 84%; ^{1}H-NMR (1*R*,2*S*)-norephedrine (600 MHz; $CDCl_3$): δ = 0.97. (d, $J_{1\text{-}2}$=6.4, 3 H, 1-H), 1.7 (brs, 3 H, OH, NH_2), 3.20 (qd, $J_{2\text{-}1}$=6.4, $J_{2\text{-}3}$=4.9, 1 H, 2-H), 4.42 (d, $J_{3\text{-}2}$=4.9, 1 H, 3-H), 7.25–7.38 ppm (m, 5 H, arom.-H); ^{13}C-NMR (1*R*,2*S*)-norephedrine (600 MHz; $CDCl_3$): 18.2 (C-1), 51.9 (C-2), 77.0 (C-3), 126.5 (arom. C-5), 127.5 (arom. C-7), 128.2 (arom. C-6), 141.3 (arom. C-4) ppm. ^{1}H-NMR (1*S*,2*S*)-norpseudoephedrine (600 MHz; $CDCl_3$): δ = 1.03 (d, $J_{1\text{-}2}$=6.4, 3 H, 1'-H), 1.7 (brs, 3 H, OH, NH_2), 3.02 (m, 1 H, 2'-H), 4.24 (d, $J_{3\text{-}2}$=6.8, 1 H, 3'-H), 7.25–7.38 ppm (m, 5 H, arom.-H); ^{13}C-NMR (1*R*,2*S*)-norephedrine (600 MHz; $CDCl_3$): 20.6 (C-1'), 53.0 (C-2'), 78.7 (C-3'), 126.5 (arom. C-5'), 127.5 (arom. C-7'), 128.2 (arom. C-6'), 142.7 (arom. C-4') ppm.

3.3 Determination of absolute configuration of all norephedrine stereoisomers

The absolute configuration of all nor(pseudo)ephedrine isomers: (1*S*,2*S*)-norpseudpephedrine, (1*R*,2*R*)-norpseudoephedrine, (1*R*,2*S*)-norephedrine and (1*S*,2*R*)-norephedrine was determined with chiral GC analysis by comparison of the synthesized reference compound (see chapter 3.2 - diagram 4) and co-injection of the authentic compounds (diagram 5).

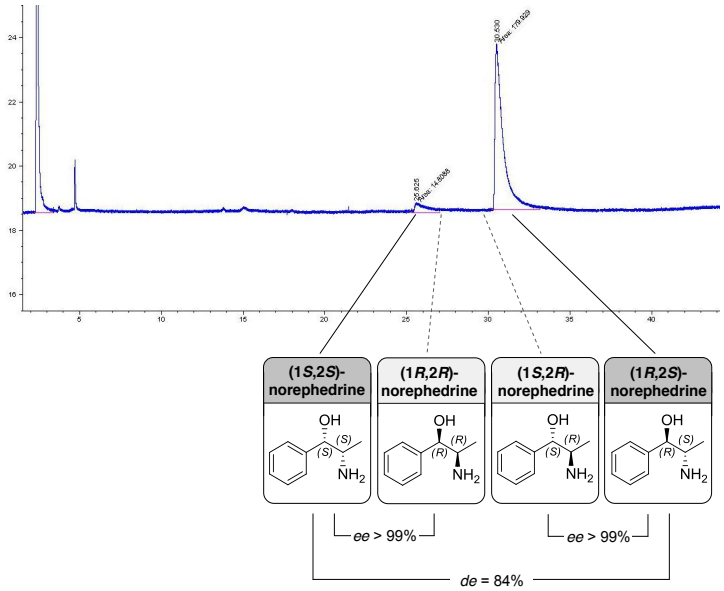


Diagram 4: Chromatogram of nor(pseudo)ephedrine isomers generated from reference compound synthesis (see chapter 3.2): (1*S*,2*S*)-norpseudoephedrine, (1*R*,2*S*)-norephedrine analyzed on a chiral GC Chirasil-DEX CB column (25 m x 0.25 mm x 0.25 µm) with a constant column temperature of 150 °C.

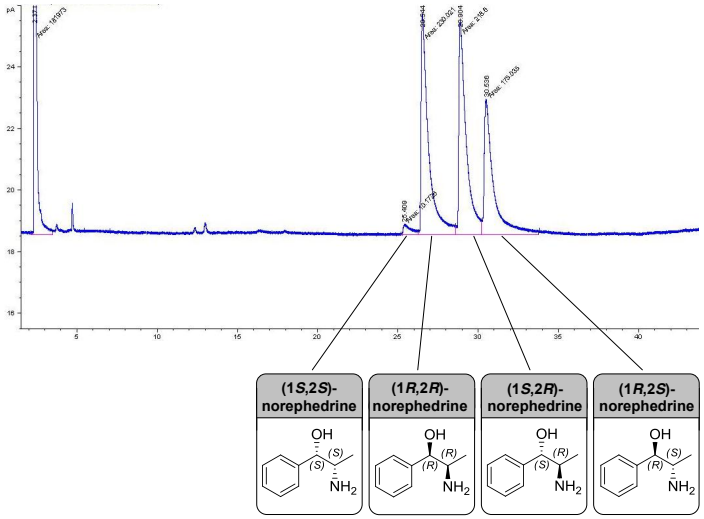


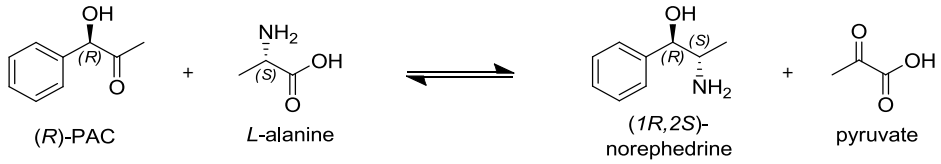
Diagram 5: Chromatogram of all nor(pseudo)ephedrine isomers: (1*R*,2*R*)-norpseudoephedrine and (1*S*,2*R*)-norephedrine as authentic samples (Sigma Aldrich) co-injected with (1*S*,2*S*)-norpseudoephedrine and (1*R*,2*S*)-norephedrine isomers synthesized reference compound (see chapter 3.2).

3.4) Quantification of benzylamine and nor(pseudo)ephedrine isomers

Nor(pseudo)ephedrine was extracted from reaction solutions upon addition of 100 µL NaOH (1 M) to a 200 µL sample obtained from the reaction. After centrifugation, 200 µL of the supernatant were extracted with 200 µL ethyl acetate containing an internal standard (0.1 µL/ml decane) for chiral gas chromatography analysis. For determination of the *ee* and *de* a Chirasil-DEX CB column (25 m x 0.25 mm x 0.25 µm) with a constant column temperature of 130 °C was used. Typical retention times of the four possible product isomers were: (1*S*,2*S*)-norpseudoephedrine = 25.4 min, (1*R*,2*R*)-norpseudoephedrine = 26.5 min, (1*S*,2*R*)-norephedrine = 28.9 min, (1*R*,2*S*)-norpseudoephedrine = 30.5 min (see diagram 5, chapter 3.3). For determination of conversion values a shorter method (linear gradient 100-150 °C with 20 °C/min, hold 150 °C for 6 min, then 180 °C for 3 min) on the same column was used. Typical retention times of the phenylpropanolamine signals were 6.9 min, benzylamine 2.4 min and the injection standard decane 1.8 min.

4.) Determination of the equilibrium constant for the reductive amination of PAC + alanine

The reaction equilibrium constant of the reaction with PAC and alanine to norephedrine and pyruvate (see scheme 2) was determined according to the method described by Tufvesson *et al.* 2012.[8]



Scheme 2: Reaction equilibrium determined for the reaction of PAC+alanine to norephedrine and pyruvate

4.1) Reaction analytics

For determination of all reactant concentrations 20 µL reaction solution was added to 180 µL acetonitrile. From this sample 5 µL or 50 µL, respectively, were injected and analyzed on a LiChrospher RP-8 reverse phase column (250 mm × 4 mm, 5 µm) with a constant column temperature of 20 °C and mixture of H_2O (containing 0.3 % (v/v) H_3PO_4) and acetonitrile (2 min 100 % (v/v) H_2O (containing 0.3 % (v/v) H_3PO_4), followed by a linear gradient to 100 % (v/v) acetonitrile over 8 min and a regeneration step to 100 % (v/v) H_2O (containing 0.3 % (v/v) H_3PO_4). Typical retention times of the reactants were: alanine: 2.43 min, pyruvate 3.74 min, norephedrine = 7.89 min, PAC = 9.37 min. The injection volume and the corresponding linear concentration range at 200 nm for all compounds diluted as described above were: alanine (5 µL) 50-250 mM, pyruvate (5 µL) 1-50 mM, pyruvate (50 µL) 0.1-5 mM, norephedrine (5 µL) 0.1-5 mM, norephedrine (50 µL) 0.01-0.5 mM, PAC (5 µL) 0.1-10 mM.

4.2) Reaction procedure

All reactions were performed in 100 mM HEPES with 200 µM PLP and 0.5 mg/ml *Cv*-(*S*)TA purified enzyme (see chapter 1.1). In order to determine the equilibrium constant (K_{eq}), reactions with different substrates and product concentrations were set up in a way that the reaction quotient (Q) is initially higher, or respectively lower than the equilibrium constant. To estimate the equilibrium constant roughly a reaction with 10 mM PAC and 10 mM alanine was set up. After 12 h reaction time a product concentration of ~0.5 mM norephedrine was observed. Kq could be estimated to be ~ 0.0025 (see equation [#1]).

[#1] $$K_{eq} = \frac{[norephedrine]\cdot[pyruvate]}{[PAC]\cdot[alanine]} \approx \frac{[0.5\,mM]\cdot[0.5\,mM]}{[9.5\,mM]\cdot[9.5\,mM]} \approx 0.0025$$

4.3) Determination of the equilibrium constant

With the roughly estimated equilibrium constant in hand (see chapter 4.2), initial substrate and product concentration for 10 reactions with different reaction quotients were chosen (see table 2). After addition of 0.5 mg/ml transaminase *Cv*-(*S*)TA, the reactant concentrations and with it the reaction quotients were determined at different time points (diagram 6-A).

Table 2: Substrate and product concentration at the initial start point of the reaction and after 24 h

	reaction time: 0.033 h					reaction time: 24 h				
reaction	Q (0 h)	PAC [mM]	alanine [mM]	norephedrine [mM]	pyruvate [mM]	Q (24 h)	PAC [mM]	alanine [mM]	norephedrine [mM]	pyruvate [mM]
#1	0.94765	0.11	100.78	1.00	10.15	0.00230	1.14	100.93	0.03	9.41
#2	0.03946	1.78	89.57	0.90	6.98	0.00230	2.63	91.81	0.09	6.46
#3	0.01991	1.76	89.58	0.69	4.58	0.00227	2.37	92.47	0.11	4.35
#4	0.01056	1.78	88.82	0.46	3.64	0.00230	2.14	91.70	0.12	3.65
#5	0.00531	1.79	89.68	0.47	1.83	0.00224	2.04	91.09	0.22	1.93
#6	0.00329	1.80	90.02	0.29	1.82	0.00235	1.90	91.72	0.19	2.11
#7	0.00145	1.79	89.74	0.25	0.93	0.00231	1.75	91.23	0.28	1.32
#8	0.00049	2.75	133.63	0.19	0.93	0.00236	2.46	136.21	0.48	1.65
#9	0.00025	2.77	135.59	0.10	0.96	0.00237	2.38	135.74	0.45	1.71
#10	0.00005	2.76	177.98	0.05	0.52	0.00225	2.15	182.48	0.60	1.47

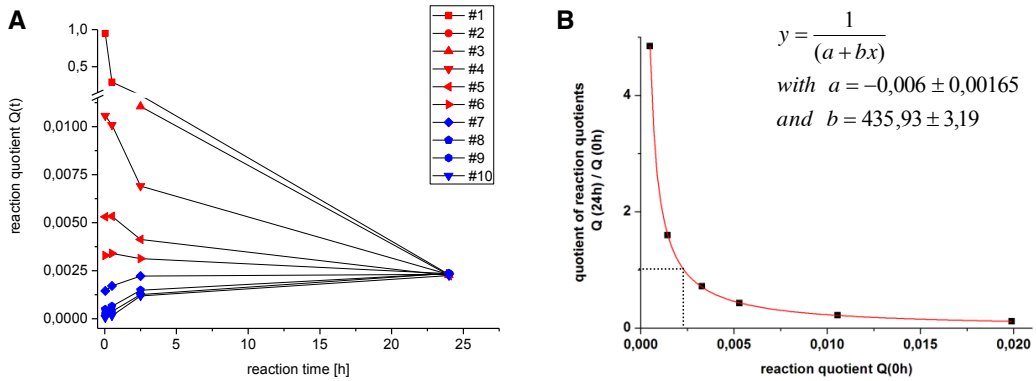


Diagram 6: Reaction quotients plotted against the reaction time (A) and quotient of the reaction quotients Q(0 h)/Q (24 h plotted against Q(0 h) (B)

After 24 h the reaction quotient converges in all reactions to the reaction equilibrium. Mathematically more precise the reaction equilibrium can be calculated from a plot where of the reaction quotients Q(0 h)/Q(24 h) were plotted against Q(0 h) (see Diagram 6-B). A reaction where no change in the reaction quotient over the time is observed, meaning Q(0 h)/Q(24 h)=1, the reaction was initially already in the thermodynamic equilibrium (Q(0 h) = K_{eq}). Thereby, the thermodynamic equilibrium constant can be calculated by solving the equation of the fit (diagram 6-B) with y = 1 (see equation [#2]). The thermodynamic equilibrium constant has a value of $2.31 \cdot 10^{-3}$.

[#2] $$K_{eq} = \frac{Q(24\,h)}{Q(0\,h)} = 1 \Rightarrow y = 1 = \frac{1}{(-0.006 + 435.93 \cdot x)}$$
$$\Leftrightarrow x = 2.31 \cdot 10^{-3} = K_{eq}$$

4.4) Calculation of the theoretical conversion from the reaction equilibrium constant

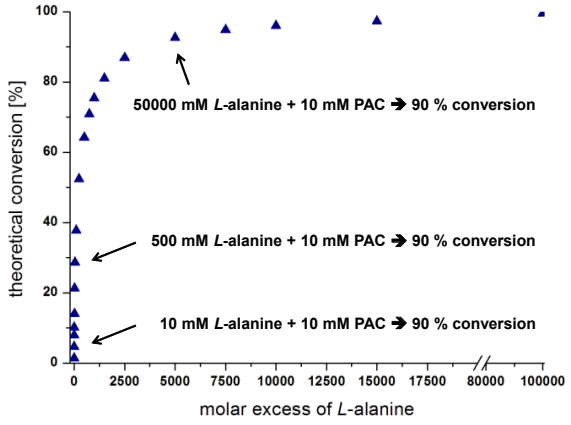
With the given reaction equilibrium constant (see chapter 4.3) the theoretical maximal conversion for reductive amination of PAC with different alanine concentrations (see scheme 2) can be calculated (see equation #3 and diagram 7).

[#3]

$$K_{eq} = \frac{[pyruvate]\cdot[norephedrine]}{[alanine]\cdot[PAC]} = \frac{[x]\cdot[x]}{[alanine-x]\cdot[PAC-x]}$$

$$\Leftrightarrow \; 0 = x^2 + \frac{[\text{alanine}]+[\text{PAC}]}{1/K_{eq}-1}x - \frac{[\text{alanine}][\text{PAC}]}{1/K_{eq}-1}$$

$$\Rightarrow \; x = -\frac{1}{2}\cdot\frac{[\text{alanine}]+[\text{PAC}]}{1/K_{eq}-1} + \sqrt{\left(\frac{1}{2}\cdot\frac{[\text{alanine}]+[\text{PAC}]}{1/K_{eq}-1}\right)^2 - \left(-\frac{[\text{alanine}][\text{PAC}]}{1/K_{eq}-1}\right)}$$

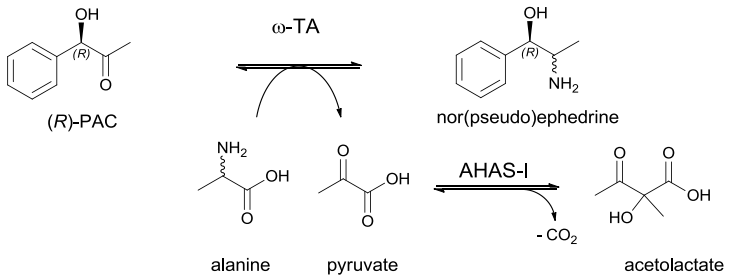


initial concentrations		concentration at reaction equilibrium					
alanine [mM]	PAC [mM]	alanine [mM]	PAC [mM]	pyruvate [mM]	NE [mM]	K_{eq}	conversion [%]
10	**10**	9,542	9,542	0,458	0,458	0,00231	4,58
30	**10**	29,21	9,212	0,788	0,788	0,00231	7,88
50	**10**	48,99	8,992	1,008	1,008	0,00231	10,08
100	**10**	98,60	8,602	1,398	1,398	0,00231	13,98
250	**10**	247,88	7,878	2,122	2,122	0,00231	21,22
500	**10**	497,14	7,139	2,861	2,861	0,00231	28,61
1000	**10**	996,22	6,220	3,780	3,780	0,00231	37,80
2500	**10**	2494,8	4,765	5,235	5,235	0,00231	52,35
5000	**10**	4993,6	3,580	6,420	6,420	0,00231	64,20
7500	**10**	7492,9	2,910	7,090	7,090	0,00231	70,90
10000	**10**	9992,5	2,464	7,536	7,536	0,00231	75,36
15000	**10**	14992	1,899	8,101	8,101	0,00231	81,01
25000	**10**	24991	1,310	8,690	8,690	0,00231	86,90
50000	**10**	49991	0,743	9,257	9,257	0,00231	92,57
75000	**10**	74991	0,520	9,480	9,480	0,00231	94,80
100000	**10**	99990	0,400	9,600	9,600	0,00231	96,00
150000	**10**	149990	0,274	9,726	9,726	0,00231	97,26
250000	**10**	249990	0,168	9,832	9,832	0,00231	98,32
500000	**10**	499990	0,085	9,915	9,915	0,00231	99,15
1000000	**10**	999990	0,043	9,957	9,957	0,00231	99,57

Diagram 7: Theoretical maximal conversion for the reaction of PAC with different alanine concentrations

5.) Optimization of the reductive aminiation of PAC with alanine and AHAS-I coupled equilibrium shift

Several parameters like pH, temperature, substrate and co-substrate concentrations as well as enzyme amount can offer potential for the optimization of an enzymatic reaction. These parameters were optimized for the reductive amination of PAC with alanine as an amine donor and AHAS coupled equilibrium shift (scheme 3). Furthermore, for an optimal process the minimal amount of enzymes (here ω-transaminase and AHAS-I) and co-substrates (here alanine) yielding the highest conversion values were identified.



Scheme 3: Reductive aminiation of PAC with alanine and AHAS-I coupled equilibrium shift

5.1) Reaction procedure and analytics

All reactions for optimizing conversion of the reductive aminiation of PAC with alanine and AHAS-I coupled equilibrium shift were set up in 100 mM HEPES containing 200 µM PLP, 50 µM FAD, 100 µM ThDP and 5 mM $MgCl_2$ and incubated for 20 h in glass vials. The varied parameters pH, incubation temperature, enzyme concentration (*Cv*-(*S*)TA and AHAS-I) and L-alanine concentration were adjusted as declared for the corresponding experiment. The conversion is determined by the product concentration (see chapter 3.4).

5.2) Statistical data analysis and interpolation with the Kriging method

The conversion is studied as a function of various process parameters, and double measurements were performed at several parameter combinations. The Kriging method was used for estimating the true conversion at these double measurements and for interpolating between the measured parameter combinations. Kriging is a statistical method in which the values of the interpolating function are determined by a weighted sum of all measurements. In a rather complicated mathematical procedure, these weights are locally computed, such that the expectation value of the interpolating function matches the true conversion, and that the variance of the difference between the interpolating function and the true conversion is minimal. Also at the measured parameter combinations, the applied variant of the Kriging method approximates the true conversion by a weighted sum of all measurements, which is generally different from the mean of the two measurement values at that specific parameter combination. Moreover, the Kriging method allows improving the interpolation quality by incorporating trend functions that describe basic characteristics but cannot accurately describe all details of the conversion as a function of the process parameters. Further details on the Kriging method can be found in the literature.[9] The following trend function [#4] was chosen for the present study:

[#4] $$conversion = k_{\max} \cdot \frac{x_1 + k_{11}}{x_1 + k_{12} + k_{13} \cdot x_2} \cdot \frac{x_2 + k_{21}}{x_2 + k_{22}}$$

The parameters k_{11} and k_{21} determine the height at the axes x_1=0 and x_2=0, respectively. The parameters k_{12} and k_{22} determine the initial slope in the direction of x_1 and x_2, respectively. The parameter k_{13} determines a decrease of the initial slope in the direction of x_1 when x_2 increases. The parameter k_{max} determines the height of the plateau. The axis x_1 denotes either alanine / PAC or *Cv*-(*S*)TA, whereas the axis x_2 always denotes AHAS-I. The parameter k_{max} is estimated in one procedure with the weights of the Kriging method. All other parameters of the trend function

are manually determined before the Kriging method is applied, such as to match the observed data as good as possible. The following parameters have been used in the presented figures:

	k_{11}	k_{12}	k_{13}	k_{12}	k_{22}
Figure 3	0	0.125	0	-0.100	0.150
Diagram 8a	0	0.050	0	0	0.100
Diagram 8b	0	0.125	0	0	0.250
Diagram 8c	0	0.125	0	0	0.250
Diagram 8d	0.125	1.375	0	0	0.150
Diagram 9	0	0.024	0.026	-0.040	2.360

Diagram 10 has been computed without using a trend function.

5.3) Parameter optimization: enzyme concentrations, alanine / PAC ratio, pH, reaction temperature

In independent reaction setups, the enzyme (*Cv*-(*S*)TA and AHAS-I) concentrations and the alanine/PAC ratio was varied while the pH (7.5) and the temperature (30 °C) of the reaction was constant (diagram 8 A-D). Reaction optima with a plateau shape were observed in all cases, where with an increase of the enzyme concentration no further increase of the product concentration was accessible (diagram 8 A-D). Here, optimal parameters in terms of reaction costs and efficiency are the minimal enzyme concentrations yielding the maximal conversion at the lowest amine donor/PAC ratio. In all cases the optimal *Cv*-(*S*)TA concentrations has a value of 1 mg/ml (diagram 8 A-D). In a plot of all values with different alanine/PAC ratio and AHAS-I concentrations at a constant *Cv*-(*S*)TA concentration of 1 mg/ml (diagram 9), the optimal parameter can be read out more easily: 0.5 mg/ml AHAS-I and a 5-fold excess of L-alanine/PAC. Moreover, with 250 mM L-alanine the significant lower conversion values indicate a putative substrate surplus inhibition.

For these optimized parameters the maximal conversion values in terms of pH and temperatures were determined to be at range pH of 7.5 - 8.0 and temperature of 25 - 30 °C (diagram 10). In a combination of all optimized parameters, the reductive aminiation of PAC to norephedrine with alanine and AHAS-I coupled equilibrium shift can be performed with a conversion of >80 %.

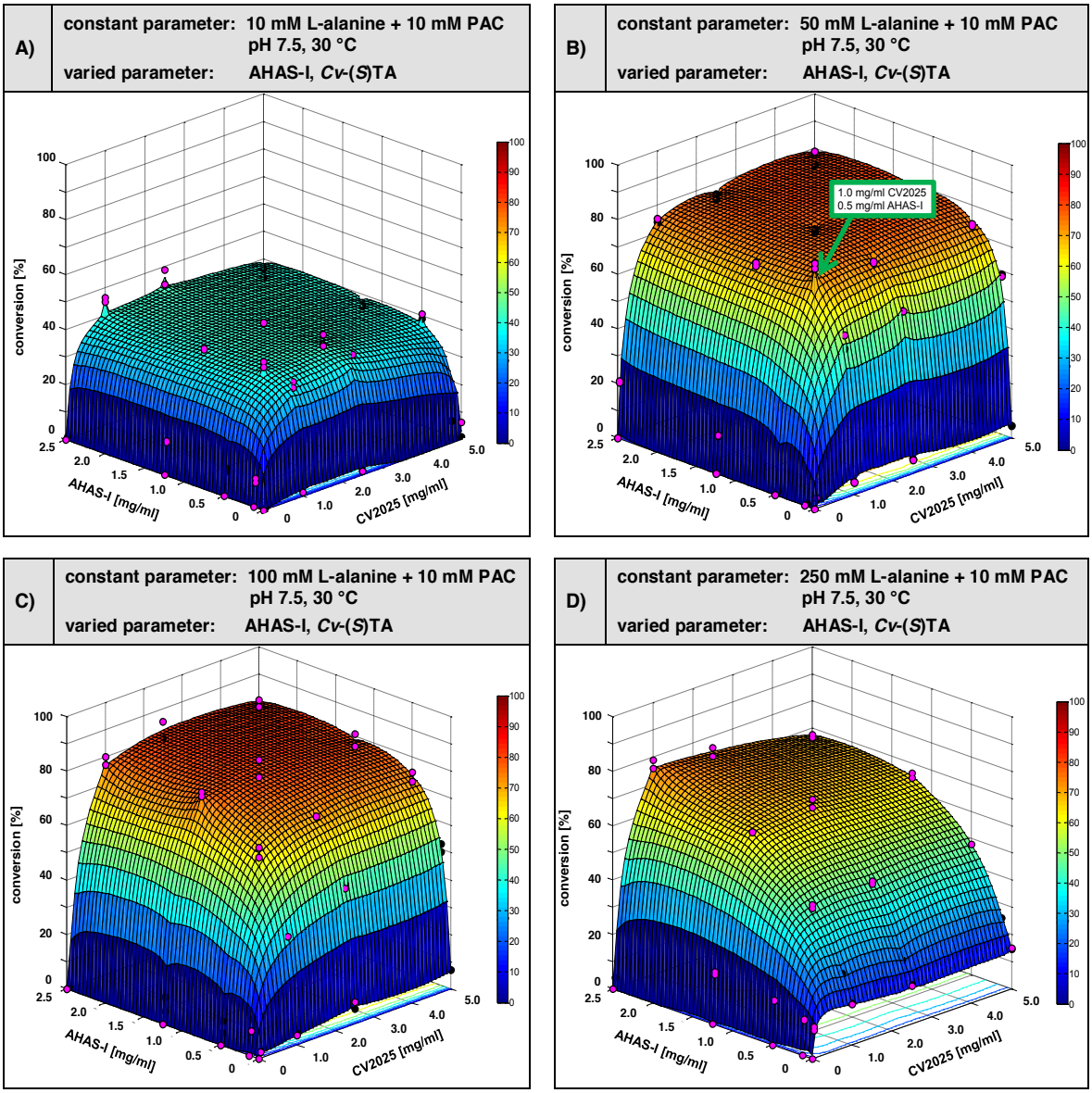


Diagram 8: Determination of optimal reaction parameters (here *Cv*-(*S*)TA, AHAS-I and L-alanine concentration) for the reductive amination of PAC with AHAS-I coupled equilibrium shift. The reactions were set up in 100 mM HEPES (pH 7.5) containing 200 µM PLP, 50 µM FAD, 100 µM ThDP and 5 mM $MgCl_2$ and incubated for 20 h at 30 °C in glass vials. For the statistical data analysis and for interpolating the data (pink dots) the 'kringing algorithm' (see chapter 5.2) was used.

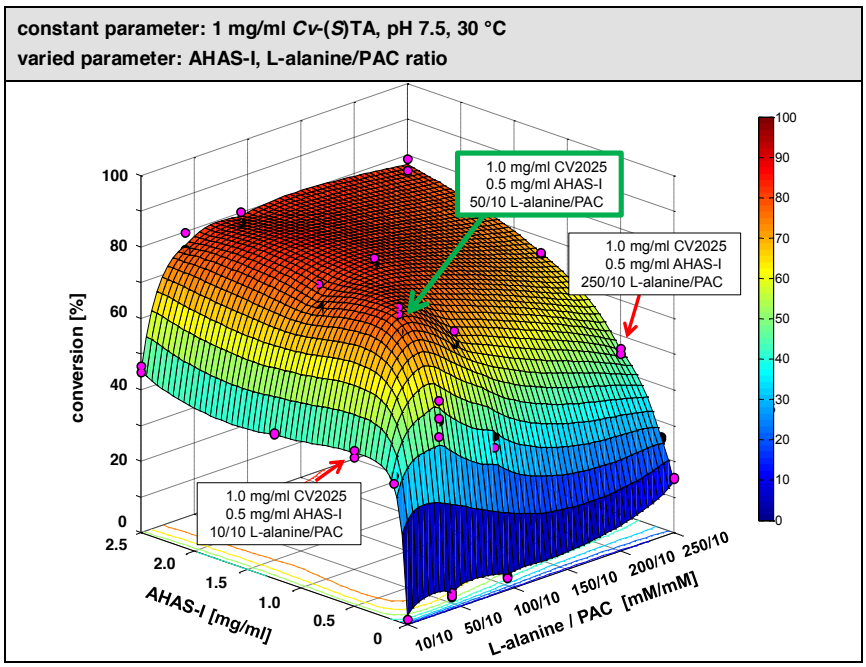


Diagram 9: Plot of all values with different L-alanine/PAC ratio and AHAS-I concentrations at constant transaminase concentrations of 1 mg/ml (see diagram 9).

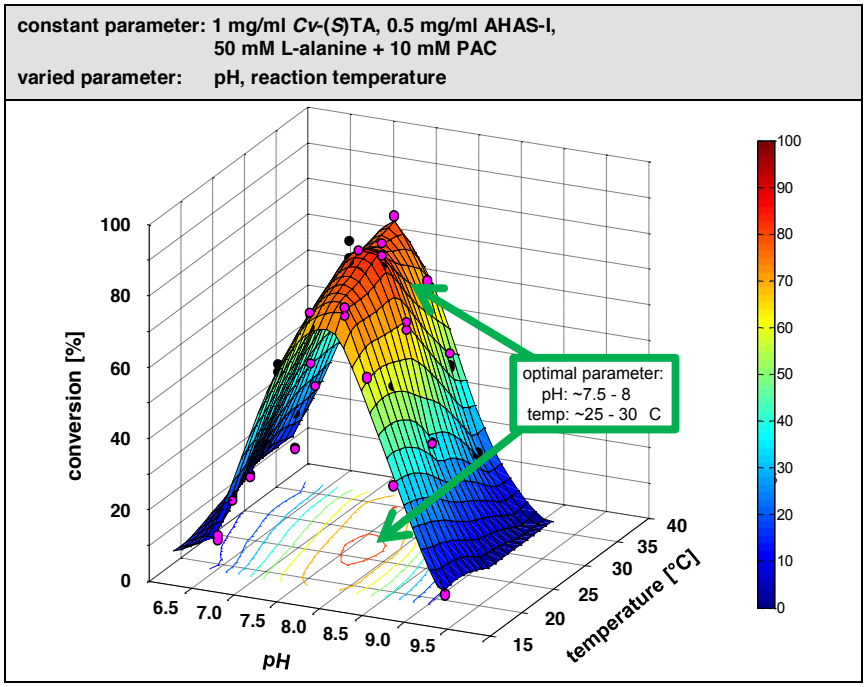
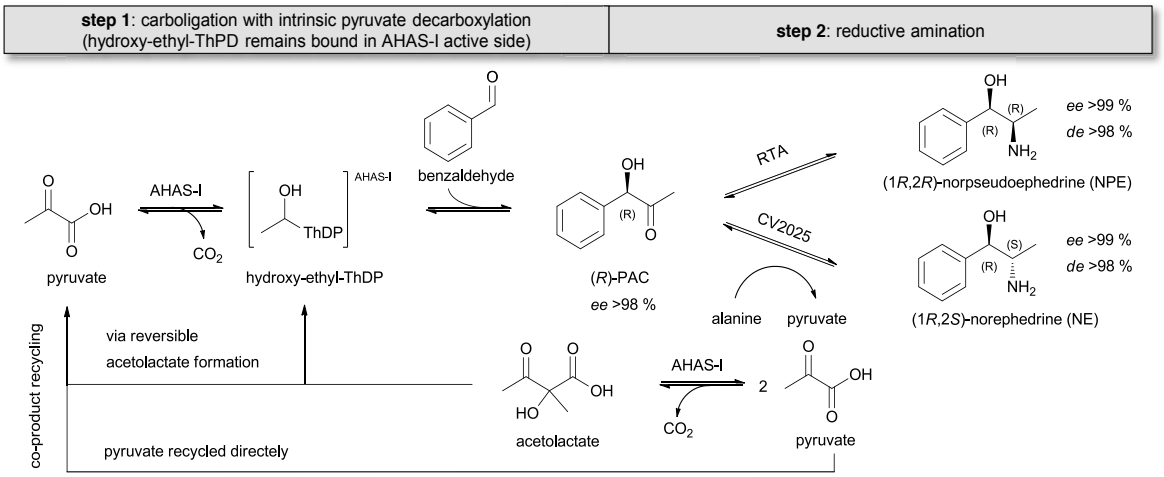


Diagram 10: Determination of optimal reaction parameters (here pH and reaction temperature) for the reductive amination of PAC with AHAS-I (0.5 mg/ml) coupled equilibrium shift and 1 mg/ml *Cv*-(*S*)TA. The reactions were set up in 100 mM HEPES containing 200 µM PLP, 50 µM FAD, 100 µM ThDP and 5 mM $MgCl_2$ and incubated for 20 h at different temperatures in glass vials. For the statistical data analysis and for interpolating the data (pink dots) the 'kringing algorithm' (see chapter 5.2) was used.

6.) 1-pot 2-step cascade reactions



Scheme 4. 1-Pot 2-step cascade for the synthesis of nor(pseudo)ephedrine performed as a recycling cascade.

6.1) Reaction procedure and analytics

A general reaction procedure for the 1-pot 2-step sequential cascade reaction contained 10 mM benzaldehyde, 10 mM pyruvate and 50 mM alanine. D-alanine was used in case of (*R*)-selective TA, and respectively L-alanine for (*S*)-selective TA. The reaction with 1 mg/ml TA with 0.5 mg/ml AHAS-I was performed in 100 mM HEPES (pH 7.5), 200 µM pyridoxal-5'-phosphate (PLP), 50 µM flavin adenine dinucleotide (FAD), 100 µM thiamine diphosphate (ThDP), 5 mM $MgCl_2$ and incubated at 25 °C in glass vials with moderate shaking (100 rpm). For the additional recycling step 10 mM benzaldehyde, 0.5 mg/ml AHAS-I and if indicated 1 mg/ml TA was added. In case of the 1-pot 2-step simultaneous cascade 1 mg/ml TA and 0.5 mg/ml AHAS-I were added simultaneously to a solution containing 10 mM benzaldehyde (respectively 20 mM benzaldehyde in case of the recycling cascade), 10 mM pyruvate and 50 mM alanine. The time-dependent reductive amination of 10 mM benzaldehyde or 10 mM PAC was carried out with 1 mg/ml TA (*Cv*-(*S*)TA or *At*-(*R*)TA, respectively) in 100 mM HEPES (pH 7.5 with 200 µM PLP, 50 µM FAD, 100 µM ThDP, 5 mM $MgCl_2$) containing either (*S*)- or (*R*)-α-methylbenzylamine (10 mM) as amine donor. The concentrations of PAC and benzaldehyde were determined as described in chapter 3.1, for quantification of benzylamine and nor(pseudo)ephedrine see chapter 3.4. The enantiomeric excess of PAC and the *de* and *ee* of nor(pseudo)ephedrine were determined as described in chapter 3.2 and 3.3.

6.2) 1-pot 2-step syntheses of nor(pseudo)ephedrine performed as a sequential cascade with additional 'recycling step' (time-dependent reaction curves)

The 1-pot 2-step syntheses of nor(pseudo)ephedrine performed as a sequential recycling cascade was described in the main text (see description of table 1 – main text). Here, the time-dependent reaction curve for the syntheses of (1*R*,2*S*)-NE (diagram 11) and (1*R*,2R)-NPE (diagram 12) are shown.

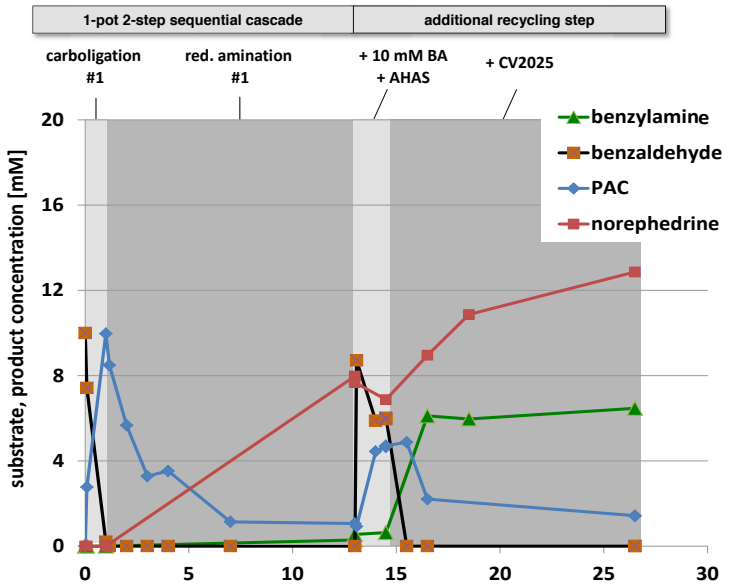


Diagram 11: Time-dependent reaction curve for the 1-pot 2-step syntheses of (1*R*,2*S*)-NE performed as a sequential cascade with additional 'recycling step' combining AHAS-I and *Cv*-(*S*)TA.
(interpolation between dots do not refer to actual concentrations)

Reaction conditions:	100 mM HEPES (pH 7.5 with 200 µM PLP, 50 µM FAD, 100 µM ThDP, 5 mM $MgCl_2$), 25 °C, 100 rpm
1-pot 2-step	carboligation #1: 10 mM benzaldehyde, 10 mM pyruvate, 0.5 mg/ml AHAS red. amination #1: + 50 mM L-alanine, + 1mg/ml ω-TA
recycling step	A: +10 mM benzaldehyde, + 0.5 mg/ml AHAS-I B: + 1 mg/ml *Cv*-(*S*)TA

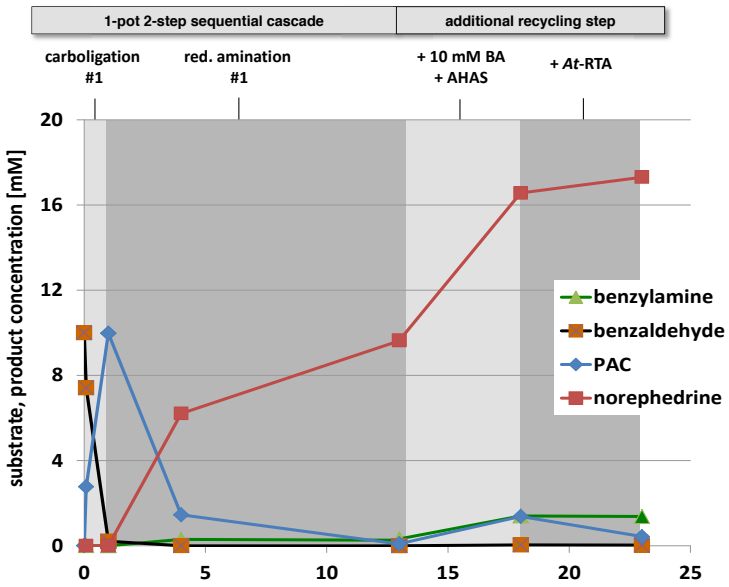


Diagram 11: Time dependent reaction curve for the 1-pot 2-step syntheses of (1*R*,2*R*)-NPE performed as a sequential cascade with additional 'recycling step' combining AHAS-I and *At*-(*R*)TA.
(interpolation between dots do not refer to actual concentrations)

Reaction conditions:	100 mM HEPES (pH 7.5 with 200 µM PLP, 50 µM FAD, 100 µM ThDP, 5 mM $MgCl_2$), 25 °C, 100 rpm
1-pot 2-step	carboligation #1: 10 mM benzaldehyde, 10 mM pyruvate, 0.5 mg/ml AHAS red. amination #1: + 50 mM D-alanine, + 1mg/ml ω-TA
recycling step	A: +10 mM benzaldehyde, + 0.5 mg/ml AHAS-I B: + 1 mg/ml *At*-(*R*)TA

7.) References

[1] a) J. H. Schrittwieser, J. Sattler, V. Resch, F. G. Mutti, W. Kroutil, *Curr. Opin. Chem. Biol.* **2011**, *15*, 249-256; b) E. Ricca, B. Brucher, J. H. Schrittwieser, *Adv. Synth. Catal.* **2011**, *353*, 2239-2262; c) S. F. Mayer, W. Kroutil, K. Faber, *Chem. Soc. Rev.* **2001**, *30*, 332-339; d) A. Bruggink, R. Schoevaart, T. Kieboom, *Org. Process Res. Dev.* **2003**, *7*, 622-640; e) S. Shanmuganathan, D. Natalia, L. Greiner, P. Dominguez de Maria, *Green Chem.* **2012**, *14*, 94-97; f) C. V. Voss, C. C. Gruber, W. Kroutil, *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 741-745.

[2] U. Schell, R. Wohlgemuth, J. M. Ward, *J. Mol. Cat. B: Enzym.* **2009**, *59*, 279-285.

[3] U. Kaulmann, K. Smithies, M. E. B. Smith, H. C. Hailes, J. M. Ward, *Enzyme Microb. Tech.* **2007**, *41*, 628-637.

[4] T. Sehl, R. C. Simon, H. C. Hailes, J. M. Ward, U. Schell, M. Pohl, D. Rother, *J. Biotechnol.* **2012**, *159*, 188-194.

[5] S. Engel, M. Vyazmensky, S. Geresh, Z. Barak, D. M. Chipman, *Biotechnol. Bioeng.* **2003**, *83*, 833-840.

[6] A. Bar-Ilan, V. Balan, K. Tittmann, R. Golbik, M. Vyazmensky, G. Hübner, Z. Barak, D. M. Chipman, *Biochem.* **2001**, *40*, 11946-11954.

[7] http://www.enzymicals.com/Enzymicals_Enzymes_v011.pdf.

[8] P. Tufvesson, J. S. Jensen, W. Kroutil, J. M. Woodley, *Biotechnol. Bioeng.* **2012**, *109*, 2159-2162.

[9] a) N. A. G. Cressie, *Statistics for Spatial Data, Revised Edition*, John Wiley & Sons, Ltd., ISBN: 9780471002550, **1993**; b) M. Sherman, *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*, John Wiley & Sons, Ltd., ISBN: 9780470699584, **2010**.

PUBLICATION II

(*S*)-Selective MenD variants from *Escherichia coli* provide access to new functionalized chiral α-hydroxy ketones

Westphal, R.; Waltzer, S.; Mackfeld, U.; Widmann, M.; Pleiss, J.;
Beigi, M.; Müller, M.; Rother, D. & Pohl, M.

Chemical Communications 2013,
49, 2061–2063
DOI: 10.1039/C3CC38607D
(http://pubs.rsc.org/en/Content/ArticleLanding/2013/CC/c3cc38607d)

ChemComm

RSC Publishing

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2013, **49**, 2061

Received 30th November 2012,
Accepted 22nd January 2013

DOI: 10.1039/c3cc38607d

www.rsc.org/chemcomm

(*S*)-Selective MenD variants from *Escherichia coli* provide access to new functionalized chiral α-hydroxy ketones†

Robert Westphal,[a] Simon Waltzer,[b] Ursula Mackfeld,[a] Michael Widmann,[c] Jürgen Pleiss,[c] Maryam Beigi,[b] Michael Müller,[b] Dörte Rother[a] and Martina Pohl*[a]

We report the first rationally designed (*S*)-selective MenD from *E. coli* for the synthesis of functionalized α-hydroxy ketones. By mutation of two amino acids in the active site stereoselectivity of the (*R*)-selective *Ec*MenD (ee > 93%) was inverted giving access to (*S*)-5-hydroxy-4-oxo-5-phenylpentanoate derivatives with stereoselectivities up to 97% ee.

Optically active α-oxyfunctionalized carbonyl compounds are valuable synthetic building blocks in preparative organic chemistry.[1] Among them, α-hydroxy ketones are of particular value for the pharmaceutical as well as the fine chemistry sector.[2] In addition, α-hydroxy ketones are putative precursors of 2-amino alcohols, or 1,2-diols, among others.[3,4] Various non-enzymatic routes to chiral α-hydroxy ketones have been reported including organocatalytic strategies.[5,6] However, high stereoselectivities are rare. Furthermore, most of the syntheses require several steps, which impair the overall yields.[2]

Thiamine diphosphate (ThDP)-dependent enzymes are well known for their catalytic potential to form various α-hydroxy ketones. The mechanism has been studied extensively[7] and several ThDP-dependent lyases, such as pyruvate decarboxylases, branched-chain keto acid decarboxylase, benzoylformate decarboxylase, and benzaldehyde lyase, have been already characterized as powerful catalysts.[8] As most of the wild type (wt) enzymes are (*R*)-selective, access to (*S*)-α-hydroxy ketones is limited. Structural basis for stereoselectivity has been investigated with the benzoylformate decarboxylase from *Pseudomonas putida*, which shows (*S*)-selectivity in the carboligation of benzaldehyde and acetaldehyde.[9,10] The study revealed a structural element called "S-pocket", which allows an antiparallel arrangement of donor and acceptor substrates, the prerequisite for (*S*)-selectivity.[11] S-pockets are present in most of the (*R*)-selective enzymes, however, not accessible to acceptor substrates due to large amino acid side chains.[12,13]

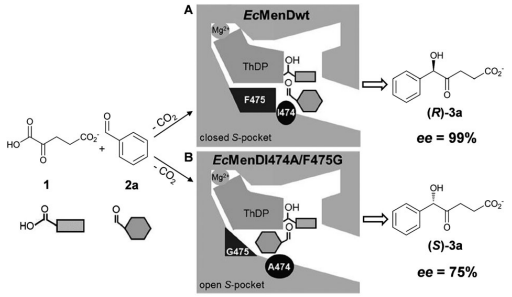
The ThDP-dependent enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase from *Escherichia coli* (*Ec*MenD) catalyzes the second step in the biosynthesis of menaquinones.[14] The carboligation potential of *Ec*MenD has recently been characterized concerning substrate diversity and stereoselectivity.[15,16] *Ec*MenD uses α-ketoglutarate (**1**) as the physiological donor, which is extraordinary among ThDP-dependent enzymes. This enables selective C4 chain elongation with a terminal carboxyl group. *Ec*MenD accepts a broad range of aldehydes as acceptors. Carboligation of **1** (upon decarboxylation) with different aromatic aldehydes gives α-hydroxy ketones with high enantiomeric excesses (ee) of >93% (*R*).[16] Here, we report the first rationally designed *Ec*MenD variants for the syntheses of functionalized (*S*)-α-hydroxy ketones starting from **1** and differently substituted benzaldehyde derivatives **2**.

We combined rational protein engineering with substrate engineering to get access to functionalized (*S*)-α-hydroxy ketones. Important residues of the S-pocket were identified based on the crystal structure of *Ec*MenD (2JLC).[17] Using docking studies (see ESI†) with benzaldehyde (**2a**), two residues were deduced to be crucial for stereoselectivity: I474 and F475 prevent the antiparallel arrangement of **2a** prior to C–C-bond formation. This explains the high (*R*)-selectivity of 99% ee of the wt enzyme in the carboligation reaction with **1** (Scheme 1A). Hence, in order to gain sufficient space for the phenyl ring of **2a**, both I474 and F475 were mutated to glycine and alanine, respectively.

Four variants were prepared: I474G/F475G, I474G/F475A, I474A/F475G, and I474A/F475A. In all cases, the formation of (*S*)-5-hydroxy-4-oxo-5-phenylpentanoate (**3a**) was observed. Whereas variants I474G/F475G and I474G/F475A showed only low selectivity for the formation of (*S*)-**3a** (ee 34% and ee 27%), the variants I474A/F475A and I474A/F475G revealed higher (*S*)-selectivities of 66% ee and 75% ee, respectively. Thus, the S-pocket could be opened such that **2a** was able to arrange

[a] *IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Leo-Brandt-Str. 1, 52425 Jülich, Germany. E-mail: ma.pohl@fz-juelich.de; Fax: +49-(0)2461-61-3870; Tel: +49-(0)2461-61-4388*
[b] *Institute of Pharmaceutical Sciences, Albert-Ludwigs-Universität Freiburg, Albertstr. 25, 79104 Freiburg, Germany*
[c] *Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany*

† Electronic supplementary information (ESI) available: Materials and experimental procedures, details of structural investigation of *Ec*MenD. See DOI: 10.1039/c3cc38607d

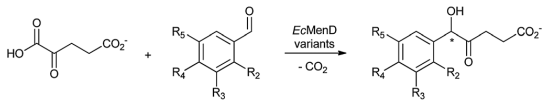


Scheme 1 Inversion of the stereoselectivity of wt *Ec*MenD by structure-based design of the active site.

predominantly in an antiparallel orientation relative to the ThDP-bound donor aldehyde (Scheme 1B). Consequently, the results support the previously developed S-pocket concept.[11,12] Besides accessibility of the S-pocket, optimal stabilization of the acceptor **2a** is crucial for high stereoselectivity. Best stabilization of antiparallely oriented **2a** seemed to be achieved in I474A/F475G. However, the catalytic activity, which may be estimated from the conversions of the *Ec*MenD variants, was lower compared to the wt enzyme (Table 1). Whereas I474G/F475A and I474A/F475A still showed a good conversion of 67% and 68%, respectively, variants I474G/F475G and I474A/F475G revealed only 26% and 15% conversion after 24 h under the tested conditions. Mutation of two adjacent amino acids in the active site might destabilize the binding of the substrate, in which the introduction of glycine at position 475 seemed to have a significant influence on the conversion.

The full catalytic potential of the three best (*S*)-selective *Ec*MenD variants I474G/F475G, I474A/F475G, and I474A/F475A was explored in the next step. Various *ortho*-, *meta*-, and *para*-substituted benzaldehydes were tested beginning with fluoro- and chlorobenzaldehydes (**2b–g**). As expected, stereoselectivity was affected with all substrates compared to the high (*R*)-selectivity of *Ec*MenDwt (ee > 93%, Table 1). *Ec*MenD variants showed the following trends: all tested *ortho*- and *para*-substituted benzaldehyde derivatives resulted in a higher amount of the (*R*)-enantiomer compared to **2a**, while the *meta*-substituted benzaldehydes led to higher (*S*)-selectivity relative to **2a**. Under non-optimized experimental conditions lowest conversions of <24% were observed for *para*-substituted derivatives **2d** and **2g**, which might be a consequence of weak substrate stabilization in the active site concomitant with low conversion rates. Compared to the *para*-substituted benzaldehydes higher conversions of 31–76% were observed with *ortho*-substituted derivatives. As in both cases the (*R*)-enantiomers are predominantly formed, a sub-optimal stabilization of *ortho*- and *para*-substituted benzaldehydes in the S-pockets of the three tested *Ec*MenD variants can be deduced. Remarkable results were obtained with *meta*-substituted benzaldehydes as acceptors: (*S*)-selectivity as well as conversion could be enhanced significantly with all variants compared to benzaldehyde **2a** as the acceptor. Employing **2c** and especially **2f** stereoselectivities of 59–89% ee (*S*) and conversions of 43–94% were obtained. Consequently, the nature of the substrate plays an important role. As *meta*-substituted benzaldehyde derivatives seem to be highly beneficial for (*S*)-selective carboligation using *Ec*MenD variants, **2h–2j** were also tested in the carboligation with **1**. Whereas substrate **2j** was also highly (*S*)-selectively transformed with an ee of 78–93%, the most striking result was obtained with **2h** and **2i**. With these substrates an almost complete inversion of stereoselectivity was possible, connected with high conversions of >86% under the tested conditions. For instance, I474A/F475A catalyzes the formation of **3h** with complete conversion and an ee of 95% (*S*). In order to examine

Table 1 Carboligation of α-ketoglutarate (**1**) and differently substituted benzaldehydes (**2**) to functionalized chiral α-hydroxy ketones (**3**) by *Ec*MenDwt and variants[a]



1 **2a-k** **3a-k**

					wt		I474G/F475G		I474A/F475A		I474A/F475G	
Product	R_2	R_3	R_4	R_5	Conv.[b] (%)	ee[c] (%)	Conv.[b] (%)	ee[c] (%)	Conv.[b] (%)	ee[c] (%)	Conv.[b] (%)	ee[c] (%)
3a	H	H	H	H	>99	>99 (*R*)	26	34 (*S*)	68	66 (*S*)	15	75 (*S*)
3b	F	H	H	H	>99	94 (*R*)	66	≤5	76	20 (*S*)	55	48 (*S*)
3c	H	F	H	H	>99	96 (*R*)	58	59 (*S*)	75	80 (*S*)	43	82 (*S*)
3d	H	H	F	H	>99	99 (*R*)	10	32 (*R*)	24	30 (*S*)	14	46 (*S*)
3e	Cl	H	H	H	98	93 (*R*)	31	70 (*R*)	66	40 (*R*)	57	≤5
3f	H	Cl	H	H	>99	96 (*R*)	75	85 (*S*)	82	88 (*S*)	94	89 (*S*)
3g	H	H	Cl	H	>99	93 (*R*)	21	89 (*R*)	12	70 (*R*)	15	48 (*R*)
3h	H	Br	H	H	99	99 (*R*)	99	94 (*S*)	99	95 (*S*)	87	97 (*S*)
3i	H	I	H	H	>99	94 (*R*)	>99	91 (*S*)	99	93 (*S*)	86	93 (*S*)
3j	H	OCH_3	H	H	99	>99 (*R*)	99	78 (*S*)	89	87 (*S*)	64	93 (*S*)
3k	H	OCH_3	H	OCH_3	99	98 (*R*)	99	85 (*S*)	98	96 (*S*)	83	96 (*S*)

[a] Carboligation reaction was performed in 50 mM potassium phosphate buffer, pH 8.0, (2 mM $MgSO_4$, 0.1 mM ThDP, and 5 vol% MTBE) with 0.70 mg of the respective *Ec*MenD variant at 30 °C and 300 rpm for 24 h using 50 mM α-ketoglutarate (**1**) and 20 mM benzaldehyde derivatives (**2a–k**). [b] Conversion was determined by GC-MS [DB-5MS Agilent] based on benzaldehyde consumption. [c] Determined by chiral HPLC analysis using a Chiralpak IC, and Chiralcel OB and OD-H, respectively. The absolute configuration was determined by circular dichroism.

the full potential of the S-pocket, 3,5-di-methoxybenzaldehyde (**2k**) was finally tested. The results obtained with variants I474A/F475G and I474A/F475A, which catalyze the formation of (*S*)-**3k** with an ee of 96% and conversions of >83%, demonstrate that even this bulky substrate was able to enter the S-pocket properly.

The designed S-pocket of the *Ec*MenD variants seems to be tailored for *meta*-substituted benzaldehydes. High stereoselectivities as well as high conversions suggested good stabilization of the acceptor in the S-pocket. Furthermore, the parallel acceptor orientation, leading to the respective (*R*)-product, might be destabilized, too (Scheme 1). Reaction engineering might further improve stereoselectivities and conversions. To explain these results as well as the selectivities with *ortho*- and *para*-substituted substrates, structural studies with these (*S*)-selective *Ec*MenD variants are currently underway. Together with molecular dynamics simulations of the binding states (parallel and antiparallel orientation of the acceptor), we want to get a deeper insight into *Ec*MenD selectivity and carboligation activity. Furthermore, we want to explore the scope of *Ec*MenD catalysis with other acceptor substrates, *e.g.* aliphatic aldehydes.

In conclusion, we have rationally designed the first MenD variants as powerful biocatalysts for the (*S*)-selective synthesis of functionalized α-hydroxy ketones starting from differently substituted benzaldehydes. Mutation of two amino acids in the active site could invert the stereoselectivity of the wt enzyme for most of the investigated substrates. Particularly the use of *meta*-substituted substrates resulted in high (*S*)-selectivities accompanied by good to excellent conversions. Therewith, we provide access to new, yet not reported (*S*)-5-hydroxy-4-oxo-5-phenylpentanoate derivatives.

Financial support by the German Research Foundation (DFG) in the frame of the research group ''Diversity of asymmetric thiamine catalysis'' FOR1296 and partial funding by the research training group ''BioNoCo'' (GK1166) are gratefully acknowledged.

Notes and references

1 W. Adam, M. Lazarus, C. R. Saha-Möller and P. Schreier, *Acc. Chem. Res.*, 1999, **32**, 837.
2 P. Hoyos, J. V. Sinisterra, F. Molinari, A. R. Alcantara and P. D. de Maria, *Acc. Chem. Res.*, 2010, **43**, 288.
3 V. B. Shukla and P. R. Kulkarni, *World J. Microbiol. Biotechnol.*, 2000, **16**, 499.
4 J. Kulig, R. C. Simon, C. A. Rose, S. M. Husain, M. Hackh, S. Lüdecke, K. Zeitler, W. Kroutil, M. Pohl and D. Rother, *Catal. Sci. Technol.*, 2012, **2**, 1580.
5 D. Enders, O. Niemeier and A. Henseler, *Chem. Rev.*, 2007, **107**, 5606.
6 S. E. O'Toole, C. A. Rose, S. Gundala, K. Zeitler and S. J. Connon, *J. Org. Chem.*, 2011, **76**, 347.
7 F. Jordan, *Nat. Prod. Rep.*, 2003, **20**, 184.
8 M. Müller, D. Gocke and M. Pohl, *FEBS J.*, 2009, **276**, 2894.
9 R. Wilcocks and O. P. Ward, *Biotechnol. Bioeng.*, 1992, **39**, 1058.
10 H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir and M. Pohl, *Chem.–Eur. J.*, 2000, **6**, 1483.
11 D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C. L. Berthold, G. Schneider, J. Pleiss, M. Müller and M. Pohl, *ChemBioChem*, 2008, **9**, 406.
12 D. Rother, G. Kolter, T. Gerhards, C. L. Berthold, E. Gauchenova, M. Knoll, J. Pleiss, M. Müller, G. Schneider and M. Pohl, *ChemCatChem*, 2011, **3**, 1587.
13 M. Pohl, D. Gocke and M. Müller, in *Handbook of Green Chemistry, Green Catalysis, Vol. 3, Biocatalysis*, ed. P. T. Anastas, Wiley-VCH, Weinheim, 2008, vol. 2009, pp. 75–114.
14 M. Jiang, Y. Cao, Z.-F. Guo, M. Chen, X. Chen and Z. Guo, *Biochemistry*, 2007, **46**, 10979.
15 J. Bongaerts, S. Esser, V. Lorbach, L. Al-Momani, M. A. Müller, D. Franke, C. Grondal, A. Kurutsch, R. Bujnicki, R. Takors, L. Raeven, M. Wubbolts, R. Bovenberg, M. Nieger, M. Schürmann, N. Trachtmann, S. Kozak, G. A. Sprenger and M. Müller, *Angew. Chem., Int. Ed.*, 2011, **50**, 7781.
16 A. Kurutsch, M. Richter, V. Brecht, G. A. Sprenger and M. Müller, *J. Mol. Catal. B: Enzym.*, 2009, **61**, 56.
17 A. Dawson, P. K. Fyfe and W. N. Hunter, *J. Mol. Biol.*, 2008, **384**, 1353.

Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

(*S*)-Selective MenD variants from *Escherichia coli* provide access to new functionalized chiral α-hydroxy ketones

Robert Westphal,[a] Simon Waltzer,[b] Ursula Mackfeld,[a] Michael Widmann,[c] Jürgen Pleiss,[c] Maryam Beigi,[b] Michael Müller,[b] Dörte Rother,[a] Martina Pohl,[*a]

[a]*IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Leo-Brandt-Str., 52425 Jülich, Germany*
[b]*Institute of Pharmaceutical Sciences, Albert-Ludwigs-Universität Freiburg, Albertstr. 25, 79104 Freiburg, Germany*
[c]*Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany*

ma.pohl@fz-juelich.de

Table of contents

Electronic Supplementary Material (ESI) for Chemical Communications

1. General experimental procedures

All chemical reagents, solvents, buffers, salts, α-ketoglutaric acid and benzaldehyde derivatives were purchased from Sigma-Aldrich and used without further purification. KOD Hot Start DNA Polymerase was obtained from Merck Millipore and the restriction enzyme *DpnI* from (Thermo Scientific).

Bradford assays were measured at 595 nm on a UV mini – 1240 spectrophotometer (SHIMADZU). The **optical density** (OD) of bacterial cultures was determined at 600 nm on a UV – 1601 spectrophotometer (SHIMADZU). **GC-MS analyses** were performed on a HP 6890 N Series GC system (EI, 70 eV) with a HP 5973 Network Mass Selective Detector (Agilent), equipped with a DB-5MS column (Agilent, 30 m × 0.25 mm, 0.25 µm thick stationary phase), using the following conditions: injector temp: 250 °C, detector temp: 300 °C, flow rate: 25 mL min^{-1}, temp program: 60 ºC for 3 min, then 20 ºC min^{-1} to 280 ºC for 4 min. **Nuclear magnetic resonance** (NMR) spectra were measured on a DRX 400 instrument (Bruker) operating at 400 and 100 MHz for ^{1}H and ^{13}C acquisitions, respectively. Chemical shifts (δ) of ^{1}H and ^{13}C NMR spectra are reported in ppm with a solvent resonance as an internal standard (^{1}H NMR: $CHCl_3$ 7.24 ppm, CHD_2OD 3.30 ppm; ^{13}C NMR: $CDCl_3$ 77.0 ppm, CD_3OD 49.0 ppm). Coupling constants (*J*) are reported in hertz (Hz). Splitting patterns are indicated as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, dd = doublet of doublet. **High-performance liquid chromatography** (HPLC-DAD) was performed on a HP 1100 and on a 1260 Infinity chromatography system (Agilent), respectively. Enantiomeric excesses were determined using the columns Diacel Chiralcel OB 10 µm (250 mm × 4.6 mm), Diacel Chiralcel OD-H 5 µm (250 mm × 4.6 mm) and Chiralpac IC 5 µm (250 mm × 4.6 mm). **Circular dichroism** (CD) spectra were recorded in order to determine absolute configuration of the products using a Jasco J-810 spectrometer (Jasco International Co).

Electronic Supplementary Material (ESI) for Chemical Communications

2. Structural investigation of *Ec*MenD and docking experiments

Thiamine diphosphate (ThDP)-dependent enzymes are well known for their catalytic potential to form various chiral α-hydroxy ketones. Most of the wild type (wt) enzymes are (*R*)-selective in the carboligation reaction, whereby the access to (*S*)-α-hydroxy ketones is limited. Mainly steric properties of the active site of ThDP-dependent enzymes are responsible for the stereoselectivity. Prerequisite for (*S*)-selectivity is a structural element called an *S*-pocket, which allows an antiparallel arrangement of donor and acceptor substrate prior to carboligation.[1] *S*-pockets are present in most ThDP-dependent enzymes, however, not accessible for acceptor substrates due to large amino acid side chains.[2,3]

The carboligation potential of MenD from *Escherichia coli* (*Ec*MenD) was recently investigated in detail concerning substrate diversity and stereoselectivity. Carboligation of α-ketoglutaric acid (**1**) as donor with different aromatic aldehydes results in the formation of α-hydroxy ketones with high enantiomeric excesses (*ee*) of >93% (*R*).[4] To explain the high (*R*)-selectivity of *Ec*MenD, the structure (pdb: 2JLC)[5] was investigated concerning important amino acid residues in the *S*-pocket region in the first step using the PyMOL software (Schrödinger). Five residues were identified in the *S*-pocket region: P30, G31, S32, I474, and F475 (Fig. S1 A). Docking of benzaldehyde (**2a**) revealed that obviously only a parallel arrangement of **2a** and **1** is possible meaning that both side chains are oriented towards the substrate channel. An antiparallel arrangement of **2a** would lead to a clash with the amino acid side chains of the *S*-pocket region (Fig. S1 B) explaining the high (*R*)-selectivity in the reaction of **1** and **2a** (*ee* = 99%).

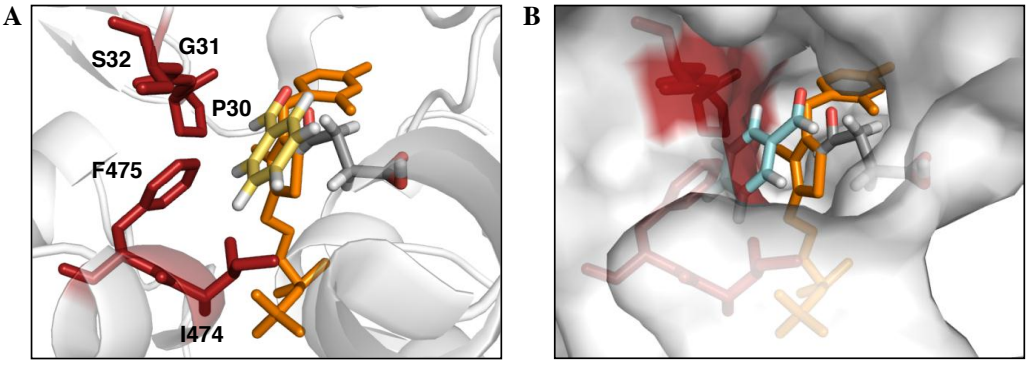


Fig. S1: Amino acids (red) defining the *S*-pocket of *Ec*MenD (**A**) and the resulting structure of the acceptor-binding site (**B**). Only a parallel arrangement of the acceptor benzaldehyde (yellow) to the donor α-ketoglutaric acid (gray) is possible. An antiparallel arrangement of benzaldehyde (cyan) would lead to a clash with the amino acids of the *S*-pocket region. The cofactor ThDP is colored in orange.

Electronic Supplementary Material (ESI) for Chemical Communications

Among the five amino acids of the *S*-pocket region, I474 and F475 were deduced to be crucial for stereoselectivity. Both amino acids prevent the antiparallel arrangement of **2a** prior to C-C-bond formation (Fig. S2 A). However, a replacement of both with the smaller amino acids glycine and alanine, respectively, opens the *S*-pocket for the phenyl ring of **2a** (Fig. S2 B, exemplarily shown for variant I474A/F475G) allowing the antiparallel arrangement of donor and acceptor substrate. The new resulting *S*-pocket should be able to stabilize **2a**, thus increasing the formation of the respective (*S*)-product. The experimental data (see Communication) support this model. Stereoselectivity of *Ec*MenD could be inverted by mutation of I474 and F475 to glycine and alanine, respectively, yielding the (*S*)-selective variants I474G/F475G, I474A/F475G, I474G/F475A, and I474A/F475A with *ee*s of 27 – 75% (*S*) in the reaction of **1** and **2a**.

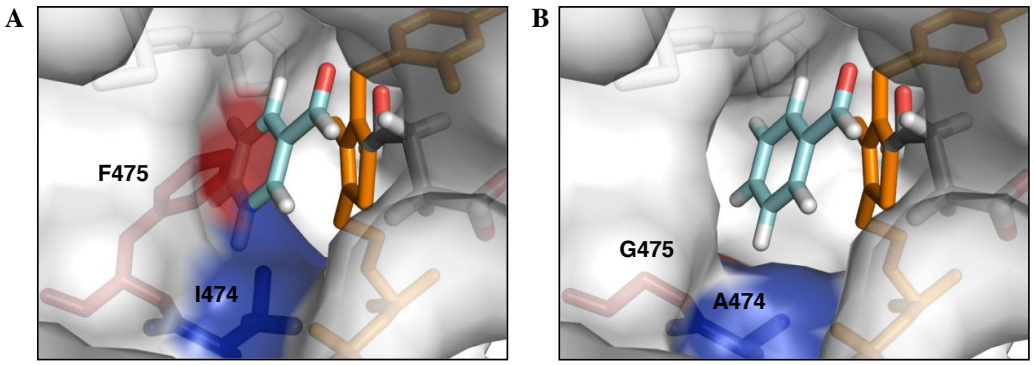


Fig. S2: The antiparallel orientation of benzaldehyde modeled inside the active side of *Ec*MenDwt (**A**) and *Ec*MenDI474A/F475G (**B**). The exchange of I474 (blue) to alanine and F475 (red) to glycine allows the antiparallel arrangement of benzaldehyde (cyan) to the donor α-ketoglutaric acid (gray). The cofactor ThDP is colored in orange.

Electronic Supplementary Material (ESI) for Chemical Communications

3. Construction, expression and purification of *Ec*MenD variants

The original *menD* gene of *Escherichia coli* K12 was cloned into pET-19b (Novagen) providing an N-terminal His_{10}-tag.[4] Site-directed mutagenesis was performed using the standard QuikChange® protocol from Stratagene. However, the KOD Hot Start DNA Polymerase (Merck Millipore) was used in the PCR reaction. First, *Ec*MenDF475G was prepared using the forward (fw) 5'-CAACAACGGCGGGCAAATTggCTCGCTGTTGC-3' and reverse (rv) primer 5'-GCAACAGCGAGccAATTTGCCCGCCGTTGTTG-3'. The mutated codons are underlined, with lower-case letters indicating base changes relative to the template. Based on *Ec*MenDF475G the double mutants *Ec*MenDI474G/F475G and *Ec*MenDI474A/F475G were prepared. *Ec*MenDI474A/F475G served as a template for the construction of *Ec*MenDI474A/F475A, whereas *Ec*MenDI474G/F475A was prepared based on *Ec*MenDI474G/F475G. The following fw and rv primers were used for mutagenesis:

*Ec*MenDI474G/F475G: 5'-CAACAACGGCGGGCAAggcGGCTCGCTGTTGC-3' (fw)
5'-GCAACAGCGAGCCcggTTGCCCGCCGTTGTTG-3' (rv)

*Ec*MenDI474A/F475G: 5'-CAACAACGGCGGGCAAgcgGGCTCGCTGTTGCC-3' (fw)
5'-GGCAACAGCGAGCCcgcTTGCCCGCCGTTGTTG-3' (rv)

*Ec*MenDI474G/F475A: 5'-CAACGGCGGGCAAGGCgcCTCGCTGTTGCCAAC-3' (fw)
5'-GTTGGCAACAGCGAGgcGCCTTGCCCGCCGTTG-3' (rv)

*Ec*MenDI474A/F475G: 5'-CAACGGCGGGCAAGCGGcgTCGCTGTTGCCAAC-3' (fw)
5'-GTTGGCAACAGCGAcgCCGCTTGCCCGCCGTTG-3' (rv).

Following mutagenesis, the template DNA was digested by *DpnI*. Finally, *E. coli* BL21(DE3) was transformed with the remaining PCR products. Gene sequences were confirmed by DNA sequencing (LGC Genomics).

Expression and purification of *Ec*MenD and its variants was carried out with minor modifications as described elsewhere.[4] After 14 h of cultivation at 120 rpm and 20 °C cells were pelleted by centrifugation at 8000 rpm and 4 °C for 30 min using Avanti J-20 XP, rotor JLA-8.1000 (Beckman Coulter). The cell pellet was resuspended in lysis buffer (20% w/v) containing lysozyme (1 mg/mL). After cell disruption by sonication cell debris was harvested by centrifugation at 20.000 rpm and 4 °C for 45 min using rotor JA-20 (Beckman Coulter). Purification of His-tagged all *Ec*MenD variants was performed by immobilized nickel chelate chromatography. The resulting enzyme was first desalted by size exclusion chromatography using Sephadex™ G25M (GE Healthcare) and finally freeze-dried using ALPHA 2-4

Electronic Supplementary Material (ESI) for Chemical Communications

(Christ). Lyophilized *Ec*MenD variants were stored at -20 °C for several months without significant loss of activity.

4. Reaction conditions for carboligation reactions with *Ec*MenD variants

Lyophilized *Ec*MenD variants were used for all preparations. The protein concentration was determined according to Bradford.[6] Reactions were performed in phosphate buffer (50 mM potassium phosphate, 0.1 mM ThDP, 2 mM $MgCl_2 \cdot 6H_2O$, pH 8.0) at 30 ºC and 300 rpm using a Thermomixer (Eppendorf). The final concentration of MenD was set to 700 µg/mL. 50 mM α-ketoglutaric acid and 20 mM benzaldehyde derivatives were incubated with enzyme in 1.5 mL reaction buffer containing 5% (v/v) MTBE for 24 h. Negative control experiments without enzyme were performed to verify enzyme-catalyzed C-C bond formation.

5. Analytical procedures

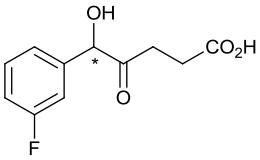
For GC-MS analysis 100 µL of the reaction mixture was acidified with 5 µL of 6 M HCl and extracted with 200 µL of ethyl acetate. For NMR analysis the complete reaction mixture of two 1.5 mL approaches was first extracted with 50 µL of 6 M HCl and then three times with 300 µL methyl *tert*-butyl ether (MTBE). Afterwards, MTBE was evaporated and the residual mixture was dissolved in 100 µL n-hexane/2-propanol (50:50). The products were purified by semi-preparative HPLC on a Reprosil Chiral-OM 5 µm (250 mm × 10 mm) column using the following conditions: n-hexane/2-propanol = 90:10 (0.25% trifluoroacetic acid (TFA)), 0.75 mL min^{-1} and 20 °C. Finally, the organic phase was dried over Na_2SO_4 and the solvent was evaporated under reduced pressure. NMR analyses were performed in $CDCl_3$. Determination of the enantiomeric excess was carried out on chiral phase HPLC. 200 µL of the reaction mixture was extracted with 10 µL 6 M HCl and 200 µL of MTBE. CD spectroscopy was performed in order to confirm the absolute configuration of the product. (*R*)-Phenylacetylcarbinol was used as reference compound. The reaction mixture was directly diluted 1:10 in ddH_2O and measured in a 2 mm cuvette. The spectra were recorded in the range of 240 - 350 nm. A blank buffer was measured to ensure that no background signal was present. (*R*)-PAC derivatives led to the appearance of a negative CD band with the maximum near 280 nm, whereas (*S*)-PAC derivatives led to the respective positive CD signal.

Electronic Supplementary Material (ESI) for Chemical Communications

6. Characterization data for selected compounds

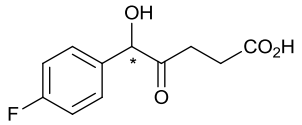
Complete characterizations of compounds **3a**, **b**, and **e - g** are given in Kurutsch *et al*. (2009) and Beigi *et al*. (2012). New compounds are described in the following.

5-(3-Fluorophenyl)-5-hydroxy-4-oxopentanoic acid **(3c)**:[4]



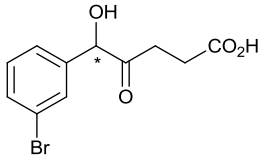
$C_{11}H_{11}FO_4$, M_r 226.20; ^{1}H NMR ($CDCl_3$): δ = 2.48–2.62 (m, 2H), 2.67–2.79 (m, 2H), 5.16 (s, 1H), 7.02–7.07 (m, 2H), 7.15 (d, 1H, J = 7.7 Hz), 7.34–7.37 (m, 2H) ppm; ^{13}C NMR ($CDCl_3$): δ = 27.5 (CH_2), 32.4 (CH_2), 79.1 (CH), 114.3 (d, CH, J = 22 Hz), 115.9 (d, CH, J = 21 Hz), 123.1 (d, CH, J = 3 Hz), 130.6 (d, CH, J = 8 Hz), 140.2 (C), 163.1 (d, CF, J = 248 Hz), 176.8 (COOH), 207.1 (C=O) ppm; HPLC-DAD: (Daicel Chiralpak IC 5 µm, 20 ºC, 1.2 mL min^{-1}: *n*-hexane/2-propanol (0.25% TFA) = 94.5:5.5): $t_R(R)$ = 26.6 min, $t_R(S)$ = 25.4 min.

5-(4-Fluorophenyl)-5-hydroxy-4-oxopentanoic acid **(3d)**:[4]



$C_{11}H_{11}FO_4$, M_r 226.20; ^{1}H NMR ($CDCl_3$): δ = 2.45–2.75 (m, 4H), 5.15 (s, 1H), AA'BB' spin system: 7.05-7.11 (m, BB', 2H, J_{HH} = $J_{HF(ortho)}$ =8.6 Hz), 7.29-7.34 (m, AA', 2H, J_{HH} = 8.6, $J_{HF(meta)}$ = 5.2 Hz) ppm; ^{13}C NMR ($CDCl_3$): δ = 27.5 (CH_2), 32.3 (CH_2), 79.0 (CH), 116.0 (d, CH, J = 22 Hz), 129.2 (d, CH, J = 8 Hz), 133.5 (d, C, J = 3 Hz), 163.0 (d, CF, J = 248 Hz), 176.9 (COOH), 207.5 (C=O) ppm; HPLC-DAD: (Daicel Chiralcel OD-H 5 µm, 20 ºC, 0.75 mL min^{-1}: *n*-hexane/2-propanol (0.25% TFA) = 95:5): $t_R(R)$ = 41.6 min, $t_R(S)$ = 44.5 min.

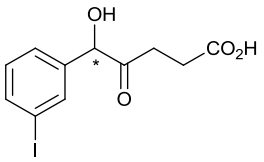
5-(3-Bromophenyl)-5-hydroxy-4-oxopentanoic acid **(3h)**:



$C_{11}H_{11}BrO_4$, M_r 287.11; ^{1}H NMR ($CDCl_3$): δ = 2.47–2.63 (m, 2H), 2.67–2.78 (m, 2H), 5.12 (s, 1H), 7.24–7.30 (m, 2H), 7.46–7.52 (m, 2H) ppm; ^{13}C NMR ($CDCl_3$): δ = 27.5 (CH_2), 32.4 (CH_2), 79.1 (CH), 123.1 (CBr), 126.0 (CH), 130.4 (CH), 130.6 (CH), 132.0 (CH), 139.9 (C), 176.9 (COOH), 207.0 (C=O) ppm; GC-MS (EI): t_R = 10.2 min, *m/z* (%) 244 (6) $[M - CO_2]^+$, 242 (6) $[M - CO_2]^+$, 185 (31) $[C_7H_4BrO]^+$, 183 (100) $[C_7H_4BrO]^+$, 157 (31) $[C_6H_4Br]^+$, 155 (31) $[C_6H_4Br]^+$, 76 (20) $[C_6H_4]^+$, 57 (25) $[C_3H_5O]^+$; $t_{R(lactone)}$ = 13.2 min, *m/z* (%) 185 (100) $[C_7H_4BrO]^+$, 183 (100) $[C_7H_4BrO]^+$, 157 (25) $[C_6H_4Br]^+$, 155 (25) $[C_6H_4Br]^+$, 85 (63) $[C_4H_5O_2]^+$; HPLC-DAD: (Daicel Chiralpac IC 5 µm, 40 ºC, 1.2 mL min^{-1}: *n*-hexane/2-propanol (0.25% TFA) = 95.5:4.5): $t_R(R)$ = 38.9 min, $t_R(S)$ = 40.1 min.

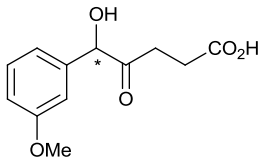
Electronic Supplementary Material (ESI) for Chemical Communications

5-Hydroxy-5-(3-iodophenyl)-4-oxopentanoic acid (**3i**):[4]



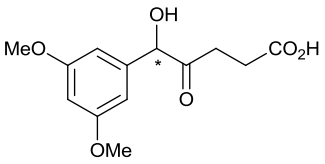
$C_{11}H_{11}IO_4$, M_r 334.11; HPLC-DAD: (Daicel Chiralpac IC 5 µm, 20 ºC, 1.2 mL min^{-1}: *n*-hexane/2-propanol (0.25% TFA) = 94.5:5.5): $t_R(R)$ = 28.1 min, $t_R(S)$ = 31.2 min.

5-Hydroxy-5-(3-methoxyphenyl)-4-oxopentanoic acid (**3j**):[4]



$C_{12}H_{14}O_5$, M_r 238.24; ^{1}H NMR ($CDCl_3$): δ = 2.51–2.59 (m, 2H), 2.66–2.76 (m, 2H), 3.80(s, 3H) 5.13 (s, 1H), 6.84–6.93 (m, 3H), 7.30 (t, 1H, *J* = 7.9 Hz) ppm; ^{13}C NMR ($CDCl_3$): δ = 27.5 (CH_2), 32.4 (CH_2), 55.3 (CH_3), 79.6 (CH), 112.6 (CH), 114.6 (CH), 119.8 (CH), 130.3 (CH), 139.2 (C), 160.1 (C), 177.0 (COOH), 207.6 (C=O) ppm; HPLC-DAD: (Daicel Chiralcel OB 10 µm, 20 ºC, 1.0 mL min^{-1}: *n*-hexane/2-propanol (0.25% TFA) = 95:5): $t_R(R)$ = 81.8 min, $t_R(S)$ = 96.8 min.

5-Hydroxy-5-(3,5-dimethoxyphenyl)-4-oxopentanoic acid (**3k**):



$C_{13}H_{16}O_6$, M_r 268.26; ^{1}H NMR ($CDCl_3$): δ = 2.52–2.60 (m, 2H), 2.65–2.79 (m, 2H), 3,78 (s, 1H), 5.07 (s, 1H), 6.43 (t, 1H, *J* = 2.2 Hz), 6.46 (d, 2H, *J* = 2.2 Hz) ppm; ^{13}C NMR ($CDCl_3$): δ = 27.5 (CH_2), 32.3 (CH_2), 55.4 (CH_3), 79.7 (CH), 100.8 (CH), 105.3 (CH), 139.9 (C), 161.3 (C), 177.1 (COOH), 207.5 (C=O) ppm; GC-MS (EI): t_R = 11.4 min, *m/z* (%) 222 (12) $[M - CO_2 - 2H]^+$, 165 (100) $[C_9H_9O_3]^+$, 137 (25) $[C_8H_9O_2]^+$, 122 (23) $[C_7H_6O_2]^+$, 107 (6) $[C_6H_3O_2]^+$, 92 (1) $[C_6H_4O]^+$, 77 (5) $[C_6H_5]^+$; $t_{R(lactone)}$ = 14.0 min, *m/z* (%) 250 (15) $[M]^+$, 165 (100) $[C_9H_9O_3]^+$, 137 (19) $[C_8H_9O_2]^+$, 122 (12) $[C_7H_6O_2]^+$, 107 (6) $[C_6H_3O_2]^+$, 85 (6) $[C_4H_5O_2]^+$, 77 (5) $[C_6H_5]^+$; HPLC-DAD: (Daicel Chiralpac IC 5 µm, 20 ºC, 1.2 mL min^{-1}: *n*-hexane/2-propanol (0.25% TFA) = 85:15): $t_R(R)$ = 18.0 min, $t_R(S)$ = 25.6 min.

Electronic Supplementary Material (ESI) for Chemical Communications

7. References

[1] D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C.L. Berthold, G. Schneider, J.Pleiss, M. Müller, and M. Pohl, *ChemBioChem*., 2008, **9**, 406.

[2] D. Rother, G. Kolter, T. Gerhards, C. L. Berthold, E. Gauchenova, M. Knoll, J. Pleiss, M. Müller, G. Schneider and M. Pohl, *ChemCatChem*., 2011, **3**, 1587.

[3] M. Pohl, D. Gocke and M. Müller, P. T. Anastas, editor, *Handbook of Green Chemistry – Green Catalysis*, Wiley-VCH, 2008, **3**, 75.

[4] A. Kurutsch, M. Richter, V. Brecht, G. A. Sprenger and M. Müller, *J*. *Mol*. *Catal*. *B: Enzym*, 2009, **61**, 56.

[5] A. Dawson, P. K. Fyte and W. N. Hunter, *J*. *Mol*. *Biol*., 2008, **384**, 1353.

[6] M. M. Bradford, *Anal*. *Biochem*., 1976, **72**, 248.

[7] M. Beigi, S. Waltzer, L. Eggeling, G. A. Sprenger and M. Müller, Org. Lett., 2012, *in press*, *DOI* 10.1021/ol3031186.

The ^{1}H and ^{13}C NMR, chiral-phase HPLC, and CD spectra are not shown. The spectra can be looked up on: http://www.rsc.org/suppdata/cc/c3/c3cc38607d/c3cc38607d.pdf.

PUBLICATION III

Tailoring the *S*-selectivity of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) from *Escherichia coli*

Westphal, R.; Hahn, D.; Mackfeld, U.; Waltzer, S.; Beigi, M.; Widmann, M.; Vogel, C.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M.

ChemCatChem 2013,
5, 3587–3594
DOI: 10.1002/cctc.201300318
(http://onlinelibrary.wiley.com/doi/10.1002/cctc.201300318/full)

CHEMCATCHEM
FULL PAPERS

ChemPubSoc Europe

DOI: 10.1002/cctc.201300318

Tailoring the *S*-Selectivity of 2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate Synthase (MenD) from *Escherichia coli*

Robert Westphal,[a] Doris Hahn,[a] Ursula Mackfeld,[a] Simon Waltzer,[b] Maryam Beigi,[b] Michael Widmann,[c] Constantin Vogel,[c] Jürgen Pleiss,[c] Michael Müller,[b] Dörte Rother,[a] and Martina Pohl*[a]

Dedicated to Professor Dr. Christian Wandrey on the occasion of his 70th birthday

The thiamine diphosphate (ThDP)-dependent enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase from *Escherichia coli* (*Ec*MenD, E.C. 2.2.1.9) catalyzes the carboligation of α-ketoglutarate (α-KG) and various benzaldehyde derivatives with excellent chemo- as well as high *R*-selectivity (enantiomeric excess (*ee*) >93%) to yield chiral α-hydroxy ketones. Based on the recently developed *S*-pocket concept, we engineered *S*-selective *Ec*MenD variants by optimizing the steric properties and stabilization of the acceptor substrate in the *S*-pocket. Moreover, the moderate *S*-selectivity of the *Ec*MenD variant I474A/F475G described recently for the carboligation of α-KG and benzaldehyde (*ee* = 75%) could be improved by selective destabilization of the *R*-pathway, which resulted in the variant I474A/F475G/R395Y (*ee* = 85% *S*). Subsequent investigation of the acceptor substrate range of this new variant revealed high *S*-selectivity especially with *meta*-substituted benzaldehydes, which gave access to 5-hydroxy-4-oxo-5-arylpentanoates with excellent enantioselectivities of up to 99% *ee S*. Thus, opening the *S*-pocket and simultaneous destabilization of the *R*-pathway provides a potential general new strategy to enhance the *S*-selectivity of ThDP-dependent enzymes.

Introduction

Catalytic asymmetric carbon–carbon bond formation is of major importance in organic synthesis to access chiral fine chemicals. Among them, α-hydroxy ketones as chiral building blocks are of particular value for the pharmaceutical as well as the fine chemistry sector.[1] Numerous nonenzymatic routes to chiral α-hydroxy ketones have been reported, although with limited stereoselectivity.[2,3]

Thiamine diphosphate (ThDP)-dependent lyases are well known for their catalytic potential to form various α-hydroxy ketones.[4] A broad range of α-hydroxy ketones with high enantiomeric excess (*ee*) values is currently available because of their different substrate preferences.[5,6] However, wild type (wt) enzymes are predominantly *R*-selective for mixed benzoin condensation-like carboligations. Thus, direct access to (*S*)-α-hydroxy ketones is limited.

An in-depth investigation of structure–function relationships in benzaldehyde lyase (BAL) from *Pseudomonas fluorescens* (*Pf*BAL) and benzoylformate decarboxylase (BFD) from *Pseudomonas putida* (*Pp*BFD) elucidated the underlying principles of chemo- and stereoselectivity on a molecular level.[7,8] *Pp*BFD shows uncommon *S*-selectivity for the formation of 2-hydroxypropiophenone (*ee* = 92% *S*) from benzaldehyde and acetaldehyde as substrates.[9] In contrast, *Pf*BAL is strictly *R*-selective for the same reaction (*ee* >99%).[10] A comparison of both enzymes on a structural level revealed the rationale behind the stereoselectivity. A structural element called the *S*-pocket was identified in *Pp*BFD, which allows the antiparallel arrangement of the benzaldehyde donor and the acetaldehyde acceptor prior to carboligation, a prerequisite for *S*-selectivity.[8] In turn, *R*-selectivity results from the parallel orientation of both substrates.

S-pockets are present in many *R*-selective ThDP-dependent enzymes, however, often they are inaccessible to acceptor substrates because of large amino acid side chains.[4] Recently, the *S*-pocket concept was successfully transferred for the first time to the *R*-selective pyruvate decarboxylase (PDC) from *Acetobacter pasteurianus* (*Ap*PDC) by mutation of the key residue E469 to glycine. Thus, the *S*-pocket could be opened for benzaldehyde to give access to different (*S*)-phenylacetylcarbinol

[a] R. Westphal, D. Hahn, U. Mackfeld, Dr. D. Rother, Prof. M. Pohl
Forschungszentrum Jülich GmbH
IBG-1: Biotechnology
52425 Jülich (Germany)
Fax: (+49) 02461-61-3870
E-mail: ma.pohl@fz-juelich.de

[b] S. Waltzer, Dr. M. Beigi, Prof. M. Müller
Institute of Pharmaceutical Sciences
Albert-Ludwigs-University Freiburg
Albertstr. 25, 79104 Freiburg (Germany)

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

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cctc.201300318.

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

(PAC) derivatives.[11] As a result we increased the enzymatically accessible α-hydroxy ketone platform by new (*S*)-α-hydroxy ketones.

Recently, the carboligation potential of a subgroup of ThDP-dependent enzymes, which use α-ketoglutarate (α-KG) as the physiological donor, was explored.[12] The use of α-KG in carboligation reactions enables selective C4 chain elongation with a terminal carboxyl group to yield chiral δ-hydroxy-γ-keto acids, which represent valuable functionalized α-hydroxy ketones. These products are potential precursors of γ- and δ-lactones,[13] themselves valuable intermediates in organic synthesis.[14] The enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase from *Escherichia coli* (*Ec*MenD, E.C. 2.2.1.9) belongs to this group of ThDP-dependent enzymes. Its physiological activity is the catalysis of a Stetter-like 1,4-addition of α-KG to isochorismate, which represents the second step of menaquinone biosynthesis.[15] Furthermore, *Ec*MenD catalyzes the 1,2-addition of α-KG and a broad range of acceptor substrates that include aliphatic, aromatic, and α,β-unsaturated aldehydes. These carboligation reactions are characterized by excellent chemo- and enantioselectivities.[12,16] However, like many of the ThDP-dependent enzymes, *Ec*MenD is strictly *R*-selective. Recently, we communicated the first results on *S*-selective *Ec*MenD variants by tailoring its *S*-pocket and entrance through mutagenesis of F475 and I474 to glycine and alanine, respectively.[17] If we used α-KG as the donor and different benzaldehyde derivatives as acceptors, we could access various (*S*)-5-hydroxy-4-oxo-5-arylpentanoates with good enantioselectivity. The highest conversions and enantioselectivities (up to 97% *ee*) were obtained with variant I474A/F475G, particularly with *meta*-substituted benzaldehyde derivatives (Table 1).

Here, we report on a new concept to further improve the *S*-selectivity of *Ec*MenD based on targeted mutagenesis of amino acid residues, which are involved in the stabilization of the *R*-pathway.

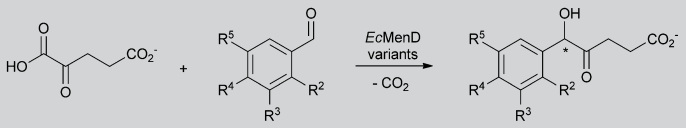
Table 2. Amino acid residues investigated in the active site of *Ec*MenD.

Amino acid residue (Standard no.)[20]	Assumed role for enantioselectivity
S32 (28)	influencing the *S*-pocket geometry
I474 (476)	blocking the entrance of the *S*-pocket
F475 (477)	blocking the *S*-pocket
L478 (480)	stabilizing the parallel orientation of benzaldehyde
R395 (393)	stabilizing the parallel orientation of benzaldehyde

Results and Discussion

The stereoselectivity of ThDP-dependent enzymes is mainly controlled by the steric properties of the active site. In addition, optimal stabilization of the acceptor substrate prior to carboligation is essential to gain high selectivity.[11,18] To support and extend the existing *S*-selectivity concept,[8] we systematically investigated the factors that control enantioselectivity in *Ec*MenD by using the 1,2-addition of α-KG to benzaldehyde as a model reaction. Therefore, the roles of five amino acid residues in the active site of *Ec*MenD were investigated with respect to their influence on the steric properties as well as their specific interaction with the acceptor benzaldehyde (Table 2). The resulting *Ec*MenD variants were characterized only by the *ee* of the product 5-hydroxy-4-oxo-5-phenylpentanoate (**1a**). Although the enantiomeric ratio *E* (k_{cat} *R*/k_{cat} *S*) offers a concise and valuable representation of the enantioselective properties of an enzyme,[19] this

Table 1. Carboligation activity of *Ec*MenD_wt, *Ec*MenD_I474A/F475G, and *Ec*MenD_I474A/F475G/R395Y using α-KG and benzaldehyde derivatives.[a]



Product	R^2	R^3	R^4	R^5	wt[d] Conversion [%][b]	*ee* [%][c]	I474A/F475G[e] Conversion [%][b] (Isolated yield [%][g])	*ee* [%][c]	I474A/F475G/R395Y Conversion [%][b] (Isolated yield [%][g])	*ee* [%][c]
1a	H	H	H	H	>99	>99 (*R*)	15	75 (*S*)	7	85 (*S*)
1b	F	H	H	H	>99	94 (*R*)	55	48 (*S*)	9	57 (*S*)
1c	H	F	H	H	>99	96 (*R*)	43	82 (*S*)	13	87 (*S*)
1d	H	H	F	H	>99	99 (*R*)	14	46 (*S*)	3	n.d.[f]
1e	Cl	H	H	H	98	93 (*R*)	57	≤5	21	9 (*S*)
1f	H	Cl	H	H	>99	96 (*R*)	94	89 (*S*)	38 (65)[h]	96 (*S*)
1g	H	H	Cl	H	>99	93 (*R*)	15	48 (*R*)	4	n.d.
1h	Br	H	H	H	82	80 (*R*)	52	≤5	24	≤5
1i	H	Br	H	H	99	99 (*R*)	87 (63)	97 (*S*)	29	>99 (*S*)
1j	H	H	Br	H	>99	96 (*R*)	17	69 (*R*)	4	n.d.
1k	I	H	H	H	40	76 (*R*)	69	≤5	20	15 (*R*)
1l	H	I	H	H	>99	94 (*R*)	86	93 (*S*)	14	96 (*S*)
1m	OCH_3	H	H	H	98	98 (*R*)	13	51 (*S*)	2	n.d.
1n	H	OCH_3	H	H	99	>99 (*R*)	64	93 (*S*)	21	96 (*S*)
1o	H	OCH_3	H	OCH_3	99	98 (*R*)	83 (76)	96 (*S*)	12	98 (*S*)

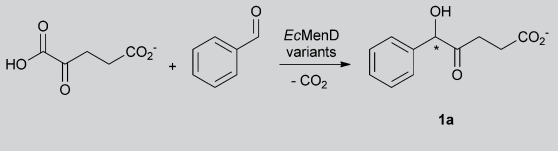
[a] Reaction conditions: 50 mM potassium phosphate buffer, pH 8.0, containing 2 mM $MgSO_4$, 0.1 mM ThDP and 5 vol% MTBE; 0.70 mg mL^{-1} *Ec*MenD; 50 mM α-KG and 20 mM benzaldehyde derivative; 30 °C; 24 h. [b] Determined by GC–MS based on benzaldehyde consumption. [c] Determined by chiral-phase HPLC. [d] Recently reported, except for **1h**, **1j**, **2k**, and **1m**.[17] [f] n.d. = not determined. [g] Preparative synthesis at 15 mL scale using 1 mg mL^{-1} of enzyme. For reaction conditions see [a]. [h] In the preparative synthesis a conversion of 90% was obtained after 24 h owing to higher enzyme concentration.

parameter could not be used for the characterization of *Ec*MenD variants because of limited access to reliable kinetic parameters such as V_{max}, the maximum rate achieved by the system, at maximum (saturating) substrate concentrations, or K_m, the substrate concentration at which the reaction rate is half of V_{max}. In addition, the activity of *Ec*MenD in the nonphysiological 1,2-addition is low, which limits the determination of low conversions at low substrate concentrations under initial rate conditions. Furthermore, the solubility limit of the aromatic substrates prevents single-substrate kinetics under saturated conditions.

Tailoring the *S*-pocket of *Ec*MenD

In the first step, the *S*-pocket of *Ec*MenD was opened by mutation of F475 to glycine. F475 corresponds to E469 in *Ap*PDC, which was identified as a key residue for stereoselectivity.[11] This position corresponds to number 477 according to the recently published standard numbering system for this enzyme class (Table 2).[20] In contrast to *Ap*PDC_E469G, for which stereoselectivity was inverted by only one amino acid exchange, the F475G mutation did not affect the stereoselectivity of *Ec*MenD to a great extent (Table 3). However, the reduced *ee* of

Table 3. Enantioselectivity of *Ec*MenD_wt and variants for the carboligation of α-KG and benzaldehyde to **1a**.[a]



1a

*Ec*MenD or variants	Conversion [%][b]	*ee* [%][c]
wt[d]	>99	>99 (*R*)
F475G	96	83 (*R*)
I474G	77	70 (*R*)
I474G/F475G[d]	26	34 (*S*)
I474G/F475A[d]	67	27 (*S*)
I474G/F475V	69	61 (*R*)
I474A/F475G[d]	15	75 (*S*)
I474V/F475G	44	86 (*R*)
I474A/F475A[d]	68	66 (*S*)

[a] Reaction conditions: 50 mM potassium phosphate buffer, pH 8.0, containing 2 mM $MgSO_4$, 0.1 mM ThDP and 5 vol% MTBE; 0.70 mg mL^{-1} enzyme; 50 mM α-KG and 20 mM benzaldehyde; 30 °C; 24 h. [b] Determined by GC–MS based on benzaldehyde consumption. [c] Determined by chiral-phase HPLC. [d] Recently reported.[17]

83% *R* for **1a** indicated that the *S*-pocket was at least partially accessible to the phenyl ring of benzaldehyde. The still predominantly *R*-selective synthesis of **1a** could be because of I474 at the entrance of the *S*-pocket, which mainly prevents the antiparallel orientation of benzaldehyde. Therefore, I474 was also exchanged by glycine. Although variant I474G still revealed *R*-selectivity (*ee* = 70%), the combination of both mutations resulted in the *S*-selective synthesis of **1a** (*ee* = 34%).[17]

The low *S*-selectivity of variant I474G/F475G indicated that the antiparallel-oriented benzaldehyde acceptor is still not optimally stabilized in the *S*-pocket relative to the parallel orientation. To improve this stabilization, positions 474 and 475 were mutated to alanine and valine, respectively. The reduction of the *S*-pocket size in position 475 did not improve the *S*-selectivity for **1a** (Table 3). Variant I474G/F475A revealed almost no change in *S*-selectivity for the formation of **1a** (*ee* = 27%),[17] whereas variant I474G/F475V was again *R*-selective (*ee* = 61%). In contrast, the exchange of glycine to alanine in position 474 improved the *S*-selectivity significantly. Variant I474A/F475G showed an *ee* of 75% for (*S*)-**1a**,[17] which can be explained by the improved stabilization of the phenyl ring of benzaldehyde by nonpolar interactions as a result of the additional methyl group of A474 (Figure 1A,B). Valine in position 474 blocks the

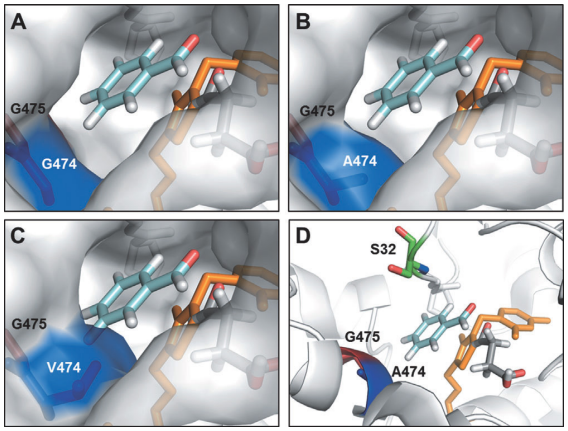


Figure 1. *S*-pocket of *Ec*MenD variants I474G/F475G, I474A/F475G, and I474V/F475G with modeled substrates: succinylsemialdehyde (gray) bound to C2-ThDP and benzaldehyde (cyan). ThDP is shown in orange. Compared to I474G/F475G (A), the *S*-pocket size of I474A/F475G (B) is smaller because of the methyl group of alanine in position 474 (blue region), which might lead to better stabilization of the antiparallel benzaldehyde orientation. Furthermore, reduction of the *S*-pocket size by valine in position 474 seems to block the entrance of the *S*-pocket again (C). S32 restricts the *S*-pocket size from the top (D).

entrance of the *S*-pocket again (Figure 1C) to result in high *R*-selectivity (*ee* = 86%). By contrast, the combination of alanine in both positions slightly reduced the *S*-selectivity and led to 65% *ee*.[17] The mutation to alanine in position 475 may already result in minor unfavorable steric interactions with benzaldehyde, which thus reduces *S*-selectivity. This result is consistent with the lower *S*-selectivity of variant I474G/F475A relative to I474G/F475G.

These results support the *S*-pocket concept developed previously.[8,11,18] The influence of steric effects could be proven by both valine variants, which revealed again *R*-selectivity because of the blocked *S*-pocket. Furthermore, the results of *S*-pocket engineering impressively demonstrated the importance of acceptor stabilization in the *S*-pocket as a prerequisite for high *S*-selectivity. Accordingly, the antiparallel benzaldehyde orienta-

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

tion is best stabilized in the *S*-pocket of variant I474A/F475G (Figure 1 B).

The catalytic activities of all variants, which may be estimated in a first approach from the conversions observed after 24 h (Table 3), were lower than that of *Ec*MenD_wt. Although variants I474G, F475G, I474G/F475A, I474G/F475V, and I474A/F475A still showed good conversions of 67–96% for **1 a**, we obtained only 15–44% conversion with variants I474G/F475G, I474A/F475G, and I474V/F475G under the tested conditions. As both amino acid residues are potentially involved in the binding of the cofactor ThDP,[21] the observed differences in conversion might also result in different stabilities of the variants because of reduced cofactor-binding stability. Moreover, the mutations probably influence the binding affinity for the substrates. Dawson et al.[22] demonstrated recently that the corresponding mutations of I489 and F490 in *Bacillus subtilis* MenD increased the K_m value for the donor α-KG as well as for the acceptor isochorismate by a factor of 12–25 concomitant with a reduction of catalytic efficiency. Consequently, the lower conversion obtained with *Ec*MenD variants might be a result of the reduced binding affinity of α-KG as well as that of benzaldehyde. Glycine in position 475 seems to have the higher impact on conversion, if both positions are mutated.

Optimization of the *S*-pocket geometry of *Ec*MenD_I474A/F475G

One option to enhance the *S*-selectivity of the best variant I474A/F475G (*ee* = 75%, **1 a**, Table 2) is to improve the stabilization of the antiparallel-oriented benzaldehyde by further optimization of the *S*-pocket geometry.

The only remaining residue in the *S*-pocket of *Ec*MenD_I474A/F475G that might influence the antiparallel orientation of benzaldehyde is S32, which is located on a flexible loop that restricts the *S*-pocket size from the top (Figure 1 D). S32 is supposed to play an important role in binding and/or the reaction of isochorismate.[23] To study the relevance of position 32, we chose site-saturation mutagenesis using NDT codon degeneracy. Therefore, an analytical system using chiral-phase HPLC was established to allow reliable screening for the enantioselectivity of biotransformations in microtiter plates using crude cell extracts (for details see Experimental Section). Additionally, variant I474A/F475G/S32A, which was not included in the NDT codon library, was prepared separately. The methyl group in the alanine side chain might lead to a better stabilization of benzaldehyde compared to the hydroxyl group of serine, although the OH group does not point into the direction of the phenyl ring (Figure 1 D).

Screening of the complete library, verified by sequencing, revealed no improved *Ec*MenD variants (Table 4). The only variant with still high *S*-selectivity for **1 a** was I474A/F475G/S32C (*ee* = 65%). The larger van der Waals volume of cysteine compared to serine might already result in higher steric restrictions in the *S*-pocket for benzaldehyde, which results in slightly reduced enantioselectivity. All other amino acid exchanges, including S32A, yielded either unselective or predominantly *R*-selective variants. As a result of the location of position 32 in a flexible loop, the interpretation of these results without structural information of such variants is not possible. Nevertheless, the results revealed that the given *S*-pocket geometry of I474A/F475G seems to be best for the *S*-selective synthesis of **1 a**.

Table 4. Carboligation activity of *Ec*MenD_I474A/F475G variants after single site-saturation mutagenesis of position 32 (S) and 478 (L), respectively.[a]

Position 32[b]	*ee* [%][c]	Position 478[b]	*ee* [%][c]
C	65 (*S*)	I	43 (*S*)
A,[d] R, N, D, H, L, Y	30 (*S*), 30 (*R*)	F, Y, V	5–30 (*R*)
I, F, G, V	40–55 (*R*)	D	77 (*R*)
		G, H, N, R, S	no activity[e]

[a] In the reaction of α-KG and benzaldehyde to **1 a**. Screening setup and conditions are described in the Experimental Section. [b] Respective amino acid substitutions using NDT codon degeneracy. [c] Determined by chiral-phase HPLC. [d] Prepared separately. [e] Verified by crude cell extract data from shaking-flask cultivation.

Extending the concept of *S*-selectivity

In addition to the strategy to tailor the *S*-pocket and enhance the *S*-pathway either by optimal stabilization of the phenyl ring in the *S*-pocket or by improving access to this pocket, in principal, there is also an alternative strategy: the suppression of the *R*-pathway. As we could show that good stabilization of the acceptor substrate in an antiparallel orientation is required for high *S*-selectivity, destabilization of the parallel-oriented acceptor should also result in improved *S*-selectivity. To suppress the *R*-pathway, the amino acids of special interest are those that potentially stabilize the parallel orientation of benzaldehyde (and derivatives thereof). In *Ec*MenD these are R395 and L478 (Figure 2). L478 is part of a hydrophobic patch (together with I474 and F475) that interacts with the nonpolar part of isochorismate.[21] Thus, this residue is also supposed to stabilize the parallel orientation of benzaldehyde.

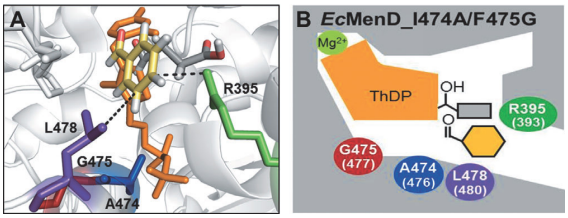


Figure 2. Potential stabilization of benzaldehyde in the parallel orientation in the active site of *Ec*MenD_I474A/F475G modeled with PyMOL (A) and as a schematic presentation (B). The parallel benzaldehyde orientation (yellow) is probably stabilized by L478 (purple) and R395 (green) putatively as a result of hydrophobic and cation–π interactions, respectively. Standard numbers of the residues for ThDP-dependent decarboxylases according to Vogel et al.[20] are given in parentheses.

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

R395 belongs to a group of four conserved arginine residues that are assumed to interact with the terminal carboxylate group of the donor α-KG (Figure 2).[21] Furthermore, R395 may stabilize the parallel orientation of benzaldehyde by cation–π interactions.[24] Both positions were subjected to site-saturation mutagenesis, again by using NDT codon degeneracy, based on variant I474A/F475G as a starting point.

Site-saturation mutagenesis of position 478 in *Ec*MenD_I474A/F475G

The introduction of small and polar amino acids, respectively, in position 478 should lead to reduced nonpolar interactions to decrease the stabilization of parallel-oriented benzaldehyde prior to C–C bond formation. However, screening of the complete library revealed that *S*-selectivity was reduced in all cases. With some mutations (e.g., L478D) the enantioselectivity was even inverted again. Moreover, five variants were inactive under the conditions tested (Table 4).

The results suggest that mutations at position 478 are not suitable to improve the *S*-selectivity. One problem might be the close proximity of L478 to the *S*-pocket (Figure 2), which thus influences its geometry. Furthermore, we cannot exclude that L478 contributes to the stabilization of the antiparallel benzaldehyde orientation as well. As shown in our model (Figure 3A), L478 is located in the middle of both possible benzaldehyde orientations. The distance of L478 to parallel- and antiparallel-oriented benzaldehyde is similar, 3.4 and 4.3 Å, respectively. In addition, biocatalysis on an analytical scale with variants in position 478 revealed a high percentage of inactive variants, which was verified by subsequent studies with crude cell extract, obtained from shaking-flask experiments. Although all variants were expressed as soluble proteins, as determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), no activity was measured for L478G, L478H, L478N, L478R, and L478S (Table 4), which confirmed the screening results obtained on an analytical scale. The high number of mutations in one α-helix might explain the high percentage of inactive variants. Like I474 and F475, L478 is located on the same short α-helix that consists of six amino acids (Figure 3B). Three mutations might lead to substantial structural rearrangements in the active site, especially in the acceptor binding region, which is also known to be involved in the stabilization of the cofactor ThDP.[21]

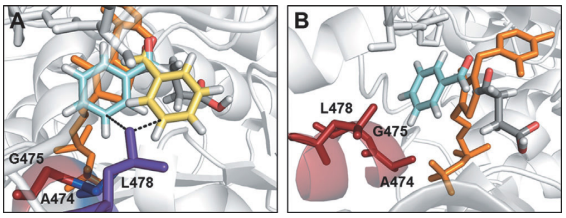


Figure 3. Role of L478 in the *S*-pocket region of *Ec*MenD_I474A/F475G. The model shows that L478 might contribute to the stabilization of the parallel (yellow) as well as the antiparallel orientation of benzaldehyde (cyan) (A). The distance of the terminal methyl group of L478 to the parallel and antiparallel benzaldehyde orientation (depicted as dashed line) is 3.4 and 4.3 Å, respectively. Moreover, L478 is located on a short α-helix (red) in the acceptor-binding region (B), which contributes to the stabilization of the cofactor ThDP.

Site-saturation mutagenesis of position 395 in *Ec*MenD_I474A/F475G

Amino acid R395, like L478, seemed to be a promising candidate for the stabilization of parallel-oriented benzaldehyde. In contrast to L478, R395 is positioned more distantly from the *S*-pocket (Figure 2B) and interacts with the acceptor substrate. Fang et al.[25] demonstrated that the mutation of R395 to alanine resulted in an increased K_m value for the physiological acceptor isochorismate. Furthermore, R395 might be able to form stabilizing cation–π interactions with the nonphysiological acceptor benzaldehyde. However, R395 might also play an important role in the stabilization of the donor α-KG (Figure 4A).[21] Consequently, mutagenesis in this position could be deleterious to the enzyme activity.

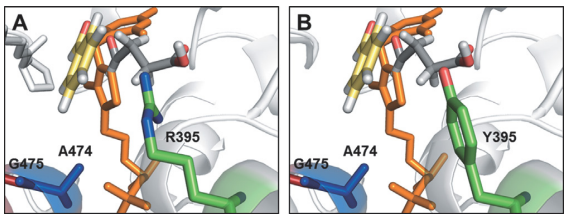


Figure 4. Influence of position 395 on the (de)stabilization of parallel-oriented benzaldehyde. R395 (green) could stabilize benzaldehyde (yellow) by the formation of cation–π interactions (A). In contrast, this stabilization is not possible with Y395 (green) (B).

As expected, screening of different variants in position 395 revealed predominantly variants with very low activity or inactive variants, respectively. Nevertheless, active variants were merely *S*-selective. However, two of these hits showed significantly improved *S*-selectivity compared to that of variant I474A/F475G. Sequence analyses of both hits yielded the same sequence: I474A/F475G/R395Y. Purified *Ec*MenD_I474A/F475G/R395Y revealed an *ee* of 85% *S* for **1a**, which, however, was accompanied by a low conversion of only 7% under the tested conditions (Table 1).

Indeed, tyrosine in position 395 has all of the required attributes to destabilize the parallel orientation of benzaldehyde (Figure 4B). Cation–π interactions with benzaldehyde are no longer possible. Furthermore, the hydroxyl group of tyrosine might additionally destabilize the parallel orientation of benzaldehyde to yield a higher *S*-selectivity, which supports our concept.

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

Substrate screening

As we have reported recently, the catalytic potential of *Ec*MenD variants I474G/F475G, I474A/F475G, and I474A/F475A was broadly investigated concerning the synthesis of diversely substituted (*S*)-5-hydroxy-4-oxo-5-arylpentanoates, which resulted in surprisingly high enantioselectivities for *meta*-substituted benzaldehydes as acceptors.[17] Thus, the acceptor spectrum of the new variant I474A/F475G/R395Y was investigated.

*Ec*MenD_I474A/F475G/R395Y revealed improved *S*-selectivity in all cases compared to the best *S*-selective variant so far I474A/F475G (Table 1). Again, the highest *S*-selectivities were observed with *meta*-substituted benzaldehydes, which resulted in excellent enantioselectivities (*ee* > 96%), except for **1c** (*ee* = 87% *S*). The best results were obtained with 3-bromobenzaldehyde, which was transformed to **1i** with > 99% *ee*. The differences in enantioselectivity between variants I474A/F475G/R395Y and I474A/F475G are 2–7%. However, for these high enantioselectivities, a large difference in activation free energy is required according to Equation (1),[26] which highlights the importance of the additional mutation in position 395.

$$\Delta G = -RT \ln[(100+ee) \times (100-ee)] \quad (1)$$

The results again support our extended model of *S*-selectivity, according to which destabilization of parallel-oriented benzaldehyde should result in a higher enantioselectivity. However, the additional mutation in position 395 led to reduced conversions (2–38%) in all reactions compared to variant I474A/F475G, and *meta*-substituted benzaldehydes gave the best results (Table 1). In addition to negative influences on the binding of α-KG, destabilization of the parallel-oriented acceptor might additionally influence the catalytic activity to a larger extent. All results were obtained without optimization of the reaction conditions, which is expected to further increase conversion rates as well as enantioselectivities.

Conclusions

With *Ec*MenD_I474A/F475G/R395Y we present the third example of the successful design of *S*-selective ThDP-dependent enzymes. The results are consistent with our previous studies on the *S*-pocket concept, which was supported by designing *S*-selective variants of *Pp*BFD and *Ap*PDC,[11,18] The results clearly demonstrate that enantioselectivity in this class of structurally related ThDP-dependent enzymes, which are summarized as the decarboxylase family,[27] is mainly controlled by steric effects. Sufficient space for the respective side chain of the acceptor aldehyde in the *S*-pocket is the first requirement to allow its antiparallel arrangement relative to the donor prior to carboligation. The better the stabilizing forces of the acceptor in the *S*-pocket the more the *S*-pathway is frequented and high enantioselectivities are achieved, which is in most cases is also accompanied by high conversions.

As the set of proteinogenic amino acids is limited, modulation of the acceptor size (substrate engineering) is additionally beneficial to increase this effect. For *Pp*BFD, we were able to tailor the pocket for larger aliphatic acceptor aldehydes by using benzaldehyde as the donor.[18] In the case of *Ec*MenD variants, *meta*-substituted benzaldehydes are clearly the best fitting acceptors if α-KG is used as the donor. Although the predictability of substrate ranges based on crystal structures is already quite good, we currently cannot explain the preference for *meta*-substituted benzaldehydes based on the structure of *Ec*MenD_wt. To elucidate this phenomenon, structural studies of *Ec*MenD variants are currently underway, which will be combined with subsequent molecular dynamics simulations. In contrast to these examples, the optimized enantioselectivity for the *S*-selective variant *Ap*PDC_E469G to give (*S*)-PAC derivatives was achieved by variation of the donor substrate range by using benzaldehyde as the acceptor. Although with acetaldehyde only a moderate enantioselectivity (*ee*) of 61% ((*S*)-PAC) was obtained, the stereoselectivity increased to 89% *ee* with propanal and to 85% *ee* with butanal.[11] The latter example demonstrates that variation of the donor may also significantly influence the population of the *S*- and *R*-pathways. In the case of *Ap*PDC, this could be explained by the steric hindrance of the *R*-pathway through the larger aliphatic side chain of propanal or butanal, respectively.

Such a donor variation is not possible with *Ec*MenD yet because α-KG is the only well-accepted donor.[16] Therefore, in the present study with *Ec*MenD, we have extended the concept of *S*-selectivity by the site-saturation mutagenesis of R395, an amino acid that influences the acceptor binding in the parallel orientation. Exchange by tyrosine was shown to be most effective to further increase the *S*-selectivity, probably because of a disturbed *R*-pathway. The combination of *S*-pocket engineering, suppression of the *R*-pathway, and substrate engineering gives access to highly *S*-selective asymmetric cross-benzoin condensations that start from readily available substrates.

Experimental Section

Materials

All chemicals were purchased from Sigma–Aldrich and used without further purification. KOD Hot Start DNA Polymerase was obtained from Merck Millipore. The restriction enzyme *DpnI* was purchased from Fermentas (Thermo Scientific).

Structural analysis

The PyMOL software[28] (Schrödinger) was used for the investigation of *Ec*MenD, generation of respective variants, and manual placement (docking) of the substrates. The acceptor benzaldehyde was docked by considering the reactive distance between the carbonyl group of benzaldehyde and the ThDP-bound succinic semialdehyde as well as the Bürgi–Dunitz angle, which defines the geometry of the nucleophilic addition to a carbonyl group.[29] All docking data were generated based on the crystal structure of *Ec*MenD (pdb: 2jlc).[21] The 3D structure of the substrates benzaldehyde and ThDP-bound succinic semialdehyde were generated by using the Avogadro software.[30] Amino acid positions were counted based on the sequence of *Ec*MenD_wt without an N-terminal His-tag (see Supporting Information).

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

Generation, expression, and purification of *Ec*MenD and variants

The original pET19b_*Ec*MenD construct that provides a 10×histidine tag at the N terminus was used as the template for site-directed mutagenesis.[16] All variants were generated by the standard QuikChange site-directed mutagenesis protocol (Stratagene) with minor modifications. To the standard reaction, $MgSO_4$ (2.5 mM, required by KOD Hot Start DNA Polymerase) and DMSO (5 vol%) were added. The cycling parameters for the PCR reaction were adjusted according to manufacturer information (Thermo Scientific). Mutagenic primers are given in the Supporting Information. Following mutagenesis, parental DNA was digested by *DpnI*. Afterwards, chemically competent *E. coli* BL21 (DE3) was transformed by the resulting PCR products by using the standard transformation protocol ($CaCl_2$ method) according to Sambrook and Russell.[31] Gene sequences were confirmed by DNA sequencing (LGC Genomics, Berlin). Expression and purification of *Ec*MenD and variants thereof was carried out as described previously.[16,17]

Site-saturation mutagenesis with *Ec*MenD_I474A/F475G

Library preparation

The mutant libraries for amino acid positions 32, 395, and 478, respectively, were prepared by using the QuikChange standard protocol (see above). Mutagenic primers that contained NDT degeneracy (in which N = any nucleotide, D = A, G, or T), which encodes for twelve amino acids (Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, and Val), are shown in the Supporting Information. The PCR products were again digested with *DpnI*. Subsequently, *E. coli* BL21 (DE3) was transformed by the resulting PCR products as described before.

Library screening

46 single colonies were transferred to a 96 deep-well plate (maximum volume of 2 mL; Nerbe Plus, Germany) filled with lysogeny broth medium (600 μL) that contained ampicillin (100 $\mu g\,\mu L^{-1}$). As a control, two wells were inoculated with a colony of *Ec*MenD_wt and *Ec*MenD_I474A/F475G, respectively. A screening of 46 colonies corresponds to 98% library completeness.[32] The master plate was sealed with a sterile oxygen sealing film (AeraSeal™, EXCEL Scientific, Inc., USA) and incubated overnight at 30 °C and 850 rpm in an INFORS Multitron Standard shaker (INFORS HT, Switzerland).

Overexpression of the variants was performed in an autoinduction medium that contained peptone (12 $g\,L^{-1}$), yeast extract (24 $g\,L^{-1}$), lactose (2 $g\,L^{-1}$), glucose (0.5 $g\,L^{-1}$), glycerin (0.5 vol%), and potassium phosphate buffer (9 vol%, 1 M, pH 7.0). 48 wells of a FlowerPlate (m2p-labs, Germany) were filled with autoinduction medium [1.5 mL; which contained ampicillin (100 $\mu g\,\mu L^{-1}$)] and then inoculated with the respective overnight culture (10 μL). The FlowerPlate was sealed again with the sterile oxygen-permeable sealing film and incubated at 20 °C and 1200 rpm for 48 h in the BioLector system (m2p-labs, Germany). The growth of each colony was monitored online. The master plate was mixed with sterile glycerin (500 μL), tightly closed with a silicon lid, and finally stored at −80 °C. To verify a working mutagenesis protocol, the completeness of the library was checked for position 32 and 478 by sequencing (LGC Genomics) of the whole plate.

The following steps were performed by the TECAN robot Freedom EVO 200 (TECAN, Switzerland). After cultivation, each culture (2×600 μL) was transferred into one 96 deep-well plate (maximum volume of 2 mL; Nerbe Plus, Germany) for duplicate analyses. Subsequently, cells were pelleted by centrifugation at 4000 rpm (Rotana 46RSC, Hettich, Germany) and 4 °C for 3 min. The supernatant was discarded and the pellets were frozen at −20 °C for at least 1 h for improved cell disruption. Afterwards, each pellet was resuspended in lysis buffer [620 μL; triethanolamine (50 mM, pH 8.0), which contained ThDP (0.1 mM), $MgSO_4$ (2 mM), and lysozyme (1 $mg\,mL^{-1}$)]. After incubation at 37 °C for 1 h, cell debris was pelleted by centrifugation at 4000 rpm and 4 °C for 10 min. Finally, the supernatant (470 μL) was transferred to a new 96 deep-well plate and assayed for carboligation.

For the carboligation reaction with *Ec*MenD variants, the respective supernatant was mixed with an α-KG stock solution (100 μL, 300 mM) and a benzaldehyde stock solution (30 μL, 400 mM in 100% methyl-*tert*-butyl ether (MTBE)) to give a final concentration of 20 mM benzaldehyde and 50 mM α-KG, respectively, in 5 vol% MTBE. Then, the plate was covered with a silicone cap-mat (Nerbe Plus, Germany) and incubated at 30 °C and 600 rpm (VARIOMAG Shaker, H+P Labortechnik, Germany) for 24 h.

Afterwards, the presence of α-hydroxy ketones was first analyzed by using the triphenyl tetrazolium chloride (TTC) assay.[33] Therefore, the reaction solution (100 μL) was transferred to a new microtiter plate (well volume: 330 μL, Nunc, Thermo Fisher Scientific, USA) and mixed with TTC solution (40 μL; TTC (10 mg) in 70% ethanol (2.5 mL) and NaOH (7.5 mL, 1 N)). Red coloration, which indicates product formation, was detected after 5 min. In case of positive hits, the product was extracted from the original carboligation reaction. Therefore, the residual volume (500 μL) was thoroughly mixed with perchloric acid (50 μL, 10% v/v) and MTBE (300 μL). Finally, the organic phase (150 μL) was transferred to HPLC vials and analyzed by using chiral HPLC (see below).

Verification of inactive variants

The inactivity of variants, obtained in the site-saturation mutagenesis of position 478, was verified by shaking-flask experiments. Therefore, autoinduction medium (50 mL; see above), which contained ampicillin (100 $\mu g\,\mu L^{-1}$), was inoculated with the respective variant (50 μL) directly from the master plate (see above). After incubation at 20 °C and 180 rpm for 48 h in an INFORS Multitron Standard shaker (INFORS HT, Switzerland), cells were pelleted at 4000 rpm and 4 °C for 30 min (Universal 32R, Hettich, Germany). The supernatant was discarded, and the pellets were resuspended in reaction buffer (triethanolamine (50 mM, pH 8.0), which contained ThDP (0.1 mM) and $MgSO_4$ (2 mM)). Then, the cells were disrupted by sonication (UP 200s, Dr. Hielscher, Germany). After removal of cell debris by centrifugation at 14000 rpm and 4 °C for 45 min (Microcentrifuge 5417R, Eppendorf, Germany), the supernatant was used for SDS-PAGE (to check the soluble protein fraction) and for the carboligation reaction. Therefore, 1, 2, and 3 $mg\,mL^{-1}$, respectively, of the crude extract was used in the reaction of α-KG (50 mM) and benzaldehyde (20 mM; see below). Product formation was checked by using the TTC-assay (see above) and chiral-phase HPLC (see below).

Carboligation reactions and analytical procedures

The mixed carboligation of α-KG and benzaldehyde derivatives was performed on an analytical scale (1.5 mL) as described elsewhere.[17] Complete characterization of the carboligation products, except for **1j**, were previously reported.[12,16,17] The identity of **1j**

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

was verified by GC–MS, NMR spectroscopy, and chiral-phase HPLC by using methods described elsewhere.[17]

5-(4-Bromophenyl)-5-hydroxy-4-oxopentanoic acid (**1 j**): $C_{11}H_{11}BrO_4$, M_r 287.11; ^{1}H NMR ($CDCl_3$): δ = 2.45–2.65 (m, 2H), 2.66–2.80 (m, 2H), 5.15 (s, 1H), 7.22–7.29 (m, 2H), 7.52–7.57 ppm (m, 2H); ^{13}C NMR ($CDCl_3$): δ = 27.5 (CH_2), 32.3 (CH_2), 79.1 (COH), 123.0 (CBr), 129.0 (CH), 132.2 (CH), 136.7 (C), 176.9 (COOH), 207.1 ppm (C=O); GC–MS (EI): t_R = 10.2 min, *m/z* (%) 244 (6) $[M-CO_2]^+$, 242 (6) $[M-CO_2]^+$, 185 (31) $[C_7H_4BrO]^+$, 183 (100) $[C_7H_4BrO]^+$, 157 (31) $[C_6H_4Br]^+$, 155 (31) $[C_6H_4Br]^+$, 76 (20) $[C_6H_4]^+$, 57 (25) $[C_3H_5O]^+$; $t_{R(lactone)}$ = 13.2 min, *m/z* (%) 185 (100) $[C_7H_4BrO]^+$, 183 (100) $[C_7H_4BrO]^+$, 157 (25) $[C_6H_4Br]^+$, 155 (25) $[C_6H_4Br]^+$, 85 (63) $[C_4H_5O_2]^+$; HPLC–DAD: (Daicel Chiralpac OD-H 5 μm, 20 °C, 0.75 mL min^{-1}: *n*-hexane/2-propanol 95:5 (0.25% TFA): t_RR = 47.9 min, t_RS = 53.7 min.

Representative examples of the enzymatic synthesis of (*S*)-5-hydroxy-4-oxo-5-arylpentanoates on a preparative scale

*Ec*MenD variants (1 mg mL^{-1} final concentration) were incubated in reaction buffer (15 mL; see above) with α-KG (50 mM), benzaldehyde derivative (20 mM, **1 f, 1 i, 1 o**) and 5% (v/v) MTBE (final concentration) at 30 °C and 400 rpm in a 15 mL falcon tube by using a Thermomixer (Eppendorf, Germany). After 24 h no more benzaldehyde derivative was detected (GC–MS) using *Ec*MenD_I474A/F475G. In the case of variant I474A/F475G/R395Y and the reaction of 3-chlorobenzaldehyde and α-KG, a conversion of 90% was obtained. The reaction mixtures were then acidified with HCl (750 μL, 5% v/v) and subsequently extracted three times with ethyl acetate (15 mL). The organic phase was dried over $MgSO_4$, and the solvent was evaporated. The products were purified by flash chromatography by using an Isolera Prime system (Biotage, Sweden) with prepacked silica columns (Biotage SNAP Ultra, Sweden). As mobile phase a mixture of ethyl acetate (10–55% v/v) and cyclohexane (0–2 column volumes (CV): 10% v/v ethyl acetate, 2–10.5 CV: gradient 10–55% v/v ethyl acetate, 10.5–13.5 CV: 55% v/v ethyl acetate) was used. Products were detected at 254 and 280 nm, respectively, after a typical retention time of 9 CVs. Isolated yields (Table 1) were calculated based on the respective benzaldehyde derivative.

Acknowledgements

Financial support from the German Research Foundation (DFG) in the frame of the research group FOR1296 "Diversity of asymmetric thiamine catalysis", and partial funding by the research training group "BioNoCo" (GK1166) is gratefully acknowledged. The authors thank Dr. Jan Marienhagen (IBG-1: Biotechnology) for helpful advice regarding the experimental design of site-saturation mutagenesis.

Keywords: asymmetric catalysis · biocatalysis · C–C coupling · enantioselectivity · protein engineering

[1] P. Hoyos, J. V. Sinisterra, F. Molinari, A. R. Alcantara, P. D. de Maria, *Acc. Chem. Res.* **2010**, *43*, 288–299.
[2] D. Enders, O. Niemeier, A. Henseler, *Chem. Rev.* **2007**, *107*, 5606–5655.
[3] S. O'Toole, C. A. Rose, S. Gundala, K. Zeitler, S. J. Connon, *J. Org. Chem.* **2011**, *76*, 347–357.
[4] M. Pohl, D. Gocke, M. Müller in *Handbook of Green Chemistry—Green Catalysis, Vol. 3* (Ed.: P. T. Anastas) Wiley-VCH, Weinheim, **2008**, pp. 75–114.
[5] M. Müller, D. Gocke, M. Pohl, *FEBS J.* **2009**, *276*, 2894–2904.
[6] M. Müller, G. A. Sprenger, M. Pohl, *Curr. Opin. Chem. Biol.* **2013**, *17*, 261–270.
[7] P. Dünkelmann, D. Kolter-Jung, A. Nitsche, A. S. Demir, P. Siegert, B. Lingen, M. Baumann, M. Pohl, M. Müller, *J. Am. Chem. Soc.* **2002**, *124*, 12084–12085.
[8] M. Knoll, M. Müller, J. Pleiss, M. Pohl, *ChemBioChem* **2006**, *7*, 1928–1934.
[9] R. Wilcocks, O. P. Ward, *Biotechnol. Bioeng.* **1992**, *39*, 1058–1063.
[10] A. S. Demir, M. Pohl, E. Janzen, M. Müller, *J. Chem. Soc. Perkin Trans. 1* **2001**, 633–635.
[11] D. Rother, G. Kolter, T. Gerhards, C. L. Berthold, E. Gauchenova, M. Knoll, J. Pleiss, M. Müller, M. Pohl, *ChemCatChem* **2011**, *3*, 1587–1596.
[12] M. Beigi, S. Waltzer, A. Fries, L. Eggeling, G. A. Sprenger, M. Müller, *Org. Lett.* **2013**, *15*, 452–455.
[13] G. Cavinato, L. Toniolo, A. Vavasori, *J. Mol. Catal.* **1994**, *89*, 93–100.
[14] J. S. Yadav, S. S. Mandal, *Tetrahedron Lett.* **2011**, *52*, 5747–5749.
[15] M. Jiang, Y. Cao, Z.-F. Guo, M. Chen, X. Chen, Z. Guo, *Biochemistry* **2007**, *46*, 10979–10989.
[16] A. Kurutsch, M. Richter, V. Brecht, G. A. Sprenger, M. Müller, *J. Mol. Catal. B* **2009**, *61*, 56–66.
[17] R. Westphal, S. Waltzer, U. Mackfeld, M. Widmann, J. Pleiss, M. Beigi, M. Müller, D. Rother, M. Pohl, *Chem. Commun.* **2013**, *49*, 2061–2063.
[18] D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C. L. Berthold, G. Schneider, J. Pleiss, M. Müller, M. Pohl, *ChemBioChem* **2008**, *9*, 406–412.
[19] A. J. J. Straathof, J. A. Jongejan, *Enzyme Microb. Technol.* **1997**, *21*, 559–571.
[20] C. Vogel, M. Widmann, M. Pohl, J. Pleiss, *BMC Biochem.* **2012**, *13*, 24.
[21] A. Dawson, P. K. Fyfe, W. N. Hunter, *J. Mol. Biol.* **2008**, *384*, 1353–1368.
[22] A. Dawson, M. Chen, P. K. Fyfe, Z. Guo, W. N. Hunter, *J. Mol. Biol.* **2010**, *401*, 253–264.
[23] M. Fang, A. Macova, K. L. Hanson, J. Kos, D. R. J. Palmer, *Biochemistry* **2011**, *50*, 8712–8721.
[24] F. Blanco, B. Kelly, I. Alkorta, I. Rozas, J. Elguero, *Chem. Phys. Lett.* **2011**, *511*, 129–134.
[25] M. Fang, R. D. Toogood, A. Macova, K. Ho, S. G. Franzblau, M. R. McNeil, D. A. R. Sanders, D. R. J. Palmer, *Biochemistry* **2010**, *49*, 2672–2679.
[26] R. S. Phillips, *Enzyme Microb. Technol.* **1992**, *14*, 417–419.
[27] M. Widmann, R. Radloff, J. Pleiss, *BMC Biochem.* **2010**, *11*, 9.
[28] W. L. DeLano, *The PyMOL Molecular Graphics System*, http://www.pymol.org, San Carlo, CA, USA, **2002**.
[29] H. B. Bürgi, J. D. Dunitz, E. Shefter, *J. Am. Chem. Soc.* **1973**, *95*, 5065–5067.
[30] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminf.* **2012**, *4*, 17.
[31] J. Sambrook, D. W. Russell, *Molecular Cloning: A Laboratory Manual, 1st ed.*, **2001**, pp. 1.116–1.118.
[32] W. M. Patrick, A. E. Firth, J. M. Blackburn, *Protein Eng.* **2003**, *16*, 451–457.
[33] M. Breuer, M. Pohl, B. Hauer, B. Lingen, *Anal. Bioanal. Chem.* **2002**, *374*, 1069–1073.

Received: April 26, 2013
Revised: July 8, 2013
Published online on September 6, 2013

Heterogeneous & Homogeneous & Bio-
CHEMCATCHEM
CATALYSIS

Supporting Information

Tailoring the *S*-Selectivity of 2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate Synthase (MenD) from *Escherichia coli*

Robert Westphal,[a] Doris Hahn,[a] Ursula Mackfeld,[a] Simon Waltzer,[b] Maryam Beigi,[b] Michael Widmann,[c] Constantin Vogel,[c] Jürgen Pleiss,[c] Michael Müller,[b] Dörte Rother,[a] and Martina Pohl*[a]

cctc_201300318_sm_miscellaneous_information.pdf

Table of contents

1. Amino acid sequence of *Ec*MenD and standard numbering

The amino acid sequence used in this work is derived from the original pET19b_*Ec*MenD construct,[1] which provides a 10×histidine tag. Absolute amino acid positions were counted based on this sequence starting with the first methionine of the original wildtype sequence (bold red).

1
MGHHHHHHHHHHSSGHIDDDDKH**M**SVSAFNRRWAAVILEALTRHGVRHICIAPGSRSTPLT
LAAAENSAFIHHTHFDERGLGHLALGLAKVSKQPVAVIVTSGTAVANLYPALIEAGLTGEKLIL
LTADRPPELIDCGANQAIRQPGMFASHPTHSISLPRPTQDIPARWLVSTIDHALGTLHAGGVH
INCPFAEPLYGEMDDTGLSWQQRLGDWWQDDKPWLREAPRLESEKQRDWFFWRQKRGV
VVAGRMSAEEGKKVALWAQTLGWPLIGDVLSQTGQPLPCADLWLGNAKATSELQQAQIVV
QLGSSLTGKRLLQWQASCEPEEYWIVDDIEGRLDPAHHRGRRLIANIADWLELHPAEKRQP
WCVEIPRLAEQAMQAVIARRDAFGEAQLAHRICDYLPEQGQLFVGNSLVVRLIDALSQLPAG
YPVYSNRGASGIDGLLSTAAGVQRASGKPTLAIVGDLSALYDLNALALLRQVSAPLVLIVVNN
NGGQIFSLLPTPQSERERFYLMPQNVHFEHAAAMFELKYHRPQNWQELETAFADAWRTPT
TTVIEMVVNDTDGAQTLQQLLAQVSHL
556

The standard numbers for ThDP-dependent decarboxylases are given according to Vogel *et al*. (Table S1).[2]

Table S1. Standard numbers of stereoselectivity affecting positions in *Ec*MenD according to Vogel *et al*..[2]

	Accession Number	Absolute position	Standard position [2]
*Ec*MenD	splP17109	32	28
		395	393
		474	476
		475	477
		478	480

2. Construction of *Ec*MenD variants

The *Ec*MenD variants F475G, I474G/F475G, I474G/F475A, I474A/F475G, and I474A/F475A were prepared as described elsewhere.[3] The original pET-19b_*Ec*MenD construct[1] was used for the preparation of *Ec*MenD_I474G applying the forward (fw) primer 5'-CAACAACGGCGGGCAAggcTTCTCGCTGTTGC-3' and the reverse (rv) primer 5'-GCAACAGCGAGAAgccTTGCCCGCCGTTGTTG-3' (mutated codons are underlined, with lower-case letters indicating base changes). Based on *Ec*MenD_I474G/F475G the mutants *Ec*MenD_I474G/F475V and *Ec*MenD_I474V/F475G were prepared. The following fw and rv primers were used for mutagenesis:

*Ec*MenD_I474G/F475V: 5'-CGGCGGGCAAGGCGtgTCGCTGTTGCCAAC-3' (fw)
5'-GTTGGCAACAGCGAcaCGCCTTGCCCGCCG-3' (rv)

*Ec*MenD_I474V/F475G: 5'-CAACGGCGGGCAAGtgGGCTCGCTGTTGC-3' (fw)
5'-GCAACAGCGAGCCcaCTTGCCCGCCGTTG-3' (rv)

For the generation of the mutant libraries for amino acid position 32, 395, and 478, respectively, *Ec*MenD_I474A/F475G was used as template. The following mutagenic primers containing NDT degeneracy (where N = any nucleotide, D = A, G, or T) were applied:

position 32: 5'-CTGTATCGCCCCAGGCndtCGTTCTACACCGTTAAC-3' (fw)
5'-GTTAACGGTGTAGAACGahnGCCTGGGGCGATACAG-3' (rv)

position 395: 5'-GGTAACAGCCTGGTGGTAndTCTGATTGATGCGCTTTC-3' (fw)
5'-GAAAGCGCATCAATCAGAhnTACCACCAGGCTGTTACC-3' (rv)

position 478: 5'-CAAGCGGGCTCGCTGndtCCAACGCCGCAAAG-3' (fw)
5'-CTTTGCGGCGTTGGahnCAGCGAGCCCGCTTG-3' (rv)

Additionally, variant I474A/F475G/S32A, which was not included in the NDT codon library, was prepared using *Ec*MenD_I474A/F475G as template, the fw primer 5'-GTATCGCCCCAGGCgCGCGTTCTACACC-3' and the rv primer 5'-GGTGTAGAACGCGcGCCTGGGGCGATAC-3'.

3. References

[1] A. Kurutsch, M. Richter, V. Brecht, G. A. Sprenger, M. Müller, *J. Mol. Catal. B: Enzym.* **2009**, *61*, 56–66.

[2] C. Vogel, M. Widmann, M. Pohl, J. Pleiss, *BMC Biochem.* **2012**, *13*:24.

[3] R. Westphal, S. Waltzer, U. Mackfeld, M. Widmann, J. Pleiss, M. Beigi, M. Müller, D. Rother, M. Pohl, *Chem. Comm.* **2013**, *49*, 2061–2063.

PUBLICATION IV

MenD from *Bacillus subtilis*: A Potent Catalyst for the Enantiocomplementary Asymmetric Synthesis of Functionalized α-Hydroxy Ketones

Westphal, R.; Jansen, S.; Vogel, C.; Pleiss, J.; Müller, M.; Rother, D. & Pohl, M.

ChemCatChem 2013,
DOI: 10.1002/cctc.201300690
(http://onlinelibrary.wiley.com/doi/10.1002/cctc.201300690/full)

Note: This manuscript represents the early view online version published before inclusion in an issue.

CHEMCATCHEM
FULL PAPERS

ChemPubSoc Europe

DOI: 10.1002/cctc.201300690

MenD from *Bacillus subtilis*: A Potent Catalyst for the Enantiocomplementary Asymmetric Synthesis of Functionalized α-Hydroxy Ketones

Robert Westphal,[a] Sascha Jansen,[a] Constantin Vogel,[b] Jürgen Pleiss,[b] Michael Müller,[c] Dörte Rother,[a] and Martina Pohl*[a]

The thiamine diphosphate-dependent enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) catalyzes a Stetter-like 1,4-addition of α-ketoglutarate to isochorismate in the biosynthesis of menaquinone (vitamin K). Here, we describe the carboligation potential of MenD from *Bacillus subtilis* (*Bs*MenD) for the nonphysiological 1,2-addition of decarboxylated α-ketoglutarate (succinylsemialdehyde) and various benzaldehyde derivatives. Furthermore, we engineer *Bs*MenD variants for the enantiocomplementary asymmetric synthesis of functionalized α-hydroxy ketones. Wild type *Bs*MenD shows an excellent chemo- as well as high (*R*)-selectivity for the carboligation of α-ketoglutarate as the donor, and different benzaldehyde derivatives as acceptor yielding (*R*)-α-hydroxy ketones with up to >99% *ee*. By engineering (*S*)-selective *Bs*MenD variants, based on the recently developed *S*-pocket concept, we provide access to most of the corresponding (*S*)-α-hydroxy ketones with up to 98% *ee*. In particular, benzaldehyde and *meta*-substituted derivatives were converted with high enantioselectivities (*ee* of 91–98% (*S*)). The significantly higher (*S*)-selectivity of *Bs*MenD variants than recently published MenD variants from *Escherichia coli*, could be attributed to a second-shell residue next to the *S*-pocket. A glycine residue, adjacent to the major *S*-pocket residues I476 and F477 (standard numbering), is assumed to result in higher structural flexibility in the *S*-pocket region of *Bs*MenD, which in turn could result in improved stabilization of the antiparallel orientation of the acceptor.

Introduction

Thiamine diphosphate (ThDP)-dependent enzymes are well-known for their catalytic potential and broad applicability in asymmetric synthesis. During the last few decades, the discovery and characterization of several ThDP-dependent enzymes that perform stereoselective C–C bond formation has paved the way for the enzymatic production of a broad range of α-hydroxy ketones,[1] which are of particular interest as chiral building blocks for the pharmaceutical as well as the fine chemical industries.[2]

ThDP-dependent enzymes are well-known for their high chemo- and stereoselectivity in mixed benzoin condensation-like carboligations. Moreover, their different substrate preferences open up access to a broad range of various chiral α-hydroxy ketones with high enantiomeric excesses (*ee*).[3] However, owing to the prevailing (*R*)-selective nature of members of the ThDP-dependent decarboxylase superfamily (family classification according to Duggleby 2006 and Widmann 2010),[4] the access to (*S*)-α-hydroxy ketones is limited. Therefore, we have developed an enzyme toolbox for the direct catalytic asymmetric synthesis of enantiocomplementary α-hydroxy ketones.

The principles of stereoselectivity can be elucidated based on the mechanistic understanding of stereoselective carboligation by comparison of benzaldehyde lyase from *Pseudomonas fluorescens* (*Pf*BAL) and benzoylformate decarboxylase from *Pseudomonas putida* (*Pp*BFD).[5] Whereas *Pf*BAL is strictly (*R*)-selective (*ee*>99%) for the formation of 2-hydroxypropiophenone (2-HPP) starting from benzaldehyde and acetaldehyde as substrates,[6] *Pp*BFD shows (*S*)-selectivity (*ee*=92%) for the same reaction.[7] The structural reason is a so-called "*S*-pocket" in *Pp*BFD, which allows for the antiparallel arrangement of the acetaldehyde acceptor relative to the ThDP-bound donor benzaldehyde prior to carboligation; the prerequisite for (*S*)-selectivity. In *Pf*BAL an *S*-pocket is missing, which explains the high (*R*)-selectivity; this in turn results from a parallel orientation of both substrates.[5]

S-pockets are present in many ThDP-dependent enzymes of the decarboxylase family. However, they are regularly blocked by bulky amino acid side chains and thus, inaccessible for acceptor substrates.[8] Based on the developed *S*-pocket concept,

[a] R. Westphal, S. Jansen, Dr. D. Rother, Prof. M. Pohl
Forschungszentrum Jülich GmbH
IBG-1: Biotechnology
Leo-Brandt-Str. 1, 52425 Jülich (Germany)
Fax: (+49) 2461-61-3870
E-mail: ma.pohl@fz-juelich.de

[b] C. Vogel, Prof. J. Pleiss
Institute of Technical Biochemistry
University of Stuttgart
Allmandring 31, 70569 Stuttgart (Germany)

[c] Prof. M. Müller
Institute of Pharmaceutical Sciences
Albert-Ludwigs-University Freiburg
Albertstr. 25, 79104 Freiburg (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cctc.201300690.

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

(*S*)-selectivity can be tailored for *Pp*BFD and pyruvate decarboxylase from *Acetobacter pasteurianus* (*Ap*PDC), which thereby expands our enzyme toolbox and broadens the enzymatically accessible α-hydroxy ketone platform by new (*S*)-2-HPP and (*S*)-phenylacetylcarbinol derivatives, respectively.[9]

Recently, the general validity of the *S*-pocket concept for ThDP-dependent enzymes has been supported by using the principles for the 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase enzyme from *Escherichia coli* (*Ec*MenD).[10] MenD is part of an extraordinary subgroup of ThDP-dependent enzymes that uses α-ketoglutarate (α-KG) as physiological donor.[11] In contrast to *Pp*BFD and *Ap*PDC, which belong to the group of lyases (EC 4), MenD is classified as a transferase (EC 2.2.1.9). The physiological activity of MenD is the Stetter-like 1,4-addition of α-KG to isochorismate, which represents the second step of the menaquinone biosynthesis.[12] The carboligation potential of *Ec*MenD has recently been explored and characterized with regards to the nonphysiological 1,2-addition of α-KG to a broad range of acceptor aldehydes.[11,13] As with most ThDP-dependent enzymes, *Ec*MenD is characterized by excellent chemoselectivity as well as high (*R*)-selectivity. The mixed carboligation of α-KG and benzaldehyde derivatives gives functionalized α-hydroxy ketones (5-hydroxy-4-oxo acids) with high *ee* values of >93% (*R*).[11] By tailoring the *S*-pocket of *Ec*MenD through mutagenesis of I474 and F475 to glycine and alanine, respectively, access to various (*S*)-5-hydroxy-4-oxo-5-arylpentanoates is possible for the first time. In particular, the use of *meta*-substituted benzaldehydes results in good to high (*S*)-selectivities (82–97% *ee*) accompanied by good to excellent conversions by using *Ec*MenD-I474A/F475G.[10a]

Recently, Dawson et al. published the structure of MenD from *Bacillus subtilis* (*Bs*MenD).[14] Although *Bs*MenD shares only 28% of the overall sequence identity with *Ec*MenD, the structures of both enzymes are closely related,[14] which is a common feature of ThDP-dependent decarboxylases.[15] Moreover, *Bs*MenD revealed the expected activity for the physiological 1,4-addition of α-KG to isochorismate.[14]

In the present study, we report on the investigation of the biocatalytic potential of *Bs*MenD with respect to the nonphysiological 1,2-addition of α-KG to various aromatic aldehydes. Furthermore, we applied the *S*-pocket concept to *Bs*MenD that was based on the mutagenesis strategy developed for *Ec*MenD.[10] Thereby, we could readily access various (*S*)-5-hydroxy-4-oxo-5-arylpentanoates with even higher *ee*-values than those that have been obtained with *Ec*MenD so far. In addition, this study revealed a second-shell amino acid close to the *S*-pocket region that significantly influences the enantioselectivity of *Bs*MenD.

Results and Discussion

To highlight the structural equivalence of *Bs*MenD and *Ec*MenD amino acid residues mentioned in this work and to enhance their comparative analysis, amino acid positions of both enzymes are given as standard numbers according to the standard numbering scheme for ThDP-dependent decarboxylases (Table 1).[15]

Table 1. Standard numbers for amino acid residues of *Bs*MenD and *Ec*MenD mentioned in this work.

Standard number[15]	*Bs*MenD[a]	*Ec*MenD[b]
26	P29	P30
27	G30	G31
28	S31	S32
388	N404	N390
393	R409	R395
410	R428	R413
413	N431	S416
475	G488	Q473
476	I489	I474
477	F490	F475
480	L493	L478

[a] Swiss-Prot Accession Number: sp|P23970. [b] Swiss-Prot Accession Number: sp|P17109.

Nonphysiological 1,2-addition

The catalytic potential of *Bs*MenD in the nonphysiological stereoselective benzoin condensation-like 1,2-addition is investigated by using α-KG and substituted benzaldehydes. As already observed for *Ec*MenD,[11,13] this kind of reaction is characterized by excellent chemo- and high enantioselectivity. The only detectable carboligation products are the desired 5-hydroxy-4-oxo-5-arylpentanoates with high *ee* values and excellent conversions under the tested conditions (Table 2). Neither the self-ligation products of α-KG and benzaldehyde, respectively, nor the 5-oxo-4-hydroxy-5-arylpentanoates derived from benzaldehyde as the donor and α-KG as the acceptor, are detectable. The preference for α-KG as the donor owes to the polar character of the donor-binding site, which is mainly defined by two arginine (R393 and R410) and asparagine residues (N388 and N413), respectively. In contrast, the acceptor-binding site is determined by a hydrophobic patch; this is formed by I476, F477, and L480 residues,[14] which is ideal for stabilizing the aromatic acceptor by nonpolar interactions.

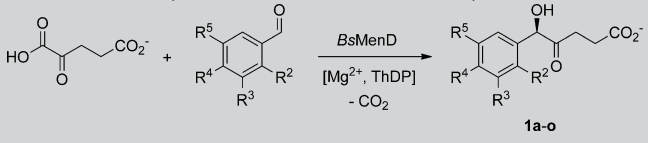
As expected, *Bs*MenD reveals high (*R*)-selectivity for the formation of 5-hydroxy-4-oxo-5-arylpentanoates (**1a**–**o**) (Table 2). The carboligation of α-KG and benzaldehyde results in the synthesis of **1a** with excellent enantioselectivity of >99% (*R*). With *meta*- and *para*-substituted benzaldehydes, high *ee*-values of >97% (*R*) for **1g**–**o** are obtained, whereas *ortho*-substituted benzaldehydes are converted with lower (*R*)-selectivity to yield **1b**–**f** with 81–94% *ee*. The same trends have recently been observed with *Ec*MenD.[11] However, the reasons for the reduced (*R*)-selectivity with *ortho*-substituted benzaldehydes are not yet known.

The *S*-pocket region of *Bs*MenD

To explain the high (*R*)-selectivity of *Bs*MenD in mixed carboligations, the *S*-pocket region is compared to *Ec*MenD, which demonstrates that the corresponding positions in both en-

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

Table 2. *Bs*MenD-catalyzed 1,2-addition of α-KG and benzaldehyde derivatives.[a]



Compound	R^2	R^3	R^4	R^5	*ee* [%][b]	Conversion [%][c]
1a	H	H	H	H	>99 (*R*)	>99
1b	F	H	H	H	94 (*R*)	>99
1c	Cl	H	H	H	81 (*R*)	>99
1d	Br	H	H	H	88 (*R*)	99
1e	I	H	H	H	81 (*R*)	99
1f	OCH_3	H	H	H	94 (*R*)	>99
1g	H	F	H	H	>99 (*R*)	>99
1h	H	Cl	H	H	97 (*R*)	>99
1i	H	Br	H	H	>99 (*R*)	>99
1j	H	I	H	H	>99 (*R*)	95
1k	H	OCH_3	H	H	>99 (*R*)	>99
1l	H	H	F	H	>99 (*R*)	99
1m	H	H	Cl	H	97 (*R*)	>99
1n	H	H	Br	H	97 (*R*)	>99
1o	H	OCH_3	H	OCH_3	99 (*R*)	>99

[a] Reaction conditions: TEA buffer (50 mM), pH 8.0, which contains $MgSO_4$ (2 mM), ThDP (0.1 mM); enzyme (0.7 mg mL^{-1}); α-KG (50 mM), and benzaldehyde derivative (20 mM); T = 30 °C, t = 22 h. [b] Determined by chiral-phase HPLC. [c] Determined by GC based on consumption of benzaldehyde derivatives.

Tailoring the *S*-pocket of *Bs*MenD

To engineer stereoselectivity of *Bs*MenD, the same strategy was applied as used before for *Ec*MenD.[10] The *S*-pocket of *Bs*MenD was opened by mutation of the I476 and F477 residues to glycine and alanine, respectively. Modeling studies reveal a large *S*-pocket with no steric restrictions for the antiparallel benzaldehyde orientation (Figure 2B).

Although the resulting *Bs*MenD variants I476G/F477G, I476A/F477G, and I476A/F477A are obtained as soluble proteins upon expression in *E. coli* BL21, variant I476G/F477A is insoluble and therefore not further characterized. In the next step, the carboligation of α-KG and benzaldehyde to **1a** was used as a model reaction to investigate the enantioselectivity of the variants.

As expected, the enantioselectivity can be inverted for all variants, which results in remarkably good (*S*)-selectivities for the formation of **1a** with 67–92% *ee* (Table 3). Thus, compared with respective (*S*)-selective *Ec*MenD variants (*ee* = 34–75%),[10a] these values are significantly higher. The lower catalytic activity of *Bs*MenD variants compared with wild type (wt) *Bs*MenD might owe to the influence of positions 476 and 477 on the binding affinity of cofactor ThDP, as well as the substrates. At this point, a reliable determination of kinetic parameters for the 1,2-addition of α-KG to benzaldehyde (and the derivatives) is hampered by the solubility of the aromatic substrate, which prevents single-substrate kinetics under saturated conditions. *Ec*MenD studies reveal the same trends with respect to decreased conversions.[10] Moreover, Dawson et al. reported on increased K_m values for the donor α-KG as well as for the natural acceptor isochorismate by a factor of 12–25, which was accompanied by a reduction in catalytic efficiency after mutagenesis of both residues to alanine.[14]

In addition, (*S*)-selectivity and conversion were influenced by the addition of the cosolvent methyl *tert*-butyl ether (MTBE). In

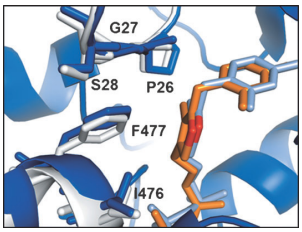


Figure 1. Structural alignment of the *S*-pocket regions of *Bs*MenD (white) and *Ec*MenD (blue). Amino acids that determine the *S*-pocket are shown as sticks. *Bs*MenD ThDP is colored in orange, *Ec*MenD ThDP in light blue. C2 as the catalytically active carbanion in ThDP is colored in red. Amino acid positions are given as standard numbers (see Table 1).

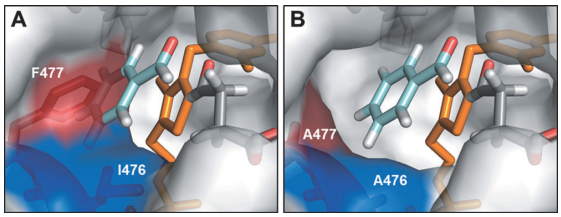
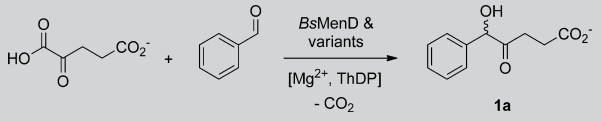


Figure 2. A) The active site of *Bs*MenD-wt and *Bs*MenD-I476A/F477A with modeled substrates: antiparallel oriented benzaldehyde (cyan) and succinylsemialdehyde (gray) bound to the C2-ThDP (orange). An antiparallel arrangement of benzaldehyde relative to succinylsemialdehyde in the active site of *Bs*MenD-wt is not possible. Therefore, the wild type enzyme is (*R*)-selective. B) Exchange of the I476 (blue) and F477 (red) residues to alanine open up the *S*-pocket, which allows for the antiparallel orientation of benzaldehyde and therewith (*S*)-selective carboligation. Amino acid positions are given as standard numbers (see Table 1).

zymes are occupied by the same amino acids (Figure 1). In addition, the *S*-pocket regions of *Bs*MenD and *Ec*MenD exhibit high structural similarity and the orientation of the cofactor ThDP is nearly the same.

In *Ec*MenD residues I476 and F477 are found to be crucial for stereoselectivity. Both amino acids prevent the antiparallel arrangement of benzaldehyde relative to the donor α-KG prior to carboligation, which explains the enzyme's high (*R*)-selectivity.[10] The same role might be realized by the I476 and F477 residues in *Bs*MenD (Figure 2A). Consequently, the principles of stereoselectivity that were elucidated by the *Ec*MenD studies[10] should be transferable to *Bs*MenD.

Table 3. Conversions and *ee*-values for the carboligation of α-KG and benzaldehyde to 5-hydroxy-4-oxo-5-phenylpentanoate (**1a**) by *Bs*MenD-wt and variants in the presence and absence of MTBE as a cosolvent.[a]



*Bs*MenD or variant	*ee* [%][b]	Conversion [%][c]
wt		
no cosolvent	>99 (*R*)	>99
5 vol% MTBE	>99 (*R*)	>99
I476G/F477G		
no cosolvent	81 (*S*)	28
5 vol% MTBE	n.d.[d]	<5
I476A/F477G		
no cosolvent	75 (*S*)	40
5 vol% MTBE	67 (*S*)	59
I476A/F477A		
no cosolvent	92 (*S*)	36
5 vol% MTBE	85 (*S*)	10

[a–b] See Table 2. [c] Determined by GC, based on consumption of benzaldehyde. [d] Not determined.

the presence of 5 vol% MTBE, *Bs*MenD-variants I476G/F477G and I476A/F477A revealed reduced catalytic activity, which may be estimated in a first approach from the conversions after 22 h, whereas *Bs*MenD-I476A/F477G showed slightly improved activity (Table 3). No influence could be detected in the case of *Bs*MenD-wt under the tested conditions.

As demonstrated in Table 3, (*S*)-selectivity for the formation of **1a** is reduced significantly in the presence of MTBE for *Bs*MenD-I476A/F477G, and I476A/F477A (from 75% to 67% *ee* and from 92% to 85% *ee*, respectively), whereas the enantioselectivity of *Bs*MenD-wt is not affected. These results are consistent with the recently published studies of Gerhards et al.,[16] who found that the stereoselectivity of ThDP-dependent enzymes is influenced by a direct interaction of the solvent molecules [which depends on the size and partition coefficient (log P)] with the active site, especially the *S*-pocket. However, (*S*)-selective *Ec*MenD variants do show contrary effects. The best results with respect to the enantioselectivity are obtained in the presence of 5 vol% MTBE. For example, *Ec*MenD-I476A/F477A catalyzes the formation of **1a** with an *ee* of 66% (*S*) in the presence of MTBE,[10a] whereas (*S*)-selectivity is slightly reduced (*ee* = 59%) in the absence of the cosolvent.

Consequently, although the *S*-pockets of *Bs*MenD and *Ec*MenD are very similar with respect to amino acid residues, there are clearly crucial differences that influence enantioselectivity. The best stabilization of the antiparallel benzaldehyde orientation to date is achieved in *Bs*MenD-I476A/F477A (*ee* = 92%); this owes to good stabilization of the acceptor in the *S*-pocket, which is a requirement for high (*S*)-selectivity.[9,10]

Substrate screening

As we have recently reported, *Ec*MenD variants I476G/F477G, I476A/F477G, and I476A/F477A convert *meta*-substituted benzaldehydes, in particular, with high (*S*)-selectivities as well as with high conversions, which indicates good stabilization of these acceptors in the *S*-pocket. In contrast, *ortho*- and *para*-substituted benzaldehydes give (*S*)-5-hydroxy-4-oxo-5-arylpentanoates with low *ee* values or with the (*R*)-enantiomer in excess.[10a]

As demonstrated in Table 4, *Bs*MenD variants I476G/F477G, I476A/F477G, and I476A/F477A show similar trends. Again, highest (*S*)-selectivities as well as conversions are obtained by using *meta*-substituted benzaldehydes, whereas *ortho*- and *para*-substituted benzaldehydes, except fluorobenzaldehydes, are converted with low enantioselectivity, which is accompanied by low conversions. As is already observed for the synthesis of **1a**, we obtain some remarkable differences between the enantioselectivities of *Bs*MenD and *Ec*MenD for different benzaldehyde derivatives. Compared with the best (*S*)-selective *Ec*MenD variant I476A/F477G, all *Bs*MenD variants show higher (*S*)-selectivity for the formation of *meta*-substituted 5-hydroxy-4-oxo-5-arylpentanoates (**1g**–**k**, **o**). For example, *Bs*MenD-I476A/F477A catalyzes the conversion of *meta*-substituted benzaldehydes with excellent enantioselectivities (*ee* > 97% (*S*)), except for **1j** (*ee* = 88% (*S*)). In contrast, *Ec*MenD-I476A/F477G catalysis only yields **1i** and **1o** with high *ee* values of 97% (*S*) and 96% (*S*), respectively.[10] In addition, the products (*S*)-**1b** and (*S*)-**1l** are now accessible with good to high *ee* values. Whereas *Bs*MenD-I476A/F477A catalyzed the transformation of 2-fluoro-benzaldehyde to (*S*)-**1b** with an *ee* of 83%, *Bs*MenD-I476A/F477G-catalyzed conversion of 4-fluorobenzaldehyde gave (*S*)-**1l** with high *ee* of 96% but low conversion under the tested conditions (Table 4). Moreover, (*S*)-selective syntheses of *para*-substituted **1m**–**n** (*ee* = 44–52%) are observed for the first time by using *Bs*MenD-I476G/F477G. However, it should be emphasized that reaction conditions were not optimized, which is expected to further increase both the enantioselectivity and conversion rates.

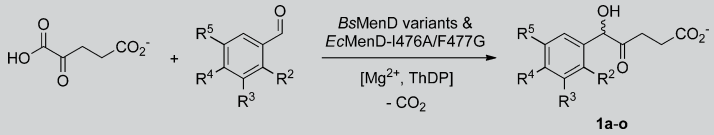
Influence of second-shell residue G475 on the enantioselectivity of *Bs*MenD

Despite an analogous *S*-pocket architecture in *Bs*MenD and *Ec*MenD (Figure 1), (*S*)-selectivity of engineered variants differs significantly (Table 4). To find an explanation for these differences, we further compare the *S*-pocket regions of both enzymes, in which we also consider second-shell residues that do not directly interact with the acceptor substrate.

The enantioselectivity-determining residues in both enzymes, I476 and F477, are located on a short α-helix, which consists of five amino acids. This region is conserved in most positions (Figure 3A, for further details see the Supporting Information, Figure S3).[14] However, the residue with standard

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

Table 4. Enantioselectivity and conversion of *Bs*MenD variant-catalyzed 1,2-addition reactions of α-KG to benzaldehyde derivatives in comparison to *Ec*MenD-I474A/F475G.[a] For results with *Bs*MenD-wt see Table 2.



Compound	R^2	R^3	R^4	R^5	I476G/F477G ee [%][b]	I476G/F477G Conv. [%][c]	I476A/F477G ee [%]	I476A/F477G Conv. [%]	I476A/F477A ee [%]	I476A/F477A Conv. [%]	G475Q/I476G/F477G ee [%]	G475Q/I476G/F477G Conv. [%]	*Ec*MenD I476A/F477G[d] ee [%][e]	*Ec*MenD I476A/F477G[d] Conv. [%][e]
1a	H	H	H	H	81 (*S*)	28	75 (*S*)	40	92 (*S*)	36	77 (*R*)	5	75 (*S*)	15
1b	F	H	H	H	72 (*S*)	36	75 (*S*)	15	83 (*S*)	33	77 (*R*)	16	48 (*S*)	55
1c	Cl	H	H	H	<5	<5	47 (*S*)	<5	45 (*S*)	<5	74 (*R*)	<5	<5	57
1d	Br	H	H	H	38 (*R*)	<5	25 (*S*)	<5	21 (*S*)	<5	90 (*R*)	<5	<5	52
1e	I	H	H	H	67 (*R*)	5	21 (*R*)	7	9 (*R*)	<5	95 (*R*)	6	<5	69
1f	OCH_3	H	H	H	69 (*S*)	8	71 (*S*)	5	<5	11	57 (*R*)	<5	51 (*S*)	13
1g	H	F	H	H	74 (*S*)	42	83 (*S*)	67	97 (*S*)	66	n.d.[f]	<5	82 (*S*)	43
1h	H	Cl	H	H	97 (*S*)	41	98 (*S*)	31	97 (*S*)	17	33 (*S*)	15	89 (*S*)	94
1i	H	Br	H	H	98 (*S*)	69	97 (*S*)	13	98 (*S*)	31	54 (*S*)	11	97 (*S*)	87
1j	H	I	H	H	91 (*S*)	22	90 (*S*)	34	88 (*S*)	13	73 (*S*)	13	93 (*S*)	86
1k	H	OCH_3	H	H	96 (*S*)	69	95 (*S*)	66	98 (*S*)	80	41 (*R*)	17	93 (*S*)	64
1l	H	H	F	H	16 (*S*)	6	96 (*S*)	<5	76 (*S*)	<5	74 (*R*)	<5	46 (*S*)	14
1m	H	H	Cl	H	44 (*S*)	34	19 (*S*)	33	<5	29	80 (*R*)	<5	48 (*R*)	15
1n	H	H	Br	H	52 (*S*)	5	50 (*R*)	<5	37 (*R*)	<5	82 (*R*)	<5	69 (*R*)	17
1o	H	OCH_3	H	OCH_3	97 (*S*)	14	98 (*S*)	6	98 (*S*)	47	<5	26	96 (*S*)	83

[a–c] See Table 2. [d] Recently reported.[10a] [e] The same reaction conditions as described under [a], but with the addition of 5 vol% MTBE and 24 h reaction time. [f] Not determined.

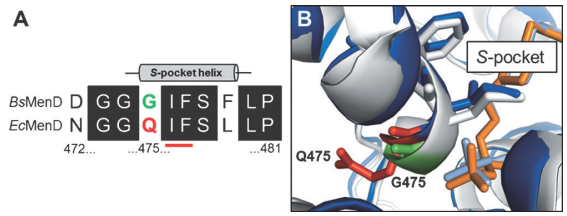


Figure 3. Sequential and structural alignment of the *S*-pocket regions of *Bs*MenD (white) and *Ec*MenD (blue). Most of the amino acid residues are strictly conserved (black background) according to Dawson et al.,[14], which includes the most prominent *S*-pocket residues I476 and F477 (underlined bold red; part A). One exception is the standard position 475. The amino acid side chain at position 475 is not directed towards the acceptor binding site in the *S*-pocket (part B), thus this does not directly influence the binding of the substrate. Amino acid positions are given as standard numbers (see Table 1).

number 475, which is a glycine in *Bs*MenD and a glutamine in *Ec*MenD, differs significantly both in size and chemical properties. Although the glutamine side chain is not directed towards the acceptor substrate (Figure 3B), it might influence the position of the *S*-pocket α-helix or its flexibility, which in turn could have an impact on enantioselectivity.

To investigate the influence of G475 on the enantioselectivity of *Bs*MenD, we exchanged the glycine to glutamine in the (*S*)-selective variant *Bs*MenD-I476G/F477G. Whereas *Bs*MenD-I476G/F477G showed high (*S*)-selectivity for the formation of **1a** (*ee* = 81%) compared to the respective *Ec*MenD variant I476G/F477G with a glutamine in position 475 (*ee* = 34% (*S*)),[10a] *Bs*MenD-I476G/F477G/G475Q reveals remarkable (*R*)-selectivity (*ee* = 77%) for the same reaction (Table 4). This result hints at significant changes in the structure of the *S*-pocket region of *Bs*MenD induced by the exchange of G475 by glutamine, such as a shift of the *S*-pocket α-helix, which might result in partial blockage of the *S*-pocket and less favorable stabilization of benzaldehyde. Alternatively, the flexibility of the *S*-pocket might be decreased by replacing one of the three consecutive glycine residues adjacent to the *S*-pocket residues 476 and 477. This observation suggests that high flexibility of the *S*-pocket region in (*S*)-selective *Bs*MenD variants, as a result of three glycine residues, leads to a better fit of benzaldehyde in the *S*-pocket and thus a higher (*S*)-selectivity than the respective (*S*)-selective *Ec*MenD variants.

The substrate screening with *Bs*MenD-I476A/F477A/G475Q also reveals inverted enantioselectivities for *ortho*- and *para*-substituted benzaldehydes (Table 4). (*R*)-Selectivity is even improved with this triple variant for the synthesis of **1d–e** compared to *Bs*MenD-wt. Remarkably, *meta*-substituted benzaldehydes are still transformed (*S*)-selectively (*ee* = 33–73%), with the exception of **1k** (*ee* = 41% (*R*)) and **1o** (*ee* < 5).

G475 is the second example of a second-shell residue in ThDP-dependent enzymes that influences the enantioselectivity without direct contact to the substrates. Previously, it was demonstrated that improved enantioselectivity of *Pp*BFD in the carboligation of benzaldehyde and acetaldehyde was obtained by mutation of the L476 residue (standard number 488).[17] The L476 residue is not directly part of the active site, however, it is in close contact to the *S*-pocket residue L461 (standard number 477), which allows for direct interactions. Thus, mutation of the L476 residue might result in

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

structural changes of the *S*-pocket region that in turn influences enantioselectivity.

Conclusions

We present a new member of our ThDP-dependent enzyme toolbox as a potent catalyst for the enantiocomplementary asymmetric synthesis of functionalized α-hydroxy ketones. Whereas *Bs*MenD-wt catalyzes the formation of diversely substituted (*R*)-5-hydroxy-4-oxo-5-arylpentanoates with high enantioselectivities and excellent conversions, engineered *Bs*MenD variants give access to most of the enantiocomplementary (*S*)-products with the highest enantioselectivities recorded so far, particularly for *meta*-substituted (*S*)-5-hydroxy-4-oxo-5-arylpentanoates. The results confirm the general validity of the existing *S*-pocket concept for ThDP-dependent enzymes of the decarboxylase superfamily,[5] which allows for a successful design of (*S*)-selective variants.

In addition, we reveal a further position, which is crucial for the enantioselectivity of *Bs*MenD (Figure 3B). The second-shell residue G475 might influence the flexibility of the *S*-pocket region as well as the position of the *S*-pocket α-helix. This glycine residue is conserved in 63% of ThDP-dependent enzymes of the decarboxylase superfamily, as was deduced from the ThDP-dependent Enzyme Engineering Database.[4b] However, for example, in pyruvate decarboxylases from *Acetobacter pasteurianus*, *Zymomonas mobilis* and *Saccharomyces cerevisiae* this position is occupied by valine (*Ap*PDC) and threonine residues (*Zm*PDC, *Sc*PDC), respectively. In the case of the (*S*)-selective *Ap*PDC variant E477G (standard number), which shows only moderate (*S*)-selectivity, substitution by glycine might be beneficial to increase (*S*)-selectivity. We therewith identify a potential further setscrew to adjust (*S*)-selectivity by influencing the *S*-pocket backbone. In addition, we have recently identified another important position (393) next to the donor binding site and the substrate channel of *Ec*MenD, in which the R393 residue is shown to stabilize the *R*-pathway and exchange to tyrosine increases the (*S*)-selectivity.[10b]

To gain a deeper insight into enantioselective catalysis with MenD, structural studies with (*S*)-selective MenD variants are underway. Based on the crystal structures, comprehensive molecular dynamic simulations of the binding states (parallel versus antiparallel orientation of the acceptor) will be performed. Furthermore, we want to explore the scope of MenD catalysis with respect to other acceptor substrates, for example, aliphatic aldehydes.

Experimental Section

Materials

All chemicals were obtained from Sigma–Aldrich. *Thermococcus kodakaraensis* (KOD) Hot Start DNA Polymerase was purchased from Merck Millipore. The restriction enzyme DpnI was obtained from Fermentas (Thermo Scientific).

Enzymes

Wild type *Bs*MenD was purchased as a synthetic gene (for the gene sequence, see the Supporting Information) from GeneArt Gene Synthesis (life technologies). By using standard molecular biology techniques,[18] the gene, which contains the restrictions sites NdeI and BamHI, was cloned into a plasmid expression T7 promotor (pET)-19b vector (Novagen); this provided an *N*-terminal 10×histidine (His)-tag. All variants were generated by the standard Quikchange site-directed mutagenesis protocol (Stratagene) as previously described.[10a] Primer sequences are given in the Supporting Information. Correct insertion of the synthetic gene as well as the gene sequences of the variants were confirmed by DNA sequencing (LGC Genomics, Berlin). Expression and purification by Ni-chelate affinity chromatography of *Bs*MenD and variants thereof was performed as described elsewhere.[13,10] Purified enzymes were used as lyophilisates.

Structural analysis

The PyMOL software (Schrödinger)[19] was used for all structural investigations of *Bs*MenD (pdb code: 2×7j) and *Ec*MenD (pdb code: 2jlc), generation of respective variants, and manual placement of the substrates. The benzaldehyde was aligned in the active site by considering the reactive distance of its carbonyl group to the ThDP-bound succinylsemialdehyde, as well as the Bürgi–Dunitz angle that defines the geometry of the nucleophilic addition to a carbonyl group.[20] The 3D-structures of the ThDP-bound succinylsemialdehyde and benzaldehyde were generated by using the Avogadro software.[21] Amino acid positions were counted based on the sequence of *Bs*MenD and *Ec*MenD, respectively, without an *N*-terminal His-tag (see Supporting Information).

Reactions

All carboligation reactions were performed at analytical scale (0.5 mL) in glass vials by using the lyophilized enzyme (0.7 mg mL^{-1}). Protein concentrations were determined according to Bradford.[22] α-KG (50 mM) and the benzaldehyde derivatives (20 mM, final concentrations) were dissolved in triethanolamine buffer (50 mM, pH 8.0) that contained ThDP (0.1 mM), $MgSO_4$ (2 mM), and MTBE (5 vol%) if indicated. The reactions were performed in 2–3 replicates (1.5 mL) in a thermomixer (Eppendorf, Germany) at $T=30\,^{\circ}C$ and at a rate of 800 rpm for 22 h. Afterwards, the conversions and *ee* values were determined by using GC and HPLC (see below), respectively. Conversions were determined with an error of $<10\%$. Errors for *ee*-data were $<5\%$. Negative controls were performed without enzyme. Besides the expected products (**1a–o**) no further byproducts were detected.

Sample preparation and analysis

Complete characterization of the products **1a–o** is described elsewhere.[10,11,13] Determination of *ee* and conversions were performed with minor alterations.

Enantiomeric excesses were determined by chiral phase HPLC. The reaction mixture (300 μL) was mixed with perchloric acid [10% (v/v)] to stop the reaction. Afterwards, the product was extracted with MTBE (200 μL). Finally, the organic phase was used for chiral analysis on a 1260 Infinity chromatography system (Agilent Technologies, USA) by using Diacel Chiralcel OD-H columns (5 μm, 250 mm×4.6 mm), and Chiralpak IC (5 μm, 250 mm×4.6 mm) (Daicel Chemical IND., France). All products were detected at $\lambda=$ 210 nm by using an *n*-hexane:2-propanol mixture. **1a**: Chiralcel

CHEMCATCHEM
FULL PAPERS
www.chemcatchem.org

OD-H, 20 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 92:8, $t_R(R)$ = 16.9 min, $t_R(S)$ = 15.4 min; **1b–f**: Chiralcel OD-H, 20 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 90:10, **1b**: $t_R(R)$ = 10.2 min, $t_R(S)$ = 9.3 min, **1c**: $t_R(R)$ = 12.3 min, $t_R(S)$ = 10.3 min, **1d**: $t_R(R)$ = 14.8 min, $t_R(S)$ = 11.6 min, **1e**: $t_R(R)$ = 17.3 min, $t_R(S)$ = 13.2 min, **1f**: $t_R(R)$ = 18.3 min, $t_R(S)$ = 15.8 min; **1g**: Chiralpak IC, 40 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 95.5:4.5, $t_R(R)$ = 38.0 min, $t_R(S)$ = 36.4 min; **1h**: Chiralpak IC, 20 °C, 1.2 mL min^{-1}, *n*-hexane:methanol (0.25% TFA) = 95.5:4.5, $t_R(R)$ = 23.9 min, $t_R(S)$ = 27.5 min; **1i**: Chiralpak IC, 40 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 95.5:4.5, $t_R(R)$ = 33.0 min, $t_R(S)$ = 35.0 min; **1j**: Chiralpak IC, 40 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 95.5:4.5, $t_R(R)$ = 34.5 min, $t_R(S)$ = 38.43 min; **1k**: Chiralpak IC, 20 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 85:15, $t_R(R)$ = 16.3 min, $t_R(S)$ = 20.3 min; **1l–n**: Chiralcel OD-H, 20 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 95:5, **1l**: $t_R(R)$ = 26.4 min, $t_R(S)$ = 28.3 min, **1m**: $t_R(R)$ = 27.2 min, $t_R(S)$ = 30.6 min, **1n**: $t_R(R)$ = 29.9 min, $t_R(S)$ = 33.5 min; **1o**: Chiralpak IC, 20 °C, 1.2 mL min^{-1}, *n*-hexane:2-propanol (0.25% TFA) = 85:15, $t_R(R)$ = 19.4 min, $t_R(S)$ = 26.7 min.

Conversions were determined by GC, based on the consumption of benzaldehyde and derivatives thereof. The reaction solution (300 μL) was mixed with ethyl acetate (200 μL), which contained decane as the internal standard. After centrifugation, the organic phase (5 μL) was used for analysis on an Agilent 6890N Network GC System with a flame ionisation detector (FID) (Agilent Technologies, USA), equipped with a CP-Chiralsil-Dex CB column (25 m × 0.25 mm × 0.25 μm) (Varian, Germany) and hydrogen as the carrier gas. The injection temperature was set to 150 °C for benzaldehyde (**1a**), 160 °C for halosubstituted benzaldehydes (**1b–j** and **1l–n**), and 170 °C for methoxybenzaldehyde derivatives (**1k**, **o**), respectively; this followed an isotherm run for 10 min.

Acknowledgements

Financial support by the German Research Foundation (DFG) in frame of the research group FOR1296 "Diversity of asymmetric thiamine catalysis", and partial funding by the research training group "BioNoCo" (GK1166) is gratefully acknowledged. The author thanks Simon Waltzer (Institute of Pharmaceutical Sciences, University of Freiburg) for his skillful technical support.

Keywords: asymmetric catalysis · biocatalysis · enantioselectivity · enzyme catalysis · protein engineering

[1] a) M. Müller, G. A. Sprenger, M. Pohl, *Curr. Opin. Chem. Biol.* **2013**, *17*, 261 – 270; b) M. Pohl, C. Dresen, M. Beigi, M. Müller in *Enzyme Catalysis in Organic Synthesis, Vol. 3* (Eds.: K. Drauz, H. Gröger, O. May), Wiley-VCH, Weinheim, **2012**, pp. 919 – 946; c) M. Müller, *Adv. Synth. Catal.* **2012**, *354*, 3161 – 3174; d) M. Brovetto, D. Gamenara, P. S. Méndez, G. A. Seoane, *Chem. Rev.* **2011**, *111*, 4346 – 4403.
[2] P. Hoyos, J. V. Sinisterra, F. Molinari, A. R. Alcantara, P. D. de Maria, *Acc. Chem. Res.* **2010**, *43*, 288 – 299.
[3] M. Müller, D. Gocke, M. Pohl, *FEBS J.* **2009**, *276*, 2894 – 2904.
[4] a) R. G. Duggleby, *Acc. Chem. Res.* **2006**, *39*, 550 – 557; b) M. Widmann, R. Radloff, J. Pleiss, *BMC Biochem.* **2010**, *11*, 9.
[5] M. Knoll, M. Müller, J. Pleiss, M. Pohl, *ChemBioChem* **2006**, *7*, 1928 – 1934.
[6] A. S. Demir, M. Pohl, E. Janzen, M. Müller, *J. Chem. Soc. Perkin Trans. 1* **2001**, 633 – 635.
[7] R. Wilcocks, O. P. Ward, *Biotechnol. Bioeng.* **1992**, *39*, 1058 – 1063.
[8] M. Pohl, D. Gocke, M. Müller in *Handbook of Green Chemistry—Green Catalysis, Vol. 3* (Ed: P. T. Anastas) Wiley-VCH, Weinheim, **2008**, pp. 75 – 114.
[9] a) D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C. L. Berthold, G. Schneider, J. Pleiss, M. Müller, M. Pohl, *ChemBioChem* **2008**, *9*, 406 – 412; b) D. Rother (neé Gocke), G. Kolter, T. Gerhards, C. L. Berthold, E. Gauchenova, M. Knoll, J. Pleiss, M. Müller, G. Schneider, M. Pohl, *ChemCatChem* **2011**, *3*, 1587 – 1596.
[10] a) R. Westphal, S. Waltzer, U. Mackfeld, M. Widmann, J. Pleiss, M. Beigi, M. Müller, D. Rother, M. Pohl, *Chem. Commun.* **2013**, *49*, 2061 – 2063; b) R. Westphal, D. Hahn, U. Mackfeld, S. Waltzer, M. Beigi, M. Widmann, C. Vogel, J. Pleiss, M. Müller, D. Rother, M. Pohl, *ChemCatChem* **2013**, DOI: 10.1002/cctc.201300318.
[11] M. Beigi, S. Waltzer, A. Fries, L. Eggeling, G. A. Sprenger, M. Müller, *Org. Lett.* **2013**, *15*, 452 – 455.
[12] M. Jiang, Y. Cao, Z.-F. Guo, M. Chen, X. Chen, Z. Guo, *Biochemistry* **2007**, *46*, 10979 – 10989.
[13] A. Kurutsch, M. Richter, V. Brecht, G. A. Sprenger, M. Müller, *J. Mol. Catal. B* **2009**, *61*, 56 – 66.
[14] A. Dawson, M. Chen, P. K. Fyfe, Z. Guo, W. N. Hunter, *J. Mol. Biol.* **2010**, *401*, 253 – 264.
[15] C. Vogel, M. Widmann, M. Pohl, J. Pleiss, *BMC Biochem.* **2012**, *13*, 24.
[16] T. Gerhards, U. Mackfeld, M. Bocola, E. von Lieres, W. Wiechert, M. Pohl, D. Rother, *Adv. Synth. Catal.* **2012**, *354*, 2805 – 2820.
[17] B. Lingen, J. Grötzinger, D. Kolter, M.-R. Kula, M. Pohl, *Protein Eng.* **2002**, *15*, 585 – 593.
[18] J. F. Sambrook, D. W. Russel in *Molecular cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, **2001**.
[19] *The PyMOL Molecular Graphics System*, http://www.pymol.org, W. L. DeLano, San Carlo, CA, USA, **2002**.
[20] H. B. Burgi, J. D. Dunitz, E. Shefter, *J. Am. Chem. Soc.* **1973**, *95*, 5065 – 5067.
[21] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminf.* **2012**, *4*, 17.
[22] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248.

Received: August 16, 2013
Published online on ■■ ▮▮, 0000

Supporting Information

Table of contents

1. Gene sequence of codon-optimized *Bs*MenD

The original *menD* from *B. subtilis* (KEGG gene entry: bsu:BSU30820) was codon-optimized for *E. coli* using the software Optimizer.[1] The final gene (Figure S1), containing a NdeI and BamHI restriction site, was purchased from GeneArt® Gene Synthesis (life technologies™).

CATATGACGGTAAATCCGATTACCCACTACATTGGTTCTTTTATCGACGAATTCGCCCTGAGTGGC
ATCACCGATGCTGTAGTGTGCCCTGGTTCCCGCTCCACTCCGTTAGCCGTACTGTGCGCGGCTC
ACCCGGATATCTCGGTGCACGTCCAGATCGATGAACGCTCCGCGGGCTTTTTCGCTCTGGGCCT
GGCGAAAGCAAAACAACGTCCGGTGCTGCTGATCTGCACCTCAGGCACTGCTGCTGCGAACTTT
TACCCGGCTGTTGTTGAAGCACACTACAGCCGTGTTCCGATCATCGTGCTGACCGCAGATCGTC
CGCACGAGCTGCGTGAGGTGGGCGCCCCGCAGGCTATCAACCAGCACTTTTTATTCGGTAACTT
CGTAAAATTTTTCACTGACTCTGCACTGCCGGAAGAGTCCCCGCAGATGCTGCGTTACATCCGTA
CTCTGGCATCGCGTGCAGCTGGCGAAGCACAGAAACGCCCGATGGGTCCTGTTCATGTAAACGT
TCCACTGCGCGAACCACTGATGCCGGATCTGTCTGACGAACCGTTCGGTCGTATGCGCACGGGC
CGCCACGTTAGCGTGAAAACTGGTACTCAGTCGGTTGATCGTGAAAGCCTGTCTGATGTTGCCGA
AATGCTGGCGGAAGCGGAAAAAGGCATGATTGTATGCGGTGAACTGCACAGCGACGCAGACAAA
GAAAACATTATCGCACTGAGTAAAGCACTGCAATACCCGATCCTGGCTGATCCGCTGTCCAATCT
CCGTAACGGTGTGCACGACAAAAGCACCGTGATCGACGCCTACGACTCTTTCCTGAAAGATGAC
GAACTGAAACGCAAGCTGCGTCCGGATGTAGTGATCCGCTTCGGTCCGATGCCTGTCTCTAAAC
CAGTATTTCTCTGGCTGAAAGACGATCCGACCATCCAGCAGATCGTAATCGATGAAGATGGAGGT
TGGCGCGATCCGACGCAGGCGTCTGCTCACATGATTCATTGCAACGCGTCGGTTTTCGCCGAGG
AAATCATGGCGGGCCTGACCGCTGCGACCCGTTCCTCTGAGTGGCTGGAAAAATGGCAGTTCGT
GAACGGTCGTTTCCGCGAACATTTGCAGACCATCTCTTCTGAAGACGTCAGCTTCGAAGGTAACC
TGTATCGTATCCTGCAGCACCTGGTTCCGGAAAACAGCTCCCTGTTTGTAGGTAACAGCATGCCG
ATCCGTGATGTCGATACGTTCTTCGAGAAGCAAGATCGCCCGTTTCGTATCTATTCTAACCGTGG
CGCAAACGGTATTGACGGCGTTGTTTCTTCTGCGATGGGCGTTTGCGAAGGGACCAAAGCGCCG
GTGACGCTGGTCATCGGAGACCTGTCCTTCTACCACGATCTGAACGGCCTGCTGGCTGCTAAAA
AACTGGGCATTCCTCTGACCGTTATCCTGGTTAACAACGACGGTGGTGGTATTTTCTCCTTCCTG
CCGCAAGCGAGTGAAAAAACCCATTTTGAAGACCTGTTCGGGACCCCGACTGGCCTGGATTTCA
AACACGCGGCTGCTCTGTACGGCGGCACTTACTCGTGCCCTGCGTCTTGGGACGAATTCAAGAC
TGCTTACGCGCCGCAAGCAGATAAACCGGGTCTGCACCTGATCGAGATTAAGACCGATCGTCAA
TCTCGTGTGCAGCTGCACCGTGACATGCTGAACGAAGCAGTTCGTGAAGTTAAGAAACAGTGGG
AGCTCTAAGGATCC

Figure S1. Codon-optimized *Bs*MenD gene. NdeI (catatg) and BamHI (ggatcc) restriction sites are highlighted in red.

2. Amino acid sequences of *Bs*MenD and *Ec*MenD

The amino acid sequences used in this work were derived from pET19b_*Bs*MenD and previously described pET19b_*Ec*MenD constructs,[2,3] which provide a 10×histidine tag. Absolute amino acid positions were counted based on this sequence starting with the first methionine of the original wild type sequences (Figure S2). Assignment of the absolute amino acid positions to the standard numbers (according to the standard numbering scheme)[4] is shown in Table 1 of the publication for all in this work relevant amino acid residues.

A Amino acid sequence of *Bs*MenD

MGHHHHHHHHHHSSGHIDDDDKH**M**TVNPITHYIGSFIDEFALSGITDAVVCPGSRSTPLAVL
CAAHPDISVHVQIDERSAGFFALGLAKAKQRPVLLICTSGTAAANFYPAVVEAHYSRVPIIVLT
ADRPHELREVGAPQAINQHFLFGNFVKFFTDSALPEESPQMLRYIRTLASRAAGEAQKRPM
GPVHVNVPLREPLMPDLSDEPFGRMRTGRHVSVKTGTQSVDRESLSDVAEMLAEAEKGMI
VCGELHSDADKENIIALSKALQYPILADPLSNLRNGVHDKSTVIDAYDSFLKDDELKRKLRPD
VVIRFGPMPVSKPVFLWLKDDPTIQQIVIDEDGGWRDPTQASAHMIHCNASVFAEEIMAGLT
AATRSSEWLEKWQFVNGRFREHLQTISSEDVSFEGNLYRILQHLVPENSSLFVGNSMPIRD
VDTFFEKQDRPFRIYSNRGANGIDGVVSSAMGVCEGTKAPVTLVIGDLSFYHDLNGLLAAKK
LGIPLTVILVNNDGGGIFSFLPQASEKTHFEDLFGTPTGLDFKHAAALYGGTYSCPASWDEFK
TAYAPQADKPGLHLIEIKTDRQSRVQLHRDMLNEAVREVKKQWEL

1 580

B Amino acid sequence of *Ec*MenD[3]

MGHHHHHHHHHHSSGHIDDDDKH**M**SVSAFNRRWAAVILEALTRHGVRHICIAPGSRSTPLT
LAAAENSAFIHHTHFDERGLGHLALGLAKVSKQPVAVIVTSGTAVANLYPALIEAGLTGEKLIL
LTADRPPELIDCGANQAIRQPGMFASHPTHSISLPRPTQDIPARWLVSTIDHALGTLHAGGVH
INCPFAEPLYGEMDDTGLSWQQRLGDWWQDDKPWLREAPRLESEKQRDWFFWRQKRGV
VVAGRMSAEEGKKVALWAQTLGWPLIGDVLSQTGQPLPCADLWLGNAKATSELQQAQIVV
QLGSSLTGKRLLQWQASCEPEEYWIVDDIEGRLDPAHHRGRRLIANIADWLELHPAEKRQP
WCVEIPRLAEQAMQAVIARRDAFGEAQLAHRICDYLPEQGQLFVGNSLVVRLIDALSQLPAG
YPVYSNRGASGIDGLLSTAAGVQRASGKPTLAIVGDLSALYDLNALALLRQVSAPLVLIVVNN
NGGQIFSLLPTPQSERERFYLMPQNVHFEHAAAMFELKYHRPQNWQELETAFADAWRTPT
TTVIEMVVNDTDGAQTLQQLLAQVSHL

1 556

Figure S2. Amino acid sequences of *Bs*MenD (**A**) and *Ec*MenD (**B**) gene. The first methionine of the original sequence is shown in bold red.

3. *S*-pocket conservation of *Bs*MenD and *Ec*MenD and homologues proteins

In addition to Dawson et al.,[5] who compared sequence homology of *Bs*MenD and *Ec*MenD, we investigated the amino acid distribution in the *S*-pocket region of all MenD homologues. Therefore, 507 homologues protein sequences of *Ec*MenD and 1126 sequence homologues of *Bs*MenD were collected from the updated version of the ThDP-dependent Enzyme Engineering Database and subsequently aligned using the standard numbering scheme for ThDP-dependent decarboxylases.[4,6] The highest global sequence similarity between two sequences of both subsets determined using the EMBOSS *needle* tool[7] was 57.6%. At standard position 475, glycine is conserved (> 77%) in *Bs*MenD homologues, whereas in *Ec*MenD homologues glycine is only present in 9% of all sequences. Instead, 51% of the *Ec*MenD homologues have glutamine and 30% carry alanine at the respective position. In contrast to the findings of Dawson et al.,[5] standard position 478 is only in subsets of both protein families occupied by serine (38% of the *Bs*MenD homologues and 55% of the *Ec*MenD homologues).

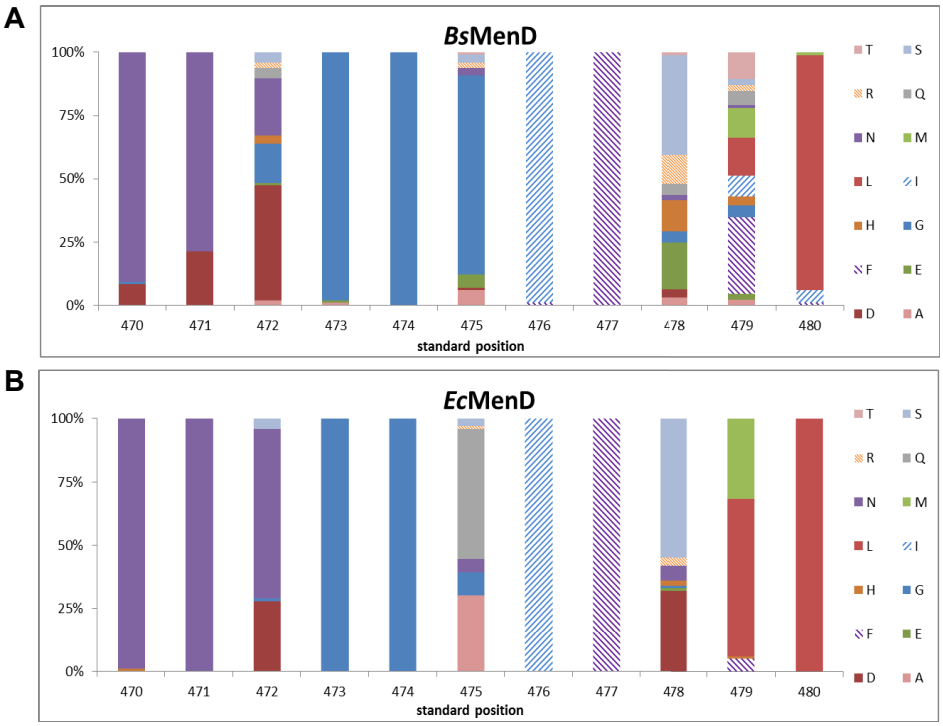


Figure S3. Amino acid distributions in the *S*-pocket regions of 1126 *Bs*MenD (**A**) and 507 *Ec*MenD (**B**) homologues. High analogy between both families is evident at positions 470, 473, 474, 476, 477 and 480 (standard numbers). The most obvious difference between both *S*-pocket regions was observed at standard position 475. Glycine is conserved in 78% of all *Bs*MenD homologous, whereas in *Ec*MenD homologous glutamine (51%) and alanine (30%) are the most prominent amino acids and glycine only plays a minor role (9%).

4. Generation of *Bs*MenD variants

The amino acid positions of generated variants are given as absolute position (see Figure S2 A). *Bs*MenD variant I489G/F490G was prepared based on the original pET-19b_*Bs*MenD construct (see above) using the forward (fw) primer 5'-CAACGACGGTGGTGGTggTggCTCCTTCCTGCCGCAAG-3' and the reverse (rv) primer 5'-CTTGCGGCAGGAAGGAGccAccACCACCACCGTCGTTG-3' (mutated codons are underlined, with lower-case letters indicating base changes). Based on *Bs*MenD-I489G/F490G the mutants *Bs*MenD-I489A/F490G, *Bs*MenD-I489G/F490A, *Bs*MenD-I489A/F490A, and *Bs*MenD-I489A/F490G/G488Q were prepared. The following fw and rv primers were used for mutagenesis:

*Bs*MenD-I489A/F490G:	5'-CAACGACGGTGGTGGTGcgGGC TCCTTCCTG-3' (fw)
	5'-CAGGAAGGAGCCcgCACCACC ACCGTCGTTG -3' (rv)
*Bs*MenD-I489G/F490A:	5'-GGTGGTGGTGGTGcCTCCTTCCTGC-3' (fw)
	5'-GCAGGAAGGAGgCACCACCACCACC-3' (rv)
*Bs*MenD-I489A/F490A:	5'-CAACGACGGTGGTGGTGcgGcCTCC TTCCTGCCGCAAG-3' (fw)
	5'-CTTGCGGCAGGAAGGAGgCcgCACC ACCACCGTCGTTG-3' (rv)
*Bs*MenD-I489G/F490G/G488Q:	5'-GTTAACAACGACGGTGGTcagGGT GGCTCCTTCCTGCCG-3' (fw)
	5'-CGGCAGGAAGGAGCCACCctgACC ACCGTCGTTGTTAAC-3' (rv).

5. References

[1] P. Puigbo, E. Guzmán, A. Romeu, S. Garcia-Vallvé, *Nucleic Acids Res.* **2007**, *35*, W126–W131.

[2] A. Kurutsch, M. Richter, V. Brecht, G. A. Sprenger, M. Müller, *J. Mol. Catal. B: Enzym.* **2009**, *61*, 56–66.

[3] R. Westphal, S. Waltzer, U. Mackfeld, M. Widmann, J. Pleiss, M. Beigi, M. Müller, D. Rother, M. Pohl, *Chem. Commun.* **2013**, *49*, 2061–2063.

[4] C. Vogel, M. Widmann, M. Pohl, J. Pleiss, *BMC Biochem.* **2012**, *13*, 24.

[5] A. Dawson, M. Chen, P. K. Fyfe, Z. Guo, W. N. Hunter, *J. Mol. Biol.* **2010**, *401*, 253–264.

[6] M. Widmann, R. Radloff, J. Pleiss, *BMC Biochem.* **2010**, *11*, 9.

[7] P. Rice, I. Longden, A. Bleasby, *Trends in Genetics* **2000**, *16*, 276–277.

PUBLICATION V

Tailor-made Pyruvate Decarboxylase for the Direct Asymmetric Synthesis of (*S*)-Benzoins

Westphal, R.; Vogel, C.; Schmitz, C.; Pleiss, J.; Müller, M.;. Pohl, M. & Rother, D.

to be submitted

Tailor-made Pyruvate Decarboxylase for the Direct Asymmetric Synthesis of (*S*)-Benzoins

Robert Westphal,[‡,] Constantin Vogel,[I,] Carlo Schmitz,[‡] Jürgen Pleiss,[I] Michael Müller,[II] Martina Pohl,[*,‡] and Dörte Rother[‡]

[‡]Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Leo-Brandt-Str. 1, D-52425 Jülich, Germany.
[I]Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, D-70569 Stuttgart, Germany.
[II]Institute of Pharmaceutical Sciences, Albert-Ludwigs-University Freiburg, Albertstr. 25, D-79104 Freiburg, Germany.
[*]Corresponding author.

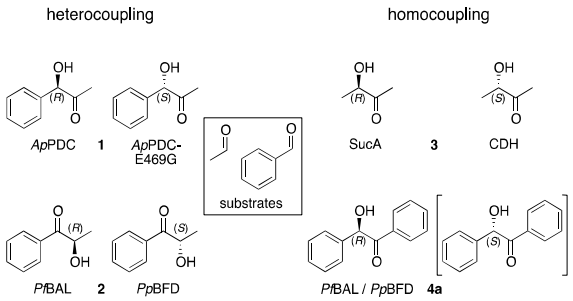
ABSTRACT: Thiamine diphosphate-dependent enzymes are well known to catalyze decarboxylations and asymmetric C–C bond formations. Many different α-hydroxy ketones are thus available through enzymatic chemo- and stereoselective ligase reactions. Nevertheless, the steric and chemical properties of the active site of the enzymes often limits the reaction and product ranges. Whereas benzaldehyde lyase (BAL) from *Pseudomonas fluorescens* is a well-known catalyst for the synthesis of (*R*)-benzoins starting from benzaldehydes, no enzymes have been identified as yet to catalyze the asymmetric synthesis of (*S*)-benzoin. Here, we report on a rational hybridization approach to solve this long-standing problem. To generate a suitable substrate-binding site for (*S*)-benzoin formation, we combined the large *S*-pocket (prerequisite for *S*-selective carboligation) of a variant of pyruvate decarboxylase (PDC) from *Acetobacter pasteurianus* with the large donor-binding site (prerequisite for benzoin synthesis) of benzaldehyde lyase. This approach resulted in a PDC variant with a hybrid substrate-binding site providing access to (*S*)-benzoin (59 % *ee*) by enzymatic asymmetric synthesis. The moderate *S*-selectivity of the initial variant was improved by selective destabilization of the "*R*-pathway". This challenging goal was achieved by two additional mutations in the active site yielding variants with high catalytic activity and excellent stereoselectivity for various (*S*)-benzoins (> 99 % *ee*).

INTRODUCTION

The challenging catalytic stereoselective synthesis of α-hydroxy ketones via asymmetric C–C bond formation can be achieved by thiamine diphosphate (ThDP)-dependent enzymes.[1-4] The chemical mechanism,[5] structure–function relationships,[6,7] and the extraordinary catalytic potential of ThDP-dependent enzymes have been intensively explored during the last decades. Scheme 1 exemplarily demonstrates the high diversity of ThDP-dependent enzymes in the asymmetric carboligation of acetaldehyde and benzaldehyde.[8] Depending on the donor, which binds to the ThDP-cofactor prior to C–C bond formation, and the acceptor substrate approaching subsequently to the carbanion-enamine complex, four different products, each in both enantiomeric forms, are possible (Scheme 1). Phenylacetylcarbinol (PAC, **1**) and acetoin (**3**) are derived from acetaldehyde as the donor and benzaldehyde or acetaldehyde as acceptor, respectively. 2-Hydroxy-propiophenone (2-HPP, **2**) and benzoin (**4a**) result from benzaldehyde as the donor and acetaldehyde or benzaldehyde as acceptor, respectively.

The chemoselectivity of ThDP-dependent enzymes belonging to the decarboxylase superfamily (classification according to Duggleby 2006 and Widmann et al. 2010)[9,10] is predominantly determined by the space accessible by the donor and acceptor substrates in the active site.[8] As a consequence, ThDP-dependent enzymes are usually highly selective for one or two of the possible products, because of their specific active site architectures (Figure 1).

Scheme 1. Diversity of ThDP-dependent enzymes in the asymmetric synthesis of α-hydroxy ketones[8]



*Pf*BAL, benzaldehyde lyase from *Pseudomonas fluorescens*; *Pp*BFD, benzoylformate decarboxylase from *Pseudomonas putida*; CDH, cyclohexane-1,2-dione hydrolase from *Azoarcus* sp.; *Ap*PDC, pyruvate decarboxylase from *Acetobacter pasteurianus*; SucA, E1 component of the α-ketoglutarate dehydrogenase subunit from *Escherichia coli* K12.

Steric properties of the active site are also responsible for the different stereoselectivity (Figure 1). Introduction of the *S*-pocket concept,[8,11] which explains stereoselectivity by the orientation of the acceptor substrate to the ThDP-bound donor substrate prior to carboligation, enabled the rational design of different *S*-selective enzyme variants with high enantiomeric excesses (*ee*).[12–15] *Inter alia*, the long-standing problem of formation of (*S*)-**1** was solved recently.[13]

Despite the tremendous number of accessible α-hydroxy ketones and the apparently unlimited synthetic possibilities of ThDP-dependent enzymes concerning C–C bond formation, a solution for the enzymatic synthesis of (*S*)-benzoins (**4**) starting from benzaldehyde (derivatives) has not been found yet (Figure 1). Instead, other enzymatic routes with purified enzyme or whole cell catalysts to (*S*)-**4** and derivatives became available, which comprise kinetic and dynamic kinetic resolution of *rac*-**4**,[16,17] asymmetric reduction of 1,2-diarylethane-1,2-diones (benzils),[18] or deracemization of *rac*-**4**.[19] Although theses enzymatic approaches are mostly characterized by high stereoselectivities and yields (except for kinetic resolution), they all share the same drawback that the racemic or prochiral starting material must be chemically synthesized beforehand. Such additional steps reduce the eco-efficiency and sustainability of the process compared to a potential direct enzymatic synthesis of (*S*)-**4** starting from readily available benzaldehyde derivatives.

Whereas the enzyme-catalyzed benzoin condensation towards (*R*)-**4** was solved successfully,[20,21,22] the *S*-selective synthesis of **4** via asymmetric C–C bond formation is a unique feature of nonenzymatic synthetic organic chemistry. In 2002, Enders and Kallfass demonstrated the chemical synthesis of homochiral (*S*)-**4** using a triazolium precatalyst.[23]

The limitations of ThDP-dependent enzyme-catalyzed synthesis of (*S*)-**4** are versatile. Although there are inherently *S*-selective enzymes, e.g. in the transketolase superfamily,[8] they do not catalyze formation of (*S*)-**4** because of their limited donor substrate spectrum. The restrictions for the formation of (*S*)-**4** in the decarboxylase superfamily, in compliance with the *S*-pocket concept, are summarized in Figure 1.

On the one hand, a highly *S*-selective variant of the pyruvate decarboxylase from *Acetobacter pasteurianus* (*Ap*PDC) for the formation of (*S*)-**1** has become available recently.[13] By replacement of E469 by glycine, the *S*-pocket was widely opened, which in turn allowed benzaldehyde as an acceptor to bind predominantly in antiparallel orientation, the prerequisite for *S*-selectivity (Figure 1 A). However, *Ap*PDC-E469G only provides a small donor-binding site that preferably stabilizes small aliphatic donor substrates (e.g. acetaldehyde or propanal). As a consequence, significant formation of **4a** via homocoupling of benzaldehyde was not observed.[13] On the other hand, powerful but strictly *R*-selective catalysts for the synthesis of **4a** are known: benzaldehyde lyase from *Pseudomonas fluorescens* (*Pf*BAL) and benzoylformate decarboxylase from *Pseudomonas putida* (*Pp*BFD).[20,21] In comparison to PDC, both enzymes provide a large donor-binding site, ideal to stabilize benzaldehyde in that position (Figure 1 B, C), but they only allow the parallel arrangement of the donor and acceptor benzaldehydes prior to carboligation. Furthermore, *S*-pocket engineering is limited in *Pf*BAL and *Pp*BFD because of the backbone positions of the respective α-helix in the *S*-pocket. Consequently, the *S*-pocket of *Pp*BFD variant L461G could only be opened for aliphatic acceptor substrates smaller than butanal yielding (*S*)-**2** and derivatives thereof (Figure 1 C).[12] Accordingly, the required space for the phenyl ring to obtain *S*-selective formation of **4a** could not be achieved through mutagenesis in both enzymes.

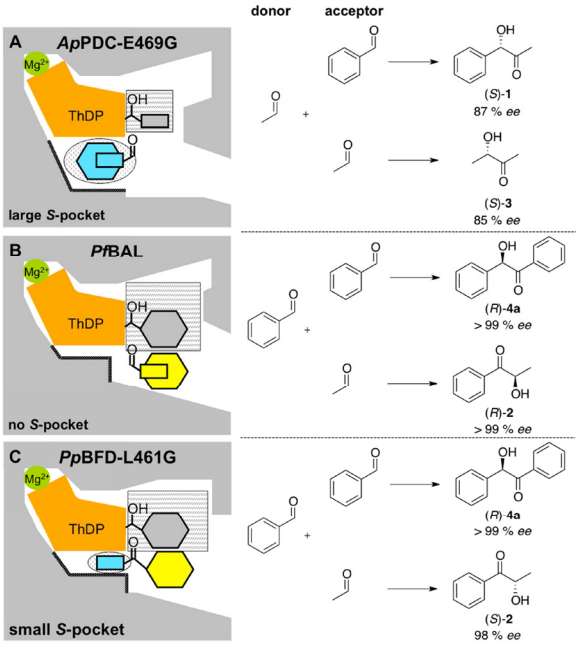


Figure 1. Scope and limitations of ThDP-dependent enzyme-catalyzed synthesis of **1–4a** schematically shown for *Ap*PDC-E469G (**A**), *Pf*BAL (**B**), and *Pp*BFD-L461G (**C**). Either an unsuitable *S*-pocket for antiparallel-oriented benzaldehyde (**B**, **C**), or a small donor-binding site (**A**), prevents (*S*)-**4a** formation. Dashed rectangle: donor binding site; dotted circle: *S*-pocket region; rectangle: acetaldehyde; hexagon: benzaldehyde; colors: gray: donor; cyan: antiparallel-oriented acceptor leading to the respective (*S*)-product; yellow: parallel-oriented acceptor leading to the respective (*R*)-product.

We hypothesized that the design of a hybrid enzyme, which combines the large *S*-pocket of the *S*-selective *Ap*PDC variant E469G and the large donor-binding site of *Pf*BAL or *Pp*BFD, could provide a solution for the long-standing problem of enzymatic (*S*)-**4** formation. Two approaches are conceivable: (i) the introduction of a large *S*-pocket into *Pf*BAL or *Pp*BFD, or (ii) the extension of the donor-binding site of *Ap*PDC-E469G. Here, we demonstrate that the second approach indeed resulted in the first tailor-made *S*-selective ThDP-dependent enzyme variant for the formation of **4** showing remarkable catalytic activity and stereoselectivity.

MATERIAL AND METHODS

Materials. All chemicals were purchased from Sigma Aldrich. Benzaldehyde was freshly distilled before use. KOD Hot Start DNA Polymerase and the restriction enzyme *DpnI* were obtained from Merck Millipore and Fermentas (Thermo Scientific), respectively.

Sequence analysis. The homologous families containing *Ap*PDC and *Pf*BAL of an up-to-date version of the ThDP-dependent enzymes engineering database[10] were considered for conservation analysis (unpublished data). 186 sequences of *Ap*PDC homologues and 43 homologous BAL sequences were aligned using the standard numbering approach for ThDP-dependent decarboxylases[24] and subsequently analyzed regarding their amino acid distribution. The absolute position numbers of the *Ap*PDC sequence[13] and *Pf*BAL sequence (gi|9965497) were counted increasingly ordered beginning with methionine in position 1.

Structural analysis. The PyMOL software (Schrödinger)[25] was used for structural investigations of *Ap*PDC (pdb code: 2vbi), *Ap*PDC variants, and *Pf*BAL (pdb code: 2ag0). The hydroxybenzyl-ThDP (HBThDP) intermediate was constructed using the ArgusLab software (Mark A. Thompson, Planaria Software LLC, Seattle, USA) and modeled into the active sites of the superimposed *Ap*PDC and *Pf*BAL structures. Both structures were then manually superimposed by aligning the ThDP molecules co-crystallized with the respective proteins. The YASARA software (YASARA Biosciences GmbH, Austria) was applied to equilibrate the structure of *Ap*PDC-E469G/T384G (Figure 2 B) using the YASARA2 force field.[26] Therefore, the structure was embedded in a water box with 2 Å distance to the protein and prepared by applying the standard energy minimization protocol. A subsequent simulation run for 0.5 ns was performed in order to equilibrate the structure of the protein and the substrate. A constraint was set to the HBThDP intermediate to keep the planar orientation of the phenyl ring of benzaldehyde, which is due to mesomeric stabilization in the physiological structure.

Generation and Expression of *Ap*PDC Variants. The plasmid pET22b (Novagen) containing the gene encoding *Ap*PDC-E469G[13] was used as template for the generation of *Ap*PDC variants. All variants were constructed as described elsewhere using the standard Quikchange® site-directed mutagenesis protocol (Stratagene).[27] Mutagenic primers are given in the Supporting Information. After digestion of parental DNA by *DpnI*, chemically competent *E. coli* BL21 (DE3) cells were transformed by the resulting PCR products using the $CaCl_2$-method according to Sambrook and Russell.[28] Gene sequences were confirmed by DNA sequencing (LGC Genomics, Berlin, Germany).

Cultivation of the variants was performed in shaking flasks (5 L) at initially 30 °C using 1 L of lysogeny broth (LB) medium, pH 7.0, supplemented with ampicillin (100 µg µL^{-1}). Expression of the pET22b_*Ap*PDC constructs, which provide a C-terminal 6 × histidine tag, was induced by addition of isopropyl-β-D-1-thio-galactopyranoside (0.4 mM) at an OD_{600} of 0.6. After induction, cells were grown overnight at 20 °C, subsequently harvested by centrifugation (30 min, 15.000 *g*, 4 °C), and finally stored at −20 °C.

Purification of *Ap*PDC variants. For purification cells were resuspended in disintegration buffer (50 mM triethanolamine (TEA), 150 mM NaCl, 2 mM $MgSO_4$, 0.1 mM ThDP, 20 mM imidazole, pH 7.5) supplemented with lysozyme (1 mg mL^{-1}) on ice. After cell disruption by sonication, cell debris was removed by centrifugation (45 min, 48.000 *g*, 4 °C). His-tagged *Ap*PDC variants were then purified by immobilized nickel chelate chromatography: washing buffer (50 mM TEA, 150 mM NaCl, 50 mM imidazole, pH 7.5); elution buffer (50 mM TEA, 150 mM NaCl, 250 mM imidazole, pH 7.5). Subsequently, purified enzymes were desalted by size exclusion chromatography using Sephadex™ G25M (GE Healthcare), equilibrated with a low salt buffer (10 mM TEA, 2 mM $MgSO_4$, 0.1 mM ThDP, pH 7.5), and freeze-dried using the lyophilizer ALPHA 2–4 (Christ, Germany). Lyophilized *Ap*PDC variants were stored at −20 °C for several months without significant loss of activity.

Site-Saturation Mutagenesis. A single site-saturation mutagenesis was performed for amino acid position 543 in *Ap*PDC-E469G/T384G/I468A using the Quikchange® standard protocol (see above). The mutagenic primer, containing NDT degeneracy (where N= any nucleotide, D= A, G, or T), is shown in the Supporting Information. Library preparation, expression of the variants, and cell disruption were performed as previously described.[27] For the carboligation assay in a 96 deep well plate (max. volume of 2 mL; Nerbe plus, Germany), 200 µL of the respective supernatant was mixed with 300 µL of a benzaldehyde stock solution (30 mM benzaldehyde in reaction buffer; see below) giving a final concentration of 18 mM benzaldehyde. 50 µL as well as 100 µL of the reaction mixture were then transferred to a new 96 deep well plate. The plates were covered with a silicone cap-mat (Nerbe plus, Germany) and incubated at 20 °C and 600 rpm (VARIOMAG Shaker, H + P Labortechnik, Germany) for 24 h. Subsequently, formation of **4a** was analyzed as described elsewhere via the triphenyl tetrazolium chloride assay using the 100 µL batches.[27, 29] In the case of positive hits, indicated by red coloration, the reaction mixture was extracted with 500 µL methyl *tert*-butyl ether (MTBE) from the residual 50 µL batch and the organic phase was analyzed by chiral-phase HPLC (see below).

Carboligation Reactions. All carboligation reactions were performed in reaction buffer (50 mM TEA, 2 mM $MgSO_4$, 0.1 mM ThDP, pH 8.0 if not otherwise indicated) using lyophilized enzyme. The protein concentration was determined according to Bradford using bovine serum albumin as standard.[30]

Analytical Scale. Carboligation reactions were performed in a total volume of 300 µL in glass vials at 800 rpm using a thermomixer (Eppendorf, Germany). The final protein concentration was set to 1 mg mL^{-1}. The enzyme was incubated with benzaldehyde and benzaldehyde derivatives (18 mM), respectively, for the homocoupling reaction, or with acetaldehyde (18 mM) and benzaldehyde (18 mM) for the heterocoupling reaction. Subsequently, products were extracted using 300 µL of MTBE. Finally, the organic phase was used for product analysis (see below). Reactions were performed in triplicates. Reaction time and temperature were varied as indicated. No C-C bond formation was observed in control experiments without enzyme.

Preparative Scale. *Ap*PDC variants E469G/T384G/I468A and E469G/T384G/I468A/W543F (1 mg mL^{-1} final concentration) were incubated in 20 mL reaction buffer (see above) with the respective benzaldehyde derivative (18 mM) at 15 °C and 750 rpm in a 50 mL Falcon tube using a thermomixer (Eppendorf, Germany). The reaction was monitored for 24–48 h until conversions > 90 % were achieved. In case of *Ap*PDC-E469G/T384G/I468A/W543F the reaction was supplemented with fresh enzyme after 12 h and 24 h. The reaction mixtures were then extracted three times using dichloromethane (20 mL). Afterwards, the products were purified as described elsewhere.[21] Crystallization afforded the respective (*S*)-**4** derivatives. Isolated yields (Table 2) were calculated based on the respective benzaldehyde derivative.

Product Analysis. The identity of (*S*)-**4** was verified by ^{1}H and ^{13}C NMR spectroscopy. Spectral data were identical with previously reported values.[21] Analytics of **1–3** were performed as described elsewhere.[31, 32]

Determination of Enantiomeric Excess and Conversion. The enantiomeric excesses of **4a–4l** were determined by chiral-phase HPLC using a 1260 Infinity chromatography system (Agilent Technologies, USA) equipped with a Daicel Chiralcel OD-H 5 µm (250 mm × 4.6 mm) column (Daicel Chemical IND., France). Products were detected at 254 nm using the eluent *n*-hexane:2-

propanol (92:8) at a flow rate of 1.2 mL min^{-1} and 20 °C. Typical retention times were: **4a**: $t_R(S)$ = 10.1 min, $t_R(R)$ = 14.8 min; **4b**: $t_R(S)$ = 8.5 min, $t_R(R)$ = 11.4 min; **4c**: $t_R(S)$ = 11.5 min, $t_R(R)$ = 14.9 min; **4d**: $t_R(S)$ = 17.5 min, $t_R(R)$ = 25.7 min; **4e**: $t_R(S)$ = 8.5 min, $t_R(R)$ = 11.1 min; **4f**: $t_R(S)$ = 9.0 min, $t_R(R)$ = 12.8 min; **4g**: $t_R(S)$ = 10.3 min, $t_R(R)$ = 14.2 min; **4h**: $t_R(S)$ = 11.5 min, $t_R(R)$ = 15.0 min; **4i**: $t_R(S)$ = 15.7 min, $t_R(R)$ = 23.1 min; **4j**: $t_R(S)$ = 8.8 min, $t_R(R)$ = 9.5 min; **4k**: $t_R(S)$ = 6.5 min, $t_R(R)$ = 9.9 min; **4l**: $t_R(S)$ = 9.7 min, $t_R(R)$ = 22.5 min.

Conversions in the carboligation of **4a–4l** were determined by chiral-phase HPLC (see above) based on the consumption of benzaldehyde or benzaldehyde derivatives. Typical retention times of benzaldehyde and derivatives were 4.5–5.5 min.

Verification of Absolute Configuration. The retention times of the respective enantiomers of **4a** could be unambiguously assigned on the basis of commercially available (*S*)-**4a** and (*R*)-**4a**. Enantiomers of **4b–4l** were assigned using the *R*-selective carboligation product catalyzed by *Pf*BAL as reference.[21] In addition, the absolute configurations of **4a** and **4e–i** were verified by determination of its optical rotations using a Jasco P-2000 polarimeter (Jasco, Germany). **4a**: $[\alpha]_D^{20}$: +154.1 (*c* 1.5, CH_3CN), [Lit.[33] $[\alpha]_D^{20}$: +138.4 (*c* 0.25, $CHCl_3$) for 89 % *ee*]. **4e**: $[\alpha]_D^{20}$: +71.5 (*c* 1.5, CH_3CN); **4f**: $[\alpha]_D^{20}$: +63.9 (*c* 1.5, CH_3CN); **4g**: $[\alpha]_D^{20}$: +49.3 (*c* 1.5, CH_3CN); **4h**: $[\alpha]_D^{20}$: +5.2 (*c* 1.5, CH_3CN); **4i**: $[\alpha]_D^{20}$: +116.4 (*c* 0.25, CH_3CN).

RESULTS AND DISCUSSION

Hybridization of the active sites of *Ap*PDC and *Pf*BAL. To enable *S*-selective synthesis of **4a** starting from benzaldehyde, we chose the previously designed *S*-selective *Ap*PDC-E469G[13] as a starting point for rational enzyme design. In the first step, relevant positions for the binding of benzaldehyde in the donor-binding site were identified by modeling the hydroxybenzyl-ThDP (HBThDP) intermediate into the superimposed crystal structures of *Ap*PDC and *Pf*BAL.

As a result, two positions were found to mainly contribute to the structure of the respective donor-binding sites: T384 (standard position 388)[24] and W388 (392) in *Ap*PDC as well as G393 (388) and Y397 (392) in *Pf*BAL (Figure 2 A, B). W388 (392) in *Ap*PDC restricts the donor-binding site in the *para*-position of the phenyl ring of HBThDP and is mainly substituted by small amino acids in 43 sequences homologous to *Pf*BAL (alanine (28 %), proline (23 %) and glycine (14 %)). But also the more bulky and aromatic tyrosine was found in 14 % of the 43 *Pf*BAL homologues. T384 (388) limits the space for benzaldehyde in the donor-binding site and is conserved in 92 % of all *Ap*PDC homologues, whereas glycine was found at the respective position in all *Pf*BAL homologues (Figure 2 A, B). Thus, standard position 388 seems to be the pivotal factor that directly influences the chemoselectivity of *Ap*PDC. Short molecular dynamics simulations of *Ap*PDC-E469G/T384G with the HBThDP intermediate supported this assumption (Figure 2 C). As a result, an enlarged donor-binding site with putatively sufficient space for benzaldehyde was obtained. Moreover, residue W388 underwent slight conformational changes, probably because of interactions with the HBThDP intermediate, which additionally opened the donor-binding site. Thus, T384 of *Ap*PDC-E469G was replaced by glycine in order to mimic the *Pf*BAL donor-binding site.

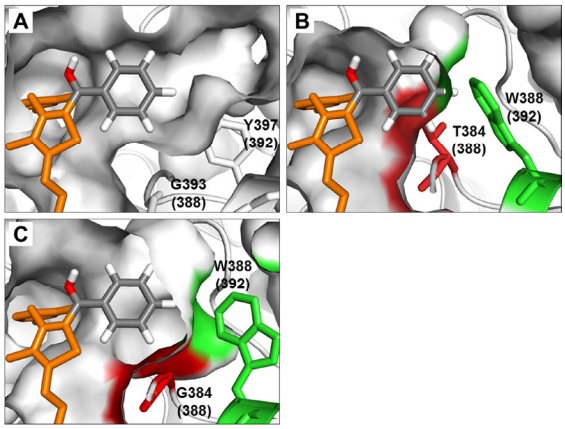


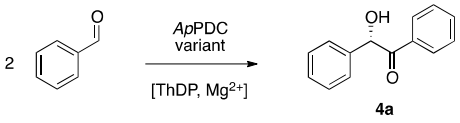
Figure 2. The donor-binding sites of *Ap*PDC-E469G, *Pf*BAL, and *Ap*PDC-E469G/T384G with hydroxybenzyl (gray) bound to the C2 of ThDP (orange). **A:** *Pf*BAL possesses a large donor-binding site, optimal to stabilize benzaldehyde as donor substrate. **B:** *Ap*PDC-E469G is not able to properly bind benzaldehyde in the small donor-binding site, which is restricted by T384 (red) and W388 (green). **C:** The donor-binding site could be opened for benzaldehyde in *Ap*PDC-E469G by replacement of T384 by glycine. Standard numbers are given in parenthesis.

Biochemical characterization of the new variant proved that already the single mutation of T384 to glycine altered the chemoselectivity of *Ap*PDC-E469G, which now revealed remarkable benzoin-forming activity. Whereas no significant formation of **4a** in the homocoupling of benzaldehyde was observed with *Ap*PDC-E469G (Table 1, entry 1), double variant E469G/T384G catalyzed the same reaction with high conversion of 52 % (entry 2). In addition, *Ap*PDC-E469G/T384G catalyzed the formation of **2** and **4a** as the main products in the heterocoupling of acetaldehyde and benzaldehyde, whereas **1** and **3** were obtained only in traces (not shown). Thus, the preference for the donor substrate was completely shifted from acetaldehyde in case of *Ap*PDC-E469G to benzaldehyde in *Ap*PDC-E469G/T384G-catalyzed carboligation.

Furthermore, the combination of both a large donor-binding site for benzaldehyde and a suitable *S*-pocket resulted in the first variant for the synthesis of (*S*)-**4a** (*ee* = 59 %, entry 2). In the heterocoupling, *Ap*PDC-E469G/T384G catalyzed highly *S*-selective the formation of **2** (*ee* = 91 %). The moderate *S*-selectivity of *Ap*PDC-E469G/T384G was subsequently further optimized.

Optimization of the *S*-selectivity. An improved *S*-selectivity can be obtained by either two strategies: stabilization of the antiparallel ("*S*-pathway") or destabilization of the parallel acceptor orientation ("*R*-pathway").[27] To suppress the *R*-pathway in *Ap*PDC-E469G/T384G, the active site was examined for residues that potentially stabilize the acceptor benzaldehyde in its parallel orientation. Two amino acid residues exhibiting such probably stabilizing effects were found in close proximity to the modeled parallel-oriented benzaldehyde:

Table 1. Formation of (*S*)-benzoin catalyzed by *Ap*PDC variants[a]



entry	*Ap*PDC variant	*ee* (%)[b]	conversion (%)[c]
1	E469G	n.d.[d]	<1
2	E469G/T384G	59 (*S*)	52
3	E469G/T384G/I468G	66 (*S*)	23
4	E469G/T384G/I468A	87 (*S*)	95
5	E469G/T384G/I468V	76 (*S*)	40
6	E469G/T384G/ I468A/W543F	95 (*S*)	36
7	E469G/T384G/ I468A/W543F	98 (*S*)[e]	26[e]

[a]Reaction conditions: 50 mM TEA buffer, pH 8.0, 2 mM $MgSO_4$ and 0.1 mM ThDP; 1 mg mL^{-1} enzyme; 18 mM benzaldehyde; 20 °C, 6 h.
[b]Determined by chiral-phase HPLC.
[c]Determined by chiral-phase HPLC based on the consumption of benzaldehyde.
[d]Not determined.
[e]Reaction conditions see *a*, but at 15 °C.

I468 (standard position 476) and W543 (559) (Figure 3 A). Both residues do not directly influence the antiparallel orientation of benzaldehyde, but are probably able to stabilize parallel-oriented benzaldehyde by nonpolar interactions or π-stacking. Consequently, both positions were subjected to mutagenesis.

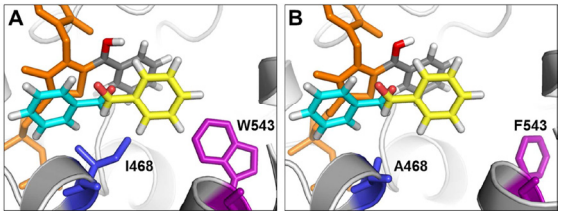


Figure 3. Potential stabilization of the parallel-oriented acceptor benzaldehyde in the active site of *Ap*PDC-variants. **A:** In E469G/T384G parallel-oriented benzaldehyde (yellow) is putatively stabilized by I468 (blue) and W543 (purple), which both do not directly influence its antiparallel orientation (cyan). **B:** In E469G/T384G/I468A/W543F stabilization of parallel-oriented benzaldehyde by A468 (blue) and F543 (purple) is not possible anymore.

I468 was replaced by valine, glycine, and alanine to increase the distance of the respective side chain to parallel-oriented benzaldehyde and thus prevent their positive effects on stabilization of the *R*-pathway (Figure 3 B). In all cases the new variants showed improved *S*-selectivity (Table 1). In the best case, the *ee* of **4a** was improved to 87 % (*S*) accompanied by an intriguingly high conversion of 95 % using triple variant E469G/T384G/I468A (entry 4). This high conversion is rather surprising, because multiple active site mutations to improve selectivity or to broaden substrate specificity frequently result in a drastic decrease of enzymatic activity, as previously demonstrated for other ThDP-dependent enzymes.[27]

To further improve the stereoselectivity of variant E469G/T384G/I468A, a fourth mutation in position 543 (standard position 559) was introduced. W543 is part of the C-terminal α-helix, which covers the entrance to the active site of *Ap*PDC (Figure 3 A). This position, probably relevant for structural stabilization, is conserved in 76 % of the homologues of *Ap*PDC. In close proximity to W543, a cluster of aromatic amino acids was found, which is conserved in *Ap*PDC homologues but not in other ThDP-dependent decarboxylases (for more details see Supporting Information, Table S1). This cluster, which includes W262 (99 % of PDCs), Y290 (91 %), W388 (98 %), F389 (98 %), and W543, probably links the C-terminal helix as well as the enzymes' middle domain (TH3-domain) to the pyrophosphate binding domain (PP-domain) by aromatic interactions.

W543 was subjected to site-saturation mutagenesis using NDT codon degeneracy. Only few variants showed significant carboligation activity in the screening, but all active variants were *S*-selective in the formation of **4a**. Among these variants, E469G/T384G/I468A/W543F revealed the highest *S*-selectivity for the synthesis of **4a** with an *ee* of 95 % and a conversion of 36 % (Table 1, entry 6). Compared to W543, the increased distance between the phenyl ring of the parallel-oriented acceptor benzaldehyde and F543 (6.8 Å compared to prior 3.7 Å) prevents stabilizing interactions (Figure 3 B). Furthermore, the physicochemical property in position 543 was preserved by the substitution of tryptophan by phenylalanine, which might be beneficial to maintain structural stability of the *Ap*PDC variant in the aromatic cluster.

In addition to improved *S*-selectivity in the homocoupling reaction of benzaldehyde, *Ap*PDC-E469G/T384G/I468A/W543F also revealed enhanced stereoselectivity for the synthesis of (*S*)-**2** (*ee* > 99 %) starting from acetaldehyde and benzaldehyde. Again, the potential products **1** and **3** were not observed in this heterocoupling reaction.

To further optimize *S*-selectivity of the *Ap*PDC variants, we investigated the influence of pH, substrate concentration, temperature, and the addition of cosolvents on both, conversion and stereoselectivity. As a result, temperature and the addition of cosolvents were identified to significantly affect the stereoselectivity. Reduced temperature resulted in improved *S*-selectivity. *Ap*PDC-E469G/T384G/I468A/W543F at 30 °C yielded (*S*)-**4a** with 93 % *ee* and a conversion of 44 %, whereas at 15 °C (*S*)-**4a** was formed with 98 % *ee*, but accompanied by a reduced conversion of 26 % (Table 1, entry 7). In the presence of dimethyl sulfoxide (DMSO) and methyl *tert*-butyl ether (MTBE) as cosolvents, both common additives to enhance the solubility of aromatic substrates, *S*-selectivity was significantly reduced. *Ap*PDC-E469G/T384G/I468A gave 75 % and 80 % *ee* for the formation of (*S*)-**4a** in the presence of DMSO (20 vol%) and MTBE (5 vol%), respectively. In addition, conversion dropped from 95 % to 28 % and 42 %, respectively. These results are consistent

with recently published data of Gerhards et al.,[31] who comprehensively investigated the influence of organic solvents on the stereoselectivity of ThDP-dependent enzymes.

Variation of pH and substrate concentration influenced the catalytic activity of both variants but not the stereoselectivity. Best conversion was obtained at pH 8.0 (data not shown). High substrate concentrations (> 30 mM) led to fast precipitation of the enzyme. Thus, synthesis of **4a** and derivatives thereof (see below) was performed without addition of cosolvents at pH 8.0 and 15 °C using 18 mM substrates.

Chemoenzymatic synthesis of (*S*)-benzoin derivatives. A substrate screening with substituted benzaldehyde derivatives was performed using the optimized reaction conditions for the synthesis of (*S*)-**4a**. *Meta*-substituted benzaldehyde derivatives turned out to be the best substrates in the homocoupling reaction, yielding highest *S*-selectivities and conversions. This was observed for the syntheses of **4b–4l** using *Ap*PDC variants E469G/T384G/I468A and E469G/T384G/I468A/W543F.

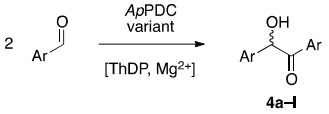
Except for 3-fluorobenzaldehyde **4e**, all *meta*-substituted benzaldehydes (entry 5–9) were transformed with higher *S*-selectivity compared to the standard reaction with benzaldehyde (entry 1). Thereby, reactions with variant E469G/T384G/I468A/W543F showed again higher *S*-selectivity and lower conversion compared to E469G/T384G/I468A. Remarkably, variant E469G/T384G/I468A/W543F catalyzed the enantiopure formation of **4f–i** with *ee* > 99 % (*S*) (entry 6–9). With *ortho*- and *para*-substituted benzaldehydes *R*-selectivity and low conversions were observed (except for 4-fluorobenzaldehyde) (Table 2).

These results confirm our findings from recent 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) studies.[15, 27] However, the reasons for the preference of *meta*-substituted benzaldehydes in *S*-selective carboligation reactions and the switch in stereoselectivity in the case of *ortho*- and *para*-substituted benzaldehydes, respectively, are still not understood. A combination of steric and electronic interactions of substrate and *S*-pocket might influence the stereoselectivity. The elucidation of this phenomenon is now subject of further investigations.

CONCLUSIONS

The rational design of highly active and selective enzymes for nonphysiological reactions is still a great challenge for biocatalytic applications. The exceptional catalytic potential of ThDP-dependent enzymes in asymmetric C–C bond formation has been explored for almost hundred years and was recently outlined to be mature for biotechnological applicability.[4] Nonetheless, the long-standing problem of (*S*)-benzoin synthesis starting from benzaldehyde as a sole substrate was not yet solved. Here, we demonstrate that the detailed understanding of principles guiding chemo- and stereoselectivity in ThDP-dependent enzymes enabled us to rationally tailor *Ap*PDC for (*S*)-benzoin synthesis. Therefore, we designed a hybrid substrate-binding site combining the large *S*-pocket of *Ap*PDC-E469G with the large donor-binding site of *Pf*BAL.

Table 2. Carboligation potential of *Ap*PDC variants[a]



			E469G/T384G/I468A		E469G/T384G/I468A/W543F	
entry	Ar	product	*ee* (%)[b]	conv. (%)[c] (yield (%)[d])	*ee* (%)[b]	conv. (%)[c] (yield (%)[d])
1	C_6H_5	4a	89 (*S*)	92 (85)	98 (*S*)	26 (66)
2	2-FC_6H_4	4b	58 (*R*)	36	21 (*R*)	<5
3	2-ClC_6H_4	4c	75 (*R*)	<5	n.d.[e]	n.d.[e]
4	2-$MeOC_6H_4$	4d	n.d.[e]	n.c.[f]	n.d.[e]	n.c.[f]
5	3-FC_6H_4	4e	87 (*S*)	80 (82)	93 (*S*)	36
6	3-ClC_6H_4	4f	91 (*S*)	97	>99 (*S*)	48 (72)
7	3-BrC_6H_4	4g	95 (*S*)	85 (84)	>99 (*S*)	30
8	3-IC_6H_4	4h	96 (*S*)	30 (70)	>99 (*S*)	11
9	3-$MeOC_6H_4$	4i	98 (*S*)	93	>99 (*S*)	58 (61)
10	4-FC_6H_4	4j	64 (*S*)	53	85 (*S*)	<5
11	4-ClC_6H_4	4k	>99 (*R*)	10	n.d.[e]	n.c.[f]
12	4-$MeOC_6H_4$	4l	n.d.[e]	<5	n.d.[e]	n.c.[f]

[a–c]See Table 1. Reaction temperature: 15 °C.
[d]Isolated yields after preparative synthesis and complete conversion (reaction conditions see Material and Methods).
[e]Not determined.
[f]No conversion.

This allows benzaldehyde to bind as donor and acceptor substrate in an antiparallel orientation. By transferring the pivotal factor of benzoin formation from *Pf*BAL to *Ap*PDC-E469G (by mutation T384G), we shifted its chemoselectivity resulting in the first *S*-selective benzoin formation (*ee* = 59 %). In addition, also the synthesis of (*S*)-2-HPP (*ee* = 91 %) in the mixed carboligation of acetaldehyde and benzaldehyde is now possible with this *Ap*PDC-variant. The formation of acetoin and PAC, the typical products of wild type *Ap*PDC, was not observed.

Furthermore, we improved the *S*-selectivity of *Ap*PDC-E469G/T384G by preventing the parallel orientation of the acceptor substrate. Two highly selective and active variants were obtained: *Ap*PDC-E469G/T384G/I468A and E469G/T384G/I468A/W453F. Thus, we successfully complemented the platform of α-hydroxy ketones (Scheme 1) with respect to (*S*)-benzoin formation (*ee* = 98 %). This provides access to various enantiopure *meta*-substituted (*S*)-benzoin derivatives starting from commercially available benzaldehyde derivatives.

These results highlight the robustness of ThDP-dependent enzymes with respect to active site mutations as well as their tremendous catalytic potential in asymmetric carboligation reactions. Moreover, this hybridization approach might open the field of combinatorial assembly of selectivity determining modules from different ThDP-dependent enzymes and offer a new perspective on thiamine catalysis with respect to the design of variants with novel catalytic activities.

ACKNOWLEDGMENT

This work is dedicated to Ayhan S. Demir (†). The authors thank Doris Hahn, Ursula Mackfeld, and Heike Offermann for skillful technical support.

ABBREVIATIONS

2-HPP, 2-hydroxypropiophenone; BAL, benzaldehyde lyase; BFD, benzoylformate decarboxylase; CDH, cyclohexane-1,2-dione hydrolase; DMSO, dimethyl sulfoxide; HBThDP, hydroxybenzyl-ThDP; *ee*, enantiomeric excess; MTBE, methyl *tert*-butyl ether; PAC, phenylacetylcarbinol; PDC, pyruvate decarboxylase; SucA, E1 component of the α-ketoglutarate dehydrogenase subunit from *Escherichia coli* K12; TEA, triethanolamine; ThDP, thiamine diphosphate.

REFERENCES

(1) Hoyos, P.; Sinisterra, J. V.; Molinari, F.; Alcantara, A. R.; de Maria, P. D. *Acc. Chem. Res.* **2010**, *43*, 288–299.

(2) Müller, M.; Gocke, D.; Pohl, M. *FEBS J.* **2009**, *276*, 2894–2904.

(3) Brovetto, M.; Gamenara, D.; Saenz Méndez, P.; Seoane, G. A. *Chem. Rev.* **2011**, *111*, 4346–4403.

(4) Müller, M.; Sprenger, G. A.; Pohl, M. *Curr. Opin. Chem. Biol.* **2013**, *17*, 261–270.

(5) Jordan, F. *Nat. Prod. Rep.* **2003**, *20*, 184–201.

(6) Frank, R. A. W.; Leeper, F. J.; Luisi, B. F. *Cell. Mol. Life Sci.* **2007**, *64*, 892–905.

(7) Kluger, R.; Tittmann, K. *Chem. Rev.* **2008**, *108*, 1797–1833.

(8) Hailes, H. C.; Rother, D.; Müller, M.; Westphal, R.; Ward, J. M.; Pleiss, J.; Vogel, C.; Pohl, M. *FEBS J.* **2013**, DOI: 10.1111/febs.12496.

(9) Duggleby, R. G. *Acc. Chem. Res.* **2006**, *39*, 550–557.

(10) Widmann, M.; Radloff, R.; Pleiss, J. *BMC Biochem.* **2010**, *11*, 9.

(11) Knoll, M.; Müller, M.; Pleiss, J.; Pohl, M. *ChemBioChem* **2006**, *7*, 1928–1934.

(12) Gocke D.; Walter, L.; Gauchenova, E.; Kolter, G.; Knoll, M.; Berthold, C. L.; Schneider, G.; Pleiss, J.; Müller, M.; Pohl, M. *ChemBioChem* **2008**, *9*, 406–412.

(13) Rother, D.; Kolter, G.; Gerhards, T.; Berthold, C. L.; Gauchenova, E.; Knoll, M.; Pleiss, J.; Müller, M.; Pohl, M. *ChemCatChem* **2011**, *3*, 1587–1596.

(14) Westphal, R.; Waltzer, S.; Mackfeld, U.; Widmann, M.; Pleiss, J.; Beigi, M.; Müller, M.; Rother, D.; Pohl, M. *Chem. Commun.* **2013**, *49*, 2061–2063.

(15) Westphal, R.; Jansen, S.; Vogel, C.; Pleiss, J.; Müller, M.; Rother, D.; Pohl, M. *ChemCatChem* **2013**,DOI: 10.1002/cctc.201300690.

(16) Demir, A. S.; Pohl, M.; Janzen, E.; Müller, M. *J. Chem. Soc., Perkin Trans. 1* **2001**, 633–635.

(17) Hoyos, P.; Fernandez, M.; Sinisterra, J. V.; Alcantara, A. R. *J. Org. Chem.* **2006**, *71*, 7632–7637.

(18) Demir, A. S.; Peruze, A.; Umut, D.; Umut, E. *J. Mol. Catal. B: Enzym.* **2008**, *55*, 164–168.

(19) Fragnelli, M. C.; Hoyos, P.; Romano, D.; Gandolfi, R.; Alcantara, A. R.; Molinari, F. *Tetrahedron* **2012**, *68*, 523–528.

(20) Demir, A. S.; Dünnwald, T.; Iding, H.; Pohl, M.; Müller, M. *Tetrahedron: Asymmetry* **1999**, *10*, 4769–4774.

(21) Demir, A. S.; Sesenoglu, Ö.; Eren, E.; Hosrik, B.; Pohl, M.; Janzen, E.; Kolter, D.; Feldmann, R.; Dünkelmann, P.; Müller, M. *Adv. Synth. Catal.* **2002**, *344*, 96–103.

(22) Dünkelmann, P.; Kolter-Jung, D.; Nitsche, A.; Demir, A. S.; Siegert, P.; Lingen, B.; Baumann, M.; Pohl, M.; Müller, M. *J. Am. Chem. Soc.* **2002**, *124*, 12084–12085.

(23) Enders, D.; Kallfass, U. *Angew. Chem. Int. Ed.* **2002**, *41*, 1743–1745.

(24) Vogel, C.; Widmann, M.; Pohl, M.; Pleiss, J. *BMC Biochem.* **2012**, *13*, 24.

(25) DeLano, W. L. *The PyMOL Molecular Graphics System* **2002**, http://www.pymol.org, San Carlo, CA, USA.

(26) Krieger, E.; Koraimann, G.; Vriend, G. *Proteins* **2002**, *47*, 393–402.

(27) Westphal, R.; Hahn, D.; Mackfeld, U.; Waltzer, S.; Beigi, M.; Widmann, M.; Vogel, C.; Pleiss, J.; Müller, M.; Rother, D.; Pohl, M. *ChemCatChem* **2013**, DOI: 10.1002/cctc.201300318.

(28) Sambrook, J.; Russell, D. W. *Molecular cloning: A Laboratory Manual 1st ed.* **2001**, pp. 1.116–1.118.

(29) Breuer, M.; Pohl, M.; Hauer, B.; Lingen, B. *Anal. Bioanal. Chem.* **2002**, *374*, 451–457.

(30) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248–254.

(31) Gerhards, T.; Mackfeld, U.; Bocola, M.; von Lieres, E.; Wiechert, W.; Pohl, M.; Rother, D. *Adv. Synth. Catal.* **2012**, *354*, 2805–2820.

(32) Gocke, D.; Graf, T.; Brosi, H.; Frindi-Wosch, I.; Walter, L.; Müller, M.; Pohl, M. *J. Mol. Catal. B: Enzym.* **2009**, *61*, 30–35.

(33) Krawczyk, E.; Koprowski, M.; Skowronska, A.; Luczak, J. *Tetrahedron: Asymmetry* **2004**, *15*, 2599–2602.

Supporting Information

Table of contents

1. Conservation of aromatic amino acids in ThDP-dependent decarboxylases

A conserved cluster of aromatic residues was found in close spatial proximity to W543 in *Ap*PDC. Conservation analysis with ThDP-dependent decarboxylases (family classification according to Duggleby 2006 and Widmann et al. 2010),[1, 2] based on an up-to-date version of the ThDP-dependent Enzyme Engineering Database (TEED) (unpublished data), revealed that this cluster is exclusively conserved in sequences homologous to *Ap*PDC (Table S1), which form the second homologous family of pyruvate decarboxylases in the TEED.[2]

Table S1. Conservation of aromatic amino acids in ThDP-dependent decarboxylases at positions corresponding to the aromatic cluster observed in *Ap*PDC.

	Standard position[a] (absolute position in *Ap*PDC)				
	264 (262)	292 (290)	392 (388)	393 (389)	559 (543)
HFAM[b]	Conserved aromatic amino acids (frequency in %)				
IPDC	-	-	-	F (73), Y (19), W (1)	-
PhePDC	-	-	-	F (89), Y (4)	-
PDC[c]	-	F (78), Y (2)	-	F (88), Y (2)	-
PDC[d]	W (99)	Y (91), F (2)	W (98)	F (98)	W (76), F (4)
KdcA	-	-	-	F (51), Y (8)	-
MenD	W (63)	-	-	-	-
AHAS	-	-	-	-	W (78), F (6)
AHAS	Y (43), F (14)	-	Y (50)	-	-
AHAS	-	-	F (81)	-	-
AHAS	-	W (61)	-	-	-
SAAT	-	F (72), W (16), Y (3)	-	-	F (57), W (10)
CDH	-	W (100)	-	-	-
THcHDOH	-	Y (100)	-	-	-

[a]Standard numbers according to the standard numbering scheme for ThDP-dependent decarboxylases.[3]
[b]Names of homologous families (HFAM): IPDC, indolepyruvate decarboxylases; PhePDC, phenylpyruvate decarboxylases; PDC, pyruvate decarboxylases; AHAS, acetohydroxyacid synthases; SAAT: sulfoacetaldehyde acetyltransferase; CDH, cyclohexane-1,2-dione hydrolase; THcHDOH: 3D-(3,5/4)-trihydroxycyclohexane-1,2-dione hydrolase. Only those homologous families with at least ten incorporated sequences and the CDH family were analyzed.
[c]Sequences homologous to the substrate inducible PDC from *Saccharomyces cerevisiae*.
[d]Sequences homologous to the *Ap*PDC and the PDC from *Zymomonas mobilis*.

2. Generation of *Ap*PDC variants

The plasmid pET22b (Novagen) containing the gene encoding for *Ap*PDC-E469G[4] was used as template for the preparation of mutant *Ap*PDC-E469G/T384G, using the forward (fw) primer 5'-CGCTGGTGGCAGAAggCGGCGATTCATGG-3' and the reverse primer 5'-CCATGAATCGCCGccTTCTGCCACCAGCG-3' (mutated codons are underlined, with lower-case letters indication base changes). Based on *Ap*PDC-E469G/T384G the mutants *Ap*PDC-E469G/T384G/I468G, *Ap*PDC-E469G/T384G/I468A, and *Ap*PDC-E469G/T384G/I468V were prepared. The following fw and rv primers were used for mutagenesis:

*Ap*PDC-E469G/T384G/I468G: 5'-CGTGGCTATGTCggcGGCATCGCCATTC-3' (fw)

5'-GAATGGCGATGCCgccGACATAGCCACG-3' (rv)

*Ap*PDC-E469G/T384G/I468A: 5'-CCGTGGCTATGTCgccGGCATCGCCATTC-3' (fw)

5'-GAATGGCGATGCCggcGACATAGCCACGG-3' (rv)

*Ap*PDC-E469G/T384G/I468V: 5'-CAACCGTGGCTATGTCgTgGGCATC GCCATTCATG-3' (fw)

5'-CATGAATGGCGATGCCcAcGACATA GCCACGGTTG-3' (rv).

For the generation of the mutant library for amino acid position W543, *Ap*PDC-E469G/T384G/I468A was used as template. The fw primer 5'-GGATATGCTGGTTCAAndtGGCCGCAAGGTTGCC-3' and the rv primer 5'-GGCAACCTTGCGGCCahnTTGAACCAGCATATCC-3' contained NDT degeneracy (where N = any nucleotide, D = A, G, or T).

3. References

[1] Duggleby, R. G. *Acc. Chem. Res.* **2006**, *39*, 550–557.

[2] Widmann, M.; Radloff, R.; Pleiss, J. *BMC Biochem.* **2010**, *11*, 9.

[3] Vogel, C.; Widmann, M.; Pohl, M.; Pleiss, J. *BMC Biochem.* **2012**, *13*, 24.

[4] Rother, D.; Kolter, G.; Gerhards, T.; Berthold, C. L.; Gauchenova, E.; Knoll, M.; Pleiss, J.; Müller, M.; Pohl, M. *ChemCatChem* **2011**, *3*, 1587–1596.

3 GENERAL DISCUSSION

Hereinafter, all publications deduced from this thesis are comparatively discussed including currently unpublished data. The review about "Engineering stereoselectivity of ThDP-dependent enzymes" (Hailes *et al*., 2013) as well as the book article "Synthetic enzyme cascades for valuable diols and amino alcohols – smart composition and optimisation strategies" (Sehl *et al*., 2013) are not part of the following discussion.

3.1 Overview of publications

PUBLICATION I reports on the application of acetohydroxyacid synthase (AHAS) from *Escherichia coli* in an enzymatic one-pot two-step reaction including a novel "recycling mode" for the synthesis of stereochemically pure nor(pseudo)ephedrine starting from inexpensive pyruvate and benzaldehyde.

PUBLICATION II describes the rational design of *S*-selective variants of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase (MenD) from *E. coli* (*Ec*MenD) for the synthesis of novel functionalized chiral α-hydroxy ketones. Thereby, a broad substrate range of diversely substituted benzaldehyde derivatives was tested in the carboligation with α-ketoglutarate (α-KG) catalyzed by three different *Ec*MenD variants.

PUBLICATION III is based on Publication II and describes the engineering strategy of *S*-selective *Ec*MenD variants as well as the influence of steric properties and stabilization of the benzaldehyde acceptor in the *S*-pocket on *S*-selectivity. Furthermore, this publication reports on a new strategy to improve the usually limited *S*-selectivity of ThDP-dependent enzymes by selective destabilization of substrate annealing in the region important for *R*-selective coupling.

PUBLICATION IV reports on the carboligation potential of MenD from *Bacillus subtilis* (*Bs*MenD) and engineered variants thereof for the enantiocomplementary asymmetric synthesis of functionalized α-hydroxy ketones. Moreover, the impact of a second-shell residue on stereoselectivity is described, which explains the significantly higher *S*-selectivity of *Bs*MenD variants relative to *Ec*MenD variants.

PUBLICATION V describes a rational hybridization approach to solve the long-standing problem of enzymatic asymmetric (*S*)-benzoin synthesis by a combination of selectivity-determining features of a variant of pyruvate decarboxylase (PDC) and benzaldehyde lyase (BAL). Furthermore, the new strategy to improve *S*-selectivity (Publication III) could be used to tailor highly *S*-selective PDC variants for the synthesis of various benzoin derivatives.

3.2 New enzymes for the enzyme toolbox

As was already addressed in the introduction, time-to-market pressure is a key issue for the choice of a process strategy in the chemical and pharmaceutical industry (Wenda *et al*., 2011), which still hampers the use of biocatalysis as sustainable alternative to classical chemical approaches. Enzyme toolboxes may help biocatalysis to become more attractive for industrial purposes (chapter 1.5). Thereby, it is of critical importance to continuously expand the toolbox of biocatalysts with new and improved enzyme variants to get access to a broad range of products as well as to provide access to *R*- and *S*-selective enzymes to enable enantiocomplementary synthesis (Meyer *et al*., 2013).

The main goal of this thesis was the expansion of the thiamine diphosphate (ThDP)-dependent enzyme toolbox for the synthesis of α-hydroxy ketones as chiral building blocks, with special focus on hardly accessible (*S*)-α-hydroxy ketones. The existing toolbox, established in the group *Biocatalysis and Biosensors* (IBG-1, Forschungszentrum Jülich GmbH), comprised 16 wild-type enzymes and about 80 active variants at the beginning of this thesis (Gocke, 2007; Kolter, 2010). In the present work, this toolbox was extended by characterization and mutagenesis of new wild-type enzymes as well as by mutagenesis of already well-characterized toolbox enzymes. Table 3 gives an overview of all enzymes used in this thesis to study carboligase activities. Except for AHAS and MenD, all other enzymes mentioned in Table 3 were already part of the ThDP-dependent enzyme toolbox and well characterized with respect to structure-function relationships (Gocke, 2007; Gocke *et al*., 2007; Rother *et al*., 2011). To allow the comparison of sterical properties of the active sites with other ThDP-dependent enzymes and to rationally design appropriate variants in order to investigate the principles of chemo- and stereoselectivity, the work started with an in-depth structural analysis of AHAS and MenD.

Table 3: Overview of recombinant enzymes used in the thesis.

Enzyme / EC number	Organism	Plasmid[1]	Reference
Acetohydroxyacid synthase I[2] 2.2.1.6	*Escherichia coli*	pET28c-BN^{Amp}	Engel *et al*., 2004a
Acetohydroxyacid synthase II[2] 2.2.1.6	*Escherichia coli*	pQEV-gM^{Amp}	Engel *et al*., 2004b
Benzaldehyde lyase 4.1.2.38	*Pseudomonas putida*	pET28a[3]	Janzen *et al*., 2006
Pyruvate decarboxylase 4.1.1.1	*Acetobacter pasteurianus*	pET22b	Rother *et al*., 2011
SEPHCHC synthase (MenD)[4] 2.2.1.9	*Escherichia coli*	pET19b	Kurutsch *et al*., 2009
SEPHCHC synthase (MenD)[5] 2.2.1.9	*Bacillus subtilis*	pET19b	Publication IV

[1]All pET vectors were expressed in *E. coli* BL21 (DE3) (Novagen®, Merck Millipore, Darmstadt, Germany), whereas pQEV-gM^{Amp} was expressed in *E. coli* SG13009 (Qiagen, Hilden, Germany).
[2]Kindly provided by Prof. Dr. K. Tittmann (Georg-August-University Göttingen, Germany).
[3]The original gene was cloned from pKK233-2 into pET28a by D. Hahn (IBG-1, Forschungszentrum Jülich GmbH, Germany).
[4]Kindly provided by Prof. Dr. M. Müller (Albert-Ludwigs-University Freiburg, Germany).
[5]Purchased from GeneArt® Gene Synthesis (life technologies™, Carlsbad, USA) as synthetic gene codon-optimized for *E. coli*.

3.3 Investigation of structure-function relationships

3.3.1 Homology models of *Ec*AHAS

The three-dimensional (3D) structure of a protein is the main precondition for the elucidation of structure-function relationships. In contrast to MenD, the crystal structures of the catalytic subunits of AHAS-I and II from *E. coli* (*Ec*AHAS) are still not available despite intensive efforts, since no appropriate crystallization conditions could be found yet. A reason might be the oligomeric heterogeneity of *Ec*AHAS due to the additional regulatory subunits (chapter 1.7.3). Up to ten different oligomers could be observed during native polyacrylamide gel electrophoresis (PAGE) (personal communication with Prof. Dr. K. Tittmann, Georg-August-University Göttingen, Germany). Even crystallization of the separated catalytic subunits of *Ec*AHAS-I and II (for sequences see Appendix 5.1) during this thesis* was yet not successful. Only so-called quasicrystals (irregular crystals) could be obtained, which cannot be analyzed and thus have to be optimized further. Up to now, no further improvement could be achieved. However, there are 3D structures available from homologous AHAS, such as the catalytic subunit of AHAS from *Saccharomyces cerevisiae* (*Sc*AHAS) (Pang *et al*., 2002). The

*Dr. J. Kulig (IBG-1, Forschungszentrum Jülich GmbH, Germany) performed crystallization studies in York under supervision of Dr. G. Grogan (Department of Chemistry, University of York, UK).

high conservation of the active sites of *Sc*AHAS and *Ec*AHAS enabled the generation of a homology model for *Ec*AHAS-II (Engel *et al*., 2004b). In this thesis, the YASARA Structure software (Krieger *et al*., 2002) was used to generate homology models of the catalytic subunits of *Ec*AHAS-I and *Ec*AHAS-II based on the structure of *Sc*AHAS (Protein Data Bank (pdb) code: 1N0H). The homology models obtained for *Ec*AHAS-I and *Ec*AHAS-II revealed the typical active site architecture of ThDP-dependent enzymes (chapter 1.7.2), including the conserved "V-conformation" of ThDP (Figure 22). Moreover, the *Ec*AHAS-II homology model was consistent with that reported previously for *Ec*AHAS-II (Engel *et al*., 2004b).

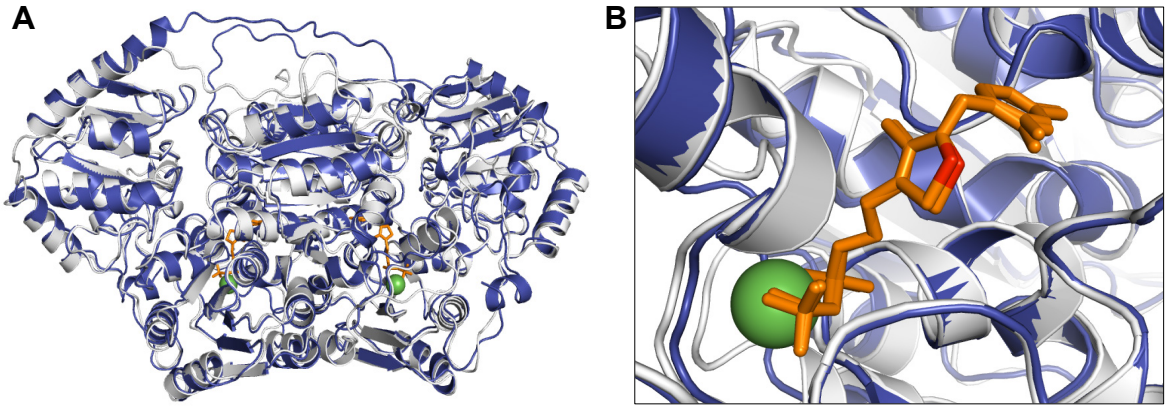


Figure 22: The dimeric structures and active sites of *Ec*AHAS-I and *Ec*AHAS-II shown as PyMOL models. The superimposed homology models of *Ec*AHAS-I (white) and *Ec*AHAS-II (blue) revealed the typically conserved structure of ThDP-dependent enzymes (**A**) with ThDP (orange) in its conserved "V-conformation" anchored *via* a magnesium cation (green) in the active site (**B**). The homology models were generated based on *Sc*AHAS (pdb: 1N0H, Pang *et al*., 2002) as template.

3.3.2 Structure-function relationships of *Ec*AHAS and MenD

AHAS and MenD belong to the group of transferases (EC 2) and catalyze a physiological C–C bond formation. By contrast, most other ThDP-dependent enzymes of the decarboxylase (DC) superfamily are lyases (EC 4), which catalyze the decarboxylation of α-keto acids as the physiological activity (chapter 1.7.3) (Hailes *et al*., 2013). Moreover, both enzymes are assumed to require an *in situ* decarboxylation of an α-keto acid donor prior to carboligation and do not accept the corresponding aldehydes, which is different compared to most other decarboxylases (Engel *et al*., 2003; Engel *et al*., 2004b; Kurutsch *et al*., 2009).

To deduce structure-function relationships for *Ec*AHAS and MenD and to complement previous studies (Gocke, 2007), the generated homology models of *Ec*AHAS and the published structures of *Ec*MenD and *Bs*MenD (Dawson *et al*., 2008, 2010) were superimposed with the structures of benzaldehyde lyase from *Pseudomonas fluorescens* (*Pf*BAL) (Mosbacher *et al*., 2005), benzoylformate decarboxylase from *Pseudomonas putida*

(*Pp*BFD) (Hasson *et al*., 1998), pyruvate decarboxylase from *Acetobacter pasteurianus* (*Ap*PDC) (Rother *et al*., 2011), and branched-chain keto acid decarboxylase from *Lactococcus lactis* (*Ll*KdcA) (Berthold *et al*., 2007). Table 4 summarizes important amino acids lining specific regions of the active site in the respective enzymes. These positions were additionally assigned to standard numbers according to the standard numbering scheme for ThDP-dependent decarboxylases (Vogel *et al*., 2012). Standard numbers highlight the structural equivalence of amino acid residues in different enzymes and thus, facilitate the comparative analysis of these positions. In this thesis, the standard numbers are written in parenthesis behind the absolute amino acid position, *e.g*. E469G (477).

Table 4: Amino acid residues lining specific regions of the active site of selected ThDP-dependent enzymes. The given standard numbers, according to the recently developed standard numbering scheme for ThDP-dependent decarboxylases (Vogel *et al*., 2012), highlight the structural equivalence of respective amino acid residues. Residues in scare brackets are only partially involved in lining the specific site. The respective areas for these positions are shown in Figure 23. The assignment of residues of *Pf*BAL, *Pp*BFD, *Ap*PDC, and *Ll*KdcA was previously reported (Gocke, 2007), but slightly modified and complemented in this thesis.

Standard number	*Pf*BAL (pdb: 2AG0)	*Pp*BFD (pdb: 1BFD)	*Ap*PDC (pdb: 2VBI)	*Ll*KdcA (pdb: 2VBF)	*Ec*AHAS-I[1]	*Ec*AHAS-II[1,2]	*Ec*MenD[2] (pdb: 2JLC)	*Bs*MenD (pdb: 2X7J)
residues lining the *S*-pocket (area S1 in Figure 23)								
26	H26	P24	A25	P24	P35	P23	P30	P29
27	G27	G25	G26	G25	G36	G24	G31	G30
28	A28	S26	D27	D26	G37	G25	S32	S31
474	[W478]	[Y458]	[Y466]	[Y459]	[L474]	[L458]	[G472]	[G487]
477[3]	T481	L461	E469	E462	V477	V461	F475	F490
480	[F484]	F464	[I472]	I465	Q480	W464	L478	L493
entrance to the *S*-pocket (area S2 in Figure 23)								
476[3]	A480	A460	I468	V461	L476	M460	I474	I489
important residues lining the donor-binding site (area D in Figure 23)								
114	L112	-	-	-	F122	F109	N117	-
388	G393	-	T384	Q377	V391	V375	N390	N404
389	-	T377	-	[G378]	[G392]	[G376]	[S391]	[S405]
392	Y397	T380	W388	F381	Q395	Q379	-	-
393	-	-	-	-	[M396]	[M380]	R395	R409
408	-	F397	-	-	-	-	-	-
410	-	-	-	-	-	-	R413	R428
413	-	-	-	-	-	-	S416	N431

Table 4: continued.

Standard number	*Pf*BAL (pdb: 2AG0)	*Pp*BFD (pdb: 1BFD)	*Ap*PDC (pdb: 2VBI)	*Ll*KdcA (pdb: 2VBF)	*Ec*AHAS-I[1]	*Ec*AHAS-II[1,2]	*Ec*MenD[2] (pdb: 2JLC)	*Bs*MenD (pdb: 2X7J)
residue stabilizing the "V-conformation" of ThDP								
415	M421	L403	I411	I404	M419	M403	I418	I433
potential proton relay systems								
29	H29[4]	-	-	-	-	-	-	-
73	T73[4]	H70	T72[4]	T71[4]	C83[4]	T70[4]	T78[4]	T77[4]
114	-	-	H113	H112	-	-	-	-
115	Q113	-	H114	H113	Q123	Q110	Q118	Q117
α-helix covering the entrance to the substrate channel (area C in Figure 23)								
	aa[5] 550–555	no α-helix	aa 538–555	aa 532–547	no α-helix	no α-helix	no α-helix	no α-helix

[1]Based on homology modeling using *Sc*AHAS (pdb: 1N0H, Pang *et al*., 2002) as template.
[2]The assignment of amino acid residues was partially reported (Hailes *et al*., 2013).
[3]Main residues defining the *S*-pocket.
[4]Proton transfer may require a water molecule.
[5]aa: amino acid

Active site maps

In order to facilitate the comparison of active sites, so-called active site maps were derived from the structural investigations for *Ec*AHAS and both MenDs. The schematic models allow the easy visualization and comparison of hotspots with respect to chemo- and stereoselectivity (Pohl *et al.*, 2009). Figure 23 demonstrates the different active site architectures of all enzymes relevant for this thesis.

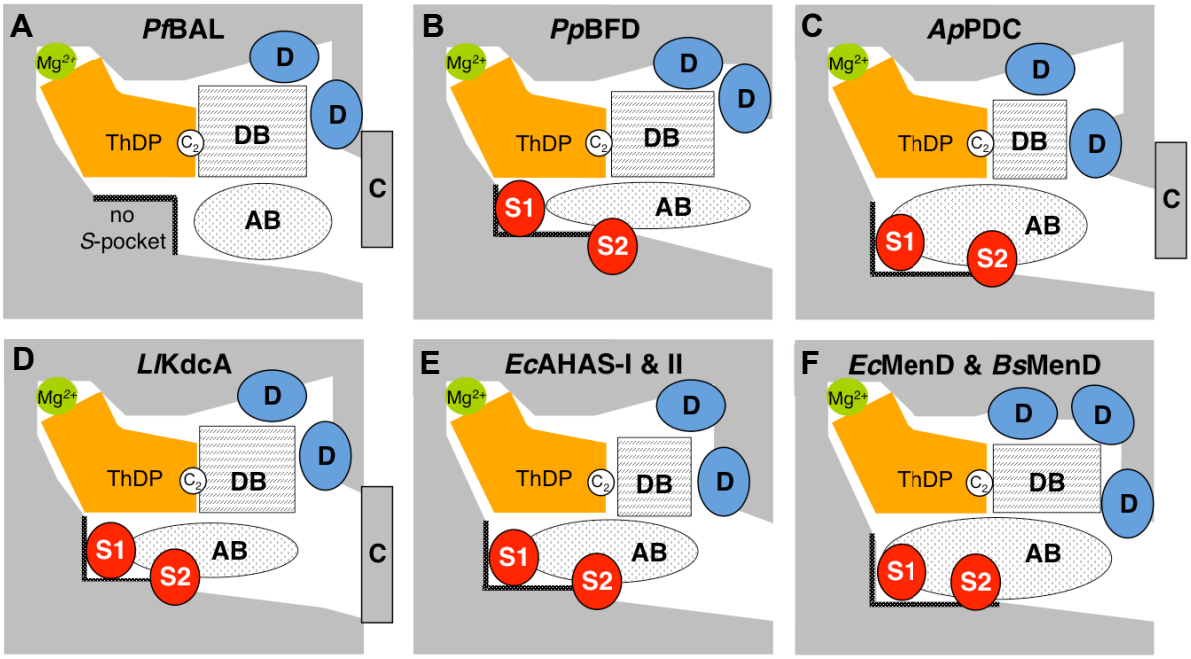


Figure 23: Active site maps of *Pf*BAL (**A**), *Pp*BFD (**B**), *Ap*PDC (**C**), *Ll*KdcA (**D**), *Ec*AHAS (**E**), and MenD (**F**) shown as schematic representation (according to Pohl *et al.*, 2009), which demonstrate the current mechanistic understanding of carboligation. AB: acceptor-binding site; C_2: C_2 atom of the thiazolium ring of ThDP, which represents the binding site for the donor substrate; C: C-terminal α-helix covering the entrance to the active site; D: residues defining the binding site for the donor; DB: donor-binding site; S1: residues lining the *S*-pocket; S2: residues defining the entrance of the *S*-pocket. The bold black line represents the position of the protein backbone, which mainly defines the potential *S*-pocket size.

Based on the gained structural information (Table 4 and Figure 23), available biochemical data, and the comparison to previously characterized ThDP-dependent enzymes, structure-function relationships were deduced for *Ec*AHAS and both MenDs.

Substrate channel

Like *Pp*BFD (Knoll *et al.*, 2006), the channel entrances of *Ec*AHAS, *Ec*MenD and *Bs*MenD are not covered by a C-terminal α-helix and thus clearly visible on the surface of the respective structures. The size of the substrate channel of *Ec*AHAS is comparable to that of *Ap*PDC (Gocke, 2007). The size of the substrate channels of *Ec*MenD and *Bs*MenD is between that of *Ap*PDC and *Ll*KdcA.

Proton relay systems

Proton relay systems are amino acid residues that function as proton donor and acceptor and thus, mediate the proton transfers during the catalytic cycle (chapter 1.7.1) (Frank *et al*., 2007). The respective amino residues, in most so far reported cases histidine residues, are found in appropriate distance to the hydroxyl group of the ThDP-bound donor.

In contrast to *Pf*BAL, *Pp*BFD, *Ap*PDC, and *Ll*KdcA (Pohl *et al*., 2009), *Ec*AHAS and both MenDs have no histidine residue in the active site that could participate in the proton transfer (Table 4). Instead, a highly conserved glutamine is found in standard position 115. As was concluded for Q202 (115) in *Sc*AHAS, the pKa of the glutamine side chain is too high to be involved in the proton relay system (Jaña *et al*., 2010). However, based on mutagenesis studies with *Ec*AHAS-II, Vyazmensky *et al*. (Vyazmensky *et al*., 2011) suggested that the glutamine residue in standard position 115 might simultaneously donate a proton to the carbonyl group of the substrate and accept a proton from the aminopyrimidine N4' (chapter 1.7.1), which would facilitate proton transfer reactions. The crucial role of this glutamine residue for the catalytic cycle was also shown for *Ec*MenD, where substitution to glutamate drastically diminished the activity of the catalyst (Fang *et al*., 2011).

Another putative proton relay system, which may require an additional water molecule, is found at standard position 73 (Hailes *et al*., 2013). A threonine occupies this position in *Ec*AHAS-II and both MenDs, and can be found in *Pf*BAL, *Ap*PDC, and *Ll*KdcA as well. *Ec*AHAS-I has a cysteine at this position.

"V-conformation"

The "V-conformation" of ThDP is conserved in *Ec*AHAS (Figure 22) and both MenDs (Dawson *et al*., 2008, 2010). As in all up to now characterized ThDP-dependent enzymes, a large hydrophobic residue next to the thiazolium ring (standard position 415) supports this essential conformation (Lindqvist *et al*., 1992; Andrews *et al*., 2013). In both *Ec*AHAS isozymes a methionine occupies this position, whereas both investigated MenDs contain an isoleucine.

Chemoselectivity & donor substrate spectrum – The donor-binding site

The binding order of the substrates determines the chemoselectivity of ThDP-dependent enzymes (chapter 1.7.4), which in turn depends on the space for the donor in the active site (area DB in Figure 23) and on the donors optimal stabilization in the respective binding pocket prior to carboligation (Pohl *et al*., 2009).

*Ec*AHAS-I and *Ec*AHAS-II

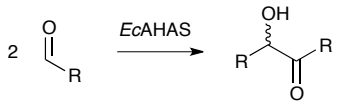
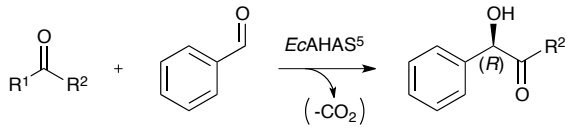
*Ec*AHAS catalyzes the formation of acetolactate (physiological reaction) and phenylacetylcarbinol (PAC) starting in both cases from pyruvate as the donor and a second pyruvate or benzaldehyde as the acceptor, respectively (chapter 1.7.3) (Chipman *et al*., 1998; Engel *et al*., 2003). The preference for pyruvate as the donor is mainly due to a small donor-binding site (Figure 23 E), which is comparable to the donor-binding pocket of *Ap*PDC (Figure 23 C). *Ap*PDC is also highly specific for pyruvate, but also accepts small aliphatic aldehydes like acetaldehyde and propanal as donor substrates (Rother *et al*., 2011). In contrast, for *Ec*AHAS decarboxylation of an α-keto acid is assumed to be essential prior to carboligation, and pyruvate was the only so far described donor substrate (Engel *et al*., 2003; Engel *et al*., 2004b; Vinogradov *et al*., 2006).

Comparison of the amino acid residues lining the donor-binding site of *Ec*AHAS and *Ap*PDC (Table 4) revealed that the donor-binding site of *Ap*PDC is restricted by T384 (388) and W388 (392). The corresponding residues in both *Ec*AHAS isozymes are valine and glutamine in standard positions 388 and 392, respectively, which were expected not to prevent the binding of small aliphatic aldehydes. Therefore, different aldehydes were tested as substrates for both *Ec*AHAS isoenzymes[*]. The results in Table 5 demonstrate that *Ec*AHAS-I and *Ec*AHAS-II are not restricted to pyruvate as donor substrate. Particularly, acetaldehyde and propanal are accepted as donors in the homocoupling to acetoin and propioin, respectively, as well as in the mixed carboligation with benzaldehyde. In the homocoupling reaction even the synthesis of small amounts of butyroin and (*R*)-benzoin could be observed starting from butanal and benzaldehyde, respectively. Although benzaldehyde can principally bind as a donor, smaller aliphatic aldehydes and α-keto acids are highly preferred as donors in the mixed carboligation reaction due to their better stabilization in the small donor-binding site. Thus, 2-hydroxypropiophenone (2-HPP) derivatives, resulting from benzaldehyde as the donor and aliphatic substrates as acceptors, could not be observed.

The use of α-keto acids donors resulted in much higher conversions in the mixed carboligation with benzaldehyde compared to the use of the corresponding aldehyde donors (*e.g*. 3 % compared to 89 %, Table 5). This is not surprising, because pyruvate is the physiological donor of *Ec*AHAS. However, the reason is not yet understood. No suitable residues were found in the active site of *Ec*AHAS, which might stabilize the carboxyl group of the pyruvate donor prior to the nucleophilic attack of ThDP.

[*]Performed by C. Schmitz in the context of his bachelor thesis (Schmitz, 2012) and a following research internship.

Table 5: *Ec*AHAS-I and *Ec*AHAS-II-catalyzed self-ligation of various aliphatic aldehydes (**A**), and the mixed carboligation of aliphatic aldehydes or the corresponding α-keto acids with benzaldehyde (**B**).[1]

Substrate / Enzyme	Conversion [%][2]	*ee* [%][3]	Conversion [%]	*ee* [%]
A				
R = CH_3	Acetoin			
*Ec*AHAS-I	22	≤ 5		
*Ec*AHAS-II	54	23 (*S*)		
R = C_2H_5	Propioin			
*Ec*AHAS-I	6	9 (*S*)		
*Ec*AHAS-II	17	31 (*S*)		
R = C_3H_7	Butyroin			
*Ec*AHAS-I	< 1	n.d.[4]		
*Ec*AHAS-II	< 1	n.d.		
R = C_6H_5	Benzoin			
*Ec*AHAS-I	< 1	98 (*R*)		
*Ec*AHAS-II	< 1	98 (*R*)		
B				
R^1 = H R^2 = CH_3	PAC		R^1 = CO_2^- R^2 = CH_3	PAC
*Ec*AHAS-I	3	98 (*R*)	89	98 (*R*)
*Ec*AHAS-II	2	97 (*R*)	42	98 (*R*)
R^1 = H R^2 = C_2H_5	PPC[6]		R^1 = CO_2^- R^2 = C_2H_5	PPC
*Ec*AHAS-I	3	> 98 (*R*)	10	> 98 (*R*)
*Ec*AHAS-II	5	> 98 (*R*)	14	> 98 (*R*)

[1]Reaction conditions: 0.1 M HEPES/KOH, pH 7.0, containing 60 mM KCl, 5 mM $MgCl_2$, 0.1 mM ThDP, 0.05 mM FAD, 10 vol% DMSO; 1 mg mL^{-1} purified enzyme; 30 mM substrates; 30 °C for 20 h at 750 rpm.
[2]Determined by GC and HPLC, respectively (Schmitz, 2012).
[3]The *ee* of aliphatic products was determined by chiral-phase GC, whereas the *ee* of araliphatic products was determined using chiral-phase HPLC (Schmitz, 2012).
[4]n.d.: not determined
[5]Self-ligation products acetoin, propioin, acetolactate, and acetohydroxybutyrate are not shown.
[6]PPC: phenylpropionylcarbinol

In summary, *Ec*AHAS closely resembles *Ap*PDC with respect to its chemoselectivity as well as the donor spectrum (Rother *et al*., 2011; Gómez Baraibar *et al*., 2013) due to comparable donor-binding sites. However, one major difference to *Ap*PDC is the ability of *Ec*AHAS to catalyze the formation and cleavage of acetolactate (chapter 1.7.3). This unique feature enabled the application of *Ec*AHAS-I in a novel recycling cascade reaction for the production

of pharmaceutically active nor(pseudo)ephedrine (Figure 24, **Publication I**). Here, two reactions are catalyzed by *Ec*AHAS-I: (i) the carboligation of pyruvate and benzaldehyde to (*R*)-PAC, and (ii) the ligation of two pyruvate molecules to acetolactate. The latter enables the efficient removal of the byproduct pyruvate, thereby shifting the reaction equilibrium* of the second reaction step to the product site.

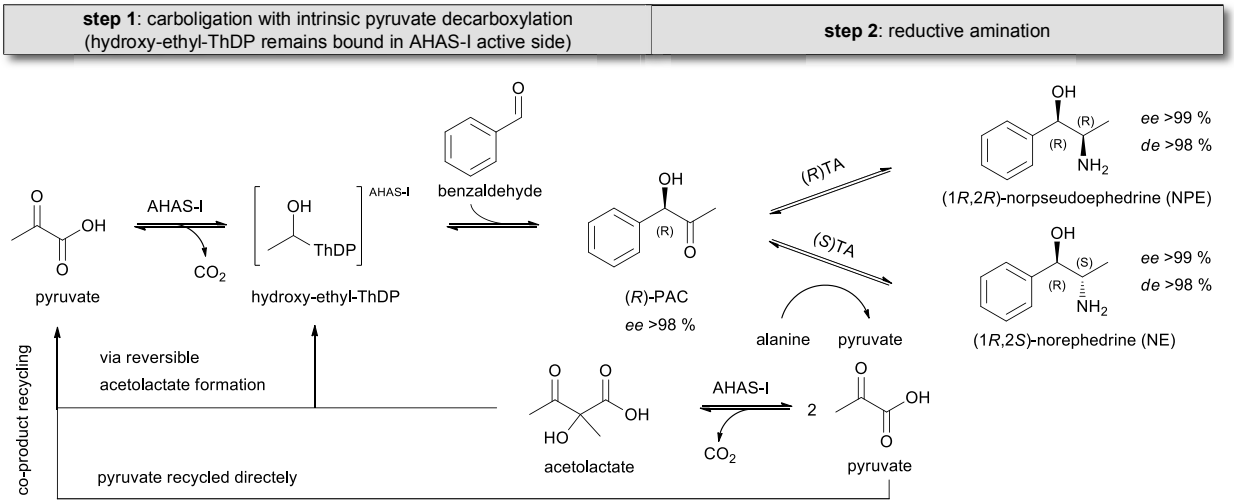


Figure 24: Recycling cascade for the synthesis of nor(pseudo)ephedrine combining *Ec*AHAS-I and ω-transaminases (**Publication I**). *Ec*AHAS-I catalyzes the carboligation of pyruvate and benzaldehyde to (*R*)-PAC in the first reaction step. A suitable ω-transaminase transforms (*R*)-PAC to the respective nor(pseudo)ephedrine in the second step. Thereby, the amine donor alanine is converted to pyruvate, which can be removed from the reaction equilibrium by *Ec*AHAS-I and recycled in the carboligation reaction of the first step.

Further, since acetolactate formation is reversible, pyruvate as well as hydroxyethyl-ThDP can be recycled in the first step of the cascade reaction, enabling very high atom efficiency and reduced byproduct formation (for detailed information see **Publication I**). Thus, the combination of *Ec*AHAS-I and suitable ω-transaminases provides access to (1*R*,2*R*)-norpseudoephedrine and (1*R*,2*S*)-norephedrine, respectively, in high optical purity (*de* > 98 %, *ee* > 99 %) and good conversions (up to 86 %) from inexpensive starting materials as visualized in Figure 24.

<u>*Ec*MenD and *Bs*MenD</u>

Both MenDs exhibit an excellent chemoselectivity catalyzing the formation of 5-hydroxy-4-oxo-5-arylpentanoates (PAC derivatives) from α-ketoglutarate (α-KG) as the donor substrate and benzaldehyde derivatives as acceptors (chapter 3.4.2, Table 7, **Publications III & IV**). Neither the self-ligation products of α-KG and benzaldehyde, respectively, nor the 5-oxo-4-

*For the reductive amination of (*R*)-PAC to (1*R*,2*S*)-norephedrine with alanine as amine donor, the thermodynamic equilibrium constant is 2.31×10^{-3}, which limits the theoretical conversion to 5 % if pyruvate is not removed (**Publication I**).

hydroxy-5-phenylpentanoate (2-HPP derivative) derived from benzaldehyde as the donor and α-KG as acceptor, could be observed (Kurutsch *et al*., 2009).

The size of the donor-binding site of MenD is between the small donor-binding sites of *Ap*PDC and *Ec*AHAS and the respective larger sites in *Pf*BAL, *Pp*BFD, and *Ll*KdcA (area DB in Figure 23). Based on the ability of *Ap*PDC and *Ec*AHAS to accept also benzaldehyde as a donor, which can be deduced from the weak benzoin formation observed with both enzymes (see above, Rother *et al*., 2011), MenD should also be able to bind benzaldehyde as the donor due to the larger donor-binding site. However, even with benzaldehyde as the sole substrate no benzoin formation could be detected. This high selectivity for α-KG is due to unique chemical characteristics of the donor-binding site (Table 4). In contrast to all other ThDP-dependent enzymes described in this thesis, the donor-binding sites of both MenDs have a highly polar character, mainly defined by two arginines in standard positions 393 and 410, respectively, and two asparagine residues in standard positions 114 and 388, respectively. Both arginine residues stabilize the terminal carboxylate of the ThDP-bound succinylsemialdehyde (Dawson *et al*., 2008, 2010). As a consequence, also other donor aldehydes, such as acetaldehyde, were not accepted by MenD (not shown)[*]. Even pyruvate and oxaloacetate were only poorly accepted by *Ec*MenD (Kurutsch, 2009; Kurutsch *et al*., 2009). Here, the higher conversions in the case of oxaloacetate (2–24 %) compared to pyruvate (1–5 %) might be due to its structural similarity to α-KG. Consequently, the donor spectrum of MenD is rather limited.

Stereoselectivity – The acceptor-binding site

After binding of the donor substrate and formation of the "activated aldehyde" (chapter 1.7.1) the acceptor can only bind in a strictly defined area of the active site (area AB in Figure 23). This binding may occur either in parallel or in antiparallel orientation relative to the ThDP-bound donor with the oxygen functions of donor and acceptor pointing in the same direction towards the proton relay system (chapter 1.7.4) (Knoll *et al*., 2006; Hailes *et al*., 2013). Here, an antiparallel arrangement is only possible in the presence of a suitable *S*-pocket, which is widespread among ThDP-dependent enzymes of the DC superfamily, but in most cases blocked by bulky amino acid residues (Gocke, 2007; Rother *et al.*, 2011).

[*] *Ec*MenD was tested with acetaldehyde, propanal and benzaldehyde in the self-ligation reaction as well as in the mixed carboligation with benzaldehyde. The presence of α-hydroxy ketones was analyzed by using the triphenyl tetrazolium chloride (TTC) assay (Breuer *et al*., 2002). Red coloration, which indicates product formation, could not be detected.

*Ec*AHAS-I and *Ec*AHAS-II

Having a closer look on the protein backbone, *S*-pockets are also present in *Ec*AHAS-I and *Ec*AHAS-II (Figure 23 E). In both cases the *S*-pockets are mainly blocked by a valine in standard position 477 (Table 4). In addition, a bulky amino acid is found at the entrance of the *S*-pockets: L476 (476) in *Ec*AHAS-I and M460 (476) in *Ec*AHAS-II, respectively. Thus, an antiparallel arrangement of the acceptor benzaldehyde relative to the ThDP-bound donor is not possible explaining the high *R*-selectivity of both *Ec*AHAS isoenzymes for the formation of PAC (*ee* = 98 %, Figure 25 A, Table 5) (Engel *et al*., 2003). In contrast, smaller aliphatic acceptor aldehydes are able to bind in antiparallel orientation indicated by the formation of (*S*)-acetoin and (*S*)-propioin, respectively (Table 5). In this case the stereocontrol is not very strict (*ee* ≤ 43% (*S*)), which is a common feature of ThDP-dependent enzymes for the carboligation of two aliphatic substrates due to missing stabilization of the small acceptor in the active site (Dominguez de Maria *et al*., 2007; Gocke *et al*., 2009; Gerhards *et al*., 2012). The potential size of the *S*-pockets of *Ec*AHAS-I and *Ec*AHAS-II is comparable to the *S*-pocket size of *Ll*KdcA and between that of *Pp*BFD and *Ap*PDC (Figure 23).

*Ec*MenD and *Bs*MenD

The structural investigation revealed also potential *S*-pockets in *Ec*MenD and *Bs*MenD that exhibit high structural similarity and exactly the same amino acid composition (Table 4, for structural alignment see **Publication IV**). Again, bulky amino acid residues block both *S*-pockets: phenylalanine in standard position 477 and isoleucine in standard position 476, respectively (Table 4, Figure 26 A). Accordingly, both MenDs are highly *R*-selective (*ee* > 99 %) in the mixed carboligation of α-KG as the donor substrate and benzaldehyde as acceptor (Table 7, **Publication III & IV**) (Beigi *et al*., 2013). The *S*-pocket sizes in both MenDs are similar to the *S*-pocket size of *Ap*PDC and thus, comparatively larger than those of *Ec*AHAS due to a slightly different position of the protein backbone (Figure 23 C, F).

Thorough analysis of the active site structures, biochemical data, and the comparison to well-characterized members of the DC superfamily enabled the deduction of structure-function relationships for *Ec*AHAS, *Ec*MenD, and *Bs*MenD. The gained information elucidated observed chemo- and stereoselectivities in the carboligation reaction and could be used as a basis for following structure-guided design of new variants with altered selectivities. Furthermore, *Ec*AHAS-I was established as a new toolbox enzyme and applied in a novel recycling cascade for the sustainable production of nor(pseudo)ephedrine from inexpensive starting materials.

3.4 Transfer and validation of the *S*-pocket concept

The introduction of the *S*-pocket concept enabled the rational design of *S*-selective ThDP-dependent enzymes (Gocke *et al*., 2008; Rother *et al*., 2011). Prerequisite is an *S*-pocket, which is large enough to harbor the side chain of the respective acceptor aldehyde in order to properly stabilize it. To confirm the general validity of the existing *S*-pocket concept for ThDP-dependent enzymes of the DC superfamily and to broaden the accessible (*S*)-α-hydroxy ketone range, the principles of stereoselectivity were transferred to *Ec*AHAS and MenD.

3.4.1 *Ec*AHAS-I and *Ec*AHAS-II

The non-physiological carboligation of (decarboxylated) pyruvate and benzaldehyde to (*R*)-PAC is described for all three isozymes of *Ec*AHAS (Engel *et al*., 2003). The two of them with the highest carboligation potential, *Ec*AHAS-I and *Ec*AHAS-II (Engel *et al*., 2003), were studied in more detail. The design of *S*-selective *Ec*AHAS variants would not broaden the platform of (*S*)-α-hydroxy ketones, since an *Ap*PDC variant for the *S*-selective PAC formation has become available recently (Rother *et al*., 2011). However, the application of *S*-selective *Ec*AHAS-I variants would give access to (1*S*,2*R*)- and (1*S*,2*S*)-nor(pseudo)ephedrine simultaneously enabling coproduct removal and recycling, which is not yet possible *via* the sustainable recycling cascade (see above, **Publication I**).

The *S*-pocket of *Ap*PDC could be opened by only a single mutation of E469 (477) to glycine, allowing benzaldehyde to bind in an antiparallel mode and thus giving access to (*S*)-PAC (Rother *et al*., 2011). As described before, the homology models of *Ec*AHAS suggested that the potential *S*-pockets are smaller compared to *Ap*PDC and thus probably not sufficiently accessible for antiparallel-oriented benzaldehyde (chapter 3.3.2). Modeling of benzaldehyde into the *S*-pockets of *Ec*AHAS-I and *Ec*AHAS-II supported this assumption, exemplarily shown for *Ec*AHAS-II in Figure 25. Although the *S*-pocket of *Ec*AHAS-II can be opened to some extent by mutation of V461 (477) and M460 (476) to glycine, antiparallel-oriented benzaldehyde still seems to clash with the *S*-pocket α-helix (protein backbone) (Figure 25 B). Compared to *Ap*PDC, this α-helix is shifted into the direction of ThDP and thus decreases the size of the resulting *S*-pocket (Figure 25 C).

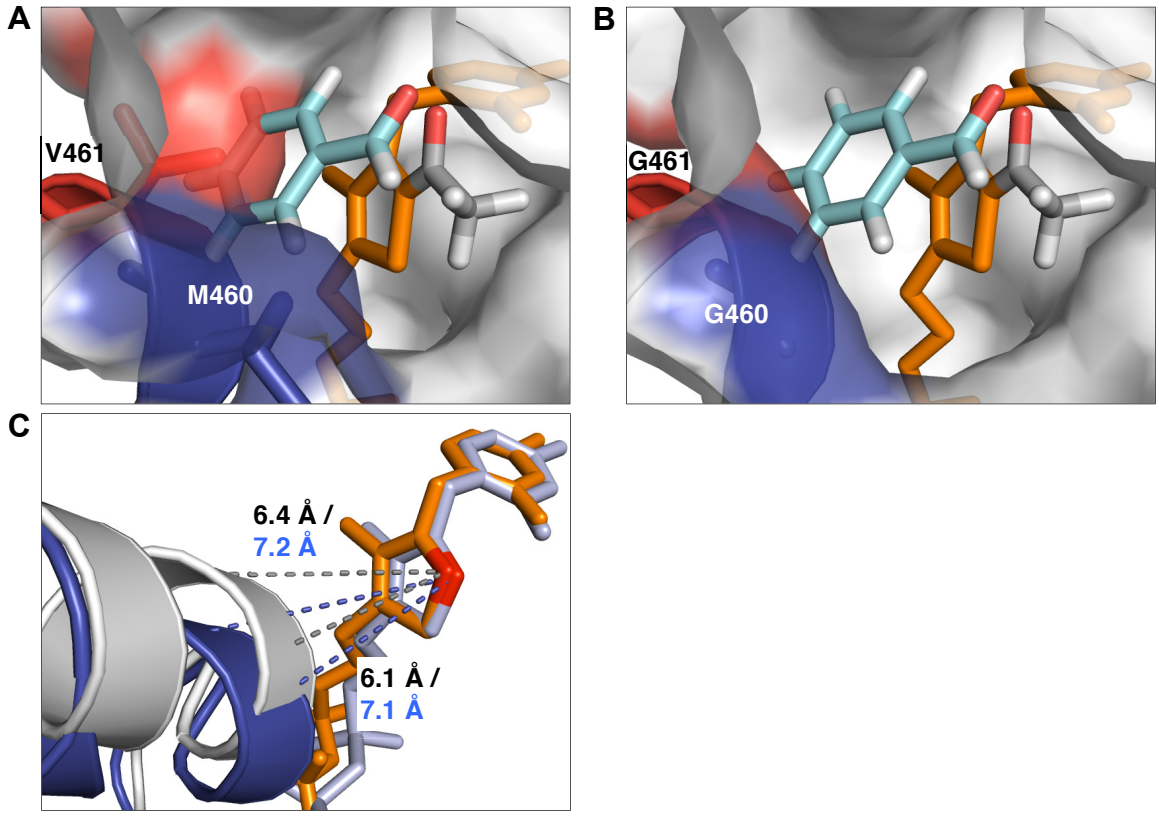
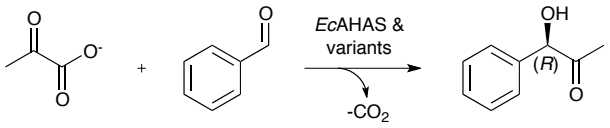


Figure 25: A: The *S*-pocket region of *Ec*AHAS-II with modeled substrates: antiparallel-oriented benzaldehyde (cyan) and hydroxyethyl ("activated aldehyde", gray) bound to the C2 of ThDP (orange). An antiparallel arrangement of the acceptor benzaldehyde relative to the ThDP-bound donor is not possible due to a blocked *S*-pocket. **B:** The exchange of M460 (blue) and V461 (red) to glycine might open the *S*-pocket, however, benzaldehyde still collides with the protein backbone of the *S*-pocket α-helix. **C:** Compared to *Ap*PDC (blue), the *S*-pocket α-helix of *Ec*AHAS (white) is slightly shifted towards ThDP (orange) decreasing the potential *S*-pocket size. ThDP of *Ap*PDC is colored in light blue. The C2 atom of ThDP is colored red.

To support the *in silico* modeling studies, which were based only on homology models, biochemical characterization of *Ec*AHAS-I and *Ec*AHAS-II variants was performed. Indeed, the high *R*-selectivity of *Ec*AHAS with respect to the formation of PAC (*ee* = 98 %) was only affected to a minor extent (Table 6). In the case of *Ec*AHAS-I, substitution of V477 (477) and L476 (476) by glycine, respectively, resulted in complete inactivation or drastic reduction of the catalytic activity, respectively. *Ec*AHAS-I-L476G exhibited only slightly reduced *R*-selectivity for the formation of PAC (*ee* = 90 %). In contrast, *Ec*AHAS-II was much more robust towards mutations in the *S*-pocket region, but again stereoselectivity was only slightly altered. The "best" result was obtained with variant M460G/V461G, which provides the largest possible *S*-pocket, yielding (*R*)-PAC with 82 % *ee*. Thus, benzaldehyde may react *via* the *S*-pathway, but is not sufficiently stabilized in the *S*-pocket due to steric restrictions through the protein backbone of the *S*-pocket α-helix (Figure 25 B, C).

Table 6: Mixed carboligation of pyruvate and benzaldehyde to PAC catalyzed by wild-type *Ec*AHAS-I, *Ec*AHAS-II and variants thereof.[1*]

Enzyme variant	Conversion [%][2]	*ee* [%][3]
		
*Ec*AHAS-I[4]		
wild-type	23	> 98 (*R*)
L476G	< 1	90 (*R*)
V477G	no conversion	n.d.[5]
*Ec*AHAS-II[6]		
wild-type	26	> 98 (*R*)
M460G	3	94 (*R*)
V461G	19	91 (*R*)
M460G/V461G	6.5	82 (*R*)

[1–3]See Table 5 for reaction conditions (but without addition of DMSO) and analytics.
[4]Purified *via* Ni-NTA Spin kit (Qiagen, Hilden, Germany). 0.1 mg mL^{-1} of the resulting eluat was used for the carboligation reaction (Schmitz, 2012).
[5]n.d.: not determined, [6]0.32 mg mL^{-1} lyophilized enzyme

In summary, it can be concluded that although the stereoselectivity of *Ec*AHAS could not be inverted, the results are in accordance with the existing *S*-pocket concept. Steric properties of the *S*-pocket are decisive for the stereoselectivity of ThDP-dependent enzymes. In the case of *Ec*AHAS, the maximally opened *S*-pocket is still too small to properly bind and stabilize antiparallel-oriented benzaldehyde preventing *S*-selective PAC formation (Figure 23 E & 25 B). Thereby, the experimental data nicely corroborated the prediction based on *in silico* modeling, which indicates the already good prediction quality of the model.

3.4.2 *Ec*MenD and *Bs*MenD[#]

As described in chapter 3.3.2, *Ec*MenD and *Bs*MenD are strictly *R*-selective for the synthesis of 5-hydroxy-4-oxo-5-phenylpentanoate (*ee* > 99 %), but both reveal large potential *S*-pockets with a similar size compared to that of *Ap*PDC by only looking at the protein backbone in the *S*-pocket of the crystal structure. Modeling studies supported the assumption that upon opening by mutagenesis of respective bulky residues the *S*-pockets should be large enough to bind antiparallel-oriented benzaldehyde (exemplarily shown for *Ec*MenD in Figure 26, **Publications II–IV**).

[*]Performed by C. Schmitz in the context of his bachelor thesis (Schmitz, 2012) and a following research internship.
[#]*Bs*MenD studies were performed by S. Jansen in the context of his bachelor thesis (Jansen, 2013).

Transfer of the S-pocket concept to EcMenD

The *S*-pocket concept was first transferred to *Ec*MenD. In *Ap*PDC, a single mutation at standard position 477 was sufficient to allow the acceptor benzaldehyde to bind in antiparallel orientation and thus to invert stereoselectivity (Rother *et al.*, 2011). In contrast, only the mutation of F475 (477) to glycine did not invert the stereoselectivity of *Ec*MenD (Table 7), because I474 (476) at the entrance of the *S*-pocket mainly prevented the antiparallel benzaldehyde orientation (Figure 26 B). As expected, the substitution of both bulky residues by glycine resulted in the first *S*-selective variant for the synthesis of 5-hydroxy-4-oxo-5-phenylpentanoate (I474G/F475G, *ee* 34 %, **Publications II–III**).

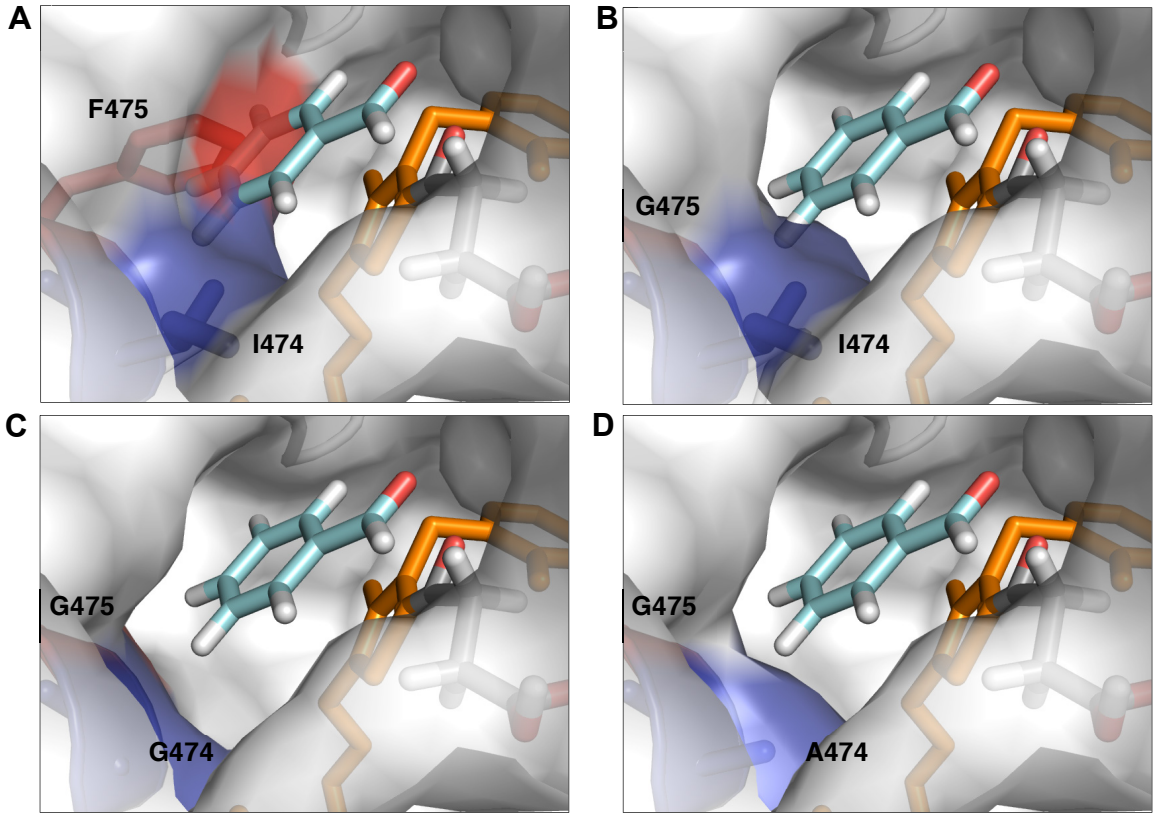
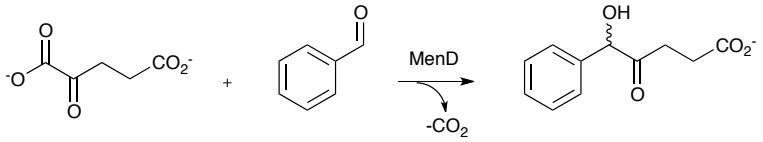


Figure 26: Tailoring the active site of *Ec*MenD (adapted from **Publications II & III**). **A**: The *S*-pocket of *Ec*MenD is blocked by F475 (red) and I474 (blue) preventing the antiparallel arrangement of benzaldehyde (cyan) relative to the ThDP-bound succinylsemialdehyde ("activated aldehyde", gray). **B:** Mutation of F475G opens the *S*-pocket, but I474 at the entrance of the *S*-pocket hinders benzaldehyde to enter properly. **C:** The additional mutation of I474 to glycine completely opens the large *S*-pocket allowing the antiparallel arrangement of benzaldehyde. **D:** Reduction of the *S*-pocket size in position 474 by mutation to alanine optimizes the stabilization of antiparallel-oriented benzaldehyde. The ThDP is colored in orange.

The *S*-pocket of *Ec*MenD was then tailored by stepwise enzyme engineering in order to enhance the low *S*-selectivity. Here, the influence of steric as well as stabilization effects on stereoselectivity could be clearly demonstrated (Table 7, **Publication III**). Whereas, the now very large *S*-pocket of *Ec*MenD-I474G/F474G stabilized benzaldehyde only suboptimal

(Figure 26 C), stabilization could be improved by substitution of glycine to alanine in position 474. This not only reduced the *S*-pocket size, but further improved stabilization of the benzaldehyde phenyl ring by nonpolar interactions as a result of the additional methyl group of A474 (Figure 26 D), which is fully consistent with the *S*-pocket concept. The better the stabilization of the acceptor side chain in the *S*-pocket the higher the *S*-selectivity (Gocke *et al*., 2008; Rother *et al*., 2011).

Table 7: Stereoselective mixed carboligation of α-ketoglutarate and benzaldehyde to 5-hydroxy-4-oxo-5-phenyl-pentanoate catalyzed by wild-type *Ec*MenD, *Bs*MenD and variants thereof (**Publications III & IV**).

*Ec*MenD variant[1]	Conversion [%][2]	*ee* [%][3]	*Bs*MenD variant[4]	Conversion [%][2]	*ee* [%][3]
					
wild-type	> 99	> 99 (*R*)	wild-type	> 99	> 99 (*R*)
F457G	96	83 (*R*)	I489G/F490G	28	81 (*S*)
I474G	77	70 (*R*)	I489A/F490G	40	75 (*S*)
I474G/F475G	26	34 (*S*)	I489G/F490A[5]	n.d.[6]	n.d.
I474A/F475G	15	75 (*S*)	I489A/F490A	36	92 (*S*)
I474S/F475G	8	23 (*R*)			
I474V/F475G	44	86 (*R*)			
I474G/F475A	67	27 (*S*)			
I474G/F475V	69	61 (*R*)			
I474A/F475A	68	66 (*S*)			

[1]Reaction conditions: 50 mM TEA buffer, pH 8.0, containing 2 mM $MgSO_4$, 0.1 mM ThDP and 5 vol% MTBE; 0.7 mg mL^{-1} enzyme; 50 mM α-KG and 20 mM benzaldehyde; 30 °C; 24 h. *Ec*MenD data were generated by S. Waltzer (Albert-Ludwigs-University Freiburg, Germany).
[2]Determined by GC based on consumption of benzaldehyde.
[3]Determined by chiral-phase HPLC.
[4]Reaction conditions: see [1], but without addition of MTBE, and 22 h reaction time.
[5]Obtained as insoluble protein upon expression in *E. coli* BL21 (DE3) and thus not further characterized.
[6]n.d.: not determined

All other substitutions in positions 474 and 475 did not improve *S*-selectivity relative to I474G/F475G or I474A/F475G, respectively, mostly due to unfavorable steric interactions, which could be nicely demonstrated by varying residues in position 475 (I474G/F475X). The results demonstrated a clear correlation between the size of the side chain and stereoselectivity for the synthesis of 5-hydroxy-4-oxo-5-phenylpentanoate, yielding higher ratios of the (*R*)-enantiomer with increasing size of the side chain: G475 (*ee* 34 % (*S*)) < A475 (*ee* 27 % (*S*)) < V475 (*ee* 61 % (*R*)) < F475 (*ee* 70 % (*R*)). Another special example is the

exchange of glycine to serine in position 474, which resulted again in an *R*-selective variant (Table 7, 23 % *ee*). Here, two effects may contribute to the change in stereoselectivity. The side chain of serine has only a slightly larger van der Waals volume relative to alanine, which may already result in minor unfavorable steric interactions with benzaldehyde. Even more important might be the polar character of serine that could lead to a reduced stabilization of benzaldehyde in the *S*-pocket, which in turn may explain the large difference in stereoselectivity between I474S/F475G (23 % *ee* (*R*)) and I474A/F475G (75 % *ee* (*S*)).

Stereoselectivity was thus shown to be sensitive towards already small changes of the side chains in the *S*-pocket (Table 7), which is fully consistent with the *S*-pocket concept. The comprehensive modeling and mutagenesis studies, and the stepwise engineering of two *S*-pocket positions at the same time with one model substrate (benzaldehyde), clearly stressed the importance of the two main aspects of the *S*-pocket concept: (i) sufficient space for the respective side chain of the acceptor substrate (= steric effects) and (ii) optimal stabilization of this side chain in the *S*-pocket according to size and hydrophobic/hydrophilic interactions.

Transfer of the S-pocket concept to BsMenD

The importance of a reliable model for the fast and successful (rational) design of *S*-selective enzyme variants could be demonstrated by the transfer of the principles of stereoselectivity to *Bs*MenD, using the mutagenesis strategy developed for *Ec*MenD (**Publication IV**). So far, *Bs*MenD was only characterized with respect to the physiological 1,4-addition of α-KG to isochorismate (Dawson *et al*., 2010), but as demonstrated in this work, it also catalyzes the non-physiological 1,2-addition of α-KG to benzaldehyde with high chemo- and stereoselectivity (*ee* > 99 %, Table 7). Due to high structural similarity of the *S*-pocket region compared to *Ec*MenD (chapter 3.3.2) and the results obtained before, respective *Bs*MenD variants were analogously prepared by mutation of I489 (476) and F490 (477) to glycine and alanine, respectively. Except for variant I489G/F475A, which was obtained as insoluble protein and thus not characterized, all variants revealed inverted stereoselectivity resulting in unexpectedly good *S*-selectivity for the formation of 5-hydroxy-4-oxo-5-phenylpentanoate (67–92 % *ee*) in comparison to the corresponding *Ec*MenD variants (34–75 % *ee*, Table 7). Thus, according to the *S*-pocket concept, so far best stabilization of the antiparallel-oriented benzaldehyde was achieved in *Bs*MenD-I489A/F490A (92 % *ee*).

Although the *S*-pocket regions of *Bs*MenD and *Ec*MenD are very similar with respect to the *S*-pocket geometry and involved amino acid residues, there are crucial differences that affect stereoselectivity. In addition to the significantly higher *S*-selectivity relative to *Ec*MenD

variants, which will be discussed in chapter 3.5.2, *Bs*MenD variants also revealed a different behavior in the presence of organic cosolvent (**Publication IV**).

Influence of organic solvents

Solvent engineering is an additional tool to alter stereoselectivity and activity of ThDP-dependent enzymes (chapter 1.7.5). Whereas stereoselectivity and catalytic activity, the latter estimated as endpoint conversions after 22 h and 24 h, were not affected by the addition of organic solvents in the case of both wild-type MenDs, *S*-selective MenD variants were susceptible to the addition of cosolvents. For example, the *S*-selectivity of I489A/F490A was reduced from 92 % *ee* in aqueous buffer to 85 % *ee* in the presence of 5 vol% MTBE, accompanied by a drop in catalytic activity (for further examples see **Publication IV**). These results are perfectly in line with the recently published studies of Gerhards *et al*., who demonstrated the influence of organic solvents on the *S*-selectivity of *Ap*PDC-E469G (chapter 1.7.5, Gerhards *et al*., 2012). Also in *S*-selective MenD variants MTBE may interact with the *S*-pocket and thus compete with benzaldehyde for the respective binding site to a certain extent. However, the *S*-selectivity of I489A/F490A was not affected in the presence of 20 vol% DMSO (91 % *ee*) and the catalytic activity was improved (from 36 % to 55 % conversion). Surprisingly, *S*-selective *Ec*MenD variants revealed contrary effects with respect to stereoselectivity. Whereas the addition of 5 vol% MTBE slightly improved the *S*-selectivity of *e.g*. variant I474A/F475G from 70 % *ee* to 75 % *ee*, the addition of 20 vol% DMSO resulted in drastically reduced *S*-selectivity (50 % *ee*).

The reasons for these different effects are yet not clear. However, the results confirm that solvent addition has to be carefully considered and evaluated for every reaction system, especially for *S*-selective ThDP-dependent enzymes, as was already stated by Gerhards *et al*. (Gerhards *et al*., 2012).

Catalytic activities of EcMenD and BsMenD

As shown in Table 7, the estimated catalytic activities of *S*-selective MenD variants were lower compared to the respective wild-type enzymes (**Publications II–IV**). These catalytic activities had to be estimated from the conversion data (endpoint detection), because the access to initial rate kinetic parameters was limited by the reaction system[*]. On the one hand,

[*]The low solubility of aromatic benzaldehydes prevents single substrate kinetics under saturated conditions. In addition, the activity of MenD in the non-physiological 1,2-addition is low, which in turn limits the determination of low conversions at low substrates concentrations under initial rate conditions. A reduced catalytic activity might also be a result of an instable enzyme variant. However, a reasonable activity assay for MenD is currently not available. The problem is that MenD does not release the respective aldehyde after decarboxylation of α-KG, which prevent the use of a convenient NADH-based coupled enzyme assay.

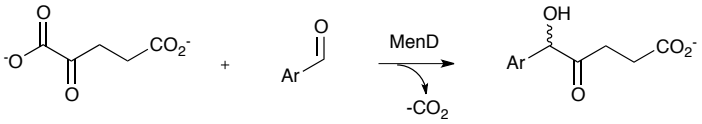
the apparently reduced catalytic activities of *S*-selective MenD variants might be caused by the close proximity of the standard positions 476 and 477 to the cofactor ThDP. The original residues isoleucine and phenylalanine stabilize ThDP by nonpolar interactions (Dawson *et al*., 2008, 2010). Substitution of both residues by smaller amino acids at the same time may reduce the binding affinity for the cofactor. On the other hand, the binding affinity of the substrates might also be reduced. Dawson *et al*. recently reported on increased K_m-values for the donor α-KG and the physiological acceptor isochorismate by a factor of 12–25 concomitant with a reduction in catalytic efficiency after mutagenesis of I489 and F490 to alanine in *Bs*MenD (Dawson *et al*., 2010).

Substrate screening

The *S*-pocket concept was successfully transferred to *Ec*MenD and *Bs*MenD enabling the *S*-selective synthesis of 5-hydroxy-4-oxo-5-phenylpentanoate, a valuable functionalized α-hydroxy ketone (chapter 1.7.3), from α-KG and benzaldehyde for the first time. Thereby, highest *S*-selectivity was observed with *Bs*MenD-I489A/F490A (92 % *ee*), whereas the best *S*-selective *Ec*MenD variant I474A/F475G only revealed moderate *S*-selectivity (75 % *ee*).

In addition to *S*-pocket engineering, the variation of the substrates is one option to further elucidate the potential of a biocatalyst with respect to higher *S*-selectivity and expansion of the α-hydroxy ketone platform. For *Pp*BFD and *Ap*PDC variants it was demonstrated that *S*-selectivity clearly depends on the donor substrates (Gocke *et al*., 2008; Rother *et al*., 2011; Gómez Baraibar *et al*., 2013). For example, *Ap*PDC-E469G catalyzed the formation of (*S*)-PAC with only a moderate *ee* of 61 % from acetaldehyde as the donor and benzaldehyde as acceptor. The stereoselectivity could be increased to 89 % *ee* (*S*) with propanal and 85 % *ee* (*S*) with butanal as the donor, probably because the larger donor substrates hamper the parallel benzaldehyde orientation (Figure 23 C, Rother *et al*., 2011). However, such a donor variation is not possible with MenD, because α-KG is the only well-accepted donor known so far (chapter 3.3.2, Kurutsch *et al*., 2009). Therefore, the benzaldehyde acceptor was varied, and different substituted benzaldehydes were tested concerning the influences of size and electronegativity of substituted substrates on stereoselectivity and conversion. Table 8 shows the results for both wild-type MenDs and the best *S*-selective variants thereof (**Publications II–IV**). From these results information concerning differences in stabilization of the benzaldehyde derivatives in the *S*-pocket of MenD were deduced.

Table 8: Carboligation of α-ketoglutarate and differently substituted benzaldehydes to 5-hydroxy-4-oxo-5-arylpentanoates catalyzed by wild-type *Ec*MenD, *Bs*MenD, and variants thereof (**Publications II– IV**).[1]

Ar	*Ec*MenD-wt[4] Conv. [%][2]	*ee* [%][3]	*Bs*MenD-wt Conv. [%]	*ee* [%]	*Ec*MenD-I474A/F475G Conv. [%]	*ee* [%]	*Bs*MenD-I489A/F490A Conv. [%]	*ee* [%]
								
C_6H_5	> 99	> 99 (*R*)	> 99	> 99 (*R*)	15	75 (*S*)	36	92 (*S*)
2-FC_6H_4	> 99	94 (*R*)	> 99	94 (*R*)	55	48 (*S*)	33	83 (*S*)
2-ClC_6H_4	98	93 (*R*)	> 99	81 (*R*)	57	< 5	< 5	45 (*S*)
2-BrC_6H_4	82	80 (*R*)	99	88 (*R*)	52	< 5	< 5	21 (*S*)
2-IC_6H_4	40	76 (*R*)	99	81 (*R*)	69	< 5	< 5	9 (*R*)
2-$MeOC_6H_4$	98	98 (*R*)	> 99	94 (*R*)	13	< 5	11	< 5
3-FC_6H_4	> 99	96 (*R*)	> 99	> 99 (*R*)	43	82 (*S*)	66	97 (*S*)
3-ClC_6H_4	> 99	96 (*R*)	> 99	97 (*R*)	94	89 (*S*)	17	97 (*S*)
3-BrC_6H_4	99	99 (*R*)	> 99	> 99 (*R*)	87	97 (*S*)	31	98 (*S*)
3-IC_6H_4	> 99	94 (*R*)	95	> 99 (*R*)	86	93 (*S*)	13	88 (*S*)
3-$MeOC_6H_4$	99	> 99 (*R*)	> 99	> 99 (*R*)	64	93 (*S*)	80	98 (*S*)
4-FC_6H_4	> 99	99 (*R*)	99	> 99 (*R*)	14	46 (*S*)	< 5	76 (*S*)
4-ClC_6H_4	> 99	93 (*R*)	> 99	97 (*R*)	15	48 (*R*)	29	< 5
4-BrC_6H_4	> 99	96 (*R*)	> 99	97 (*R*)	17	69 (*R*)	<5	37 (*R*)
3,5-di-$MeOC_6H_3$	99	98 (*R*)	> 99	99 (*R*)	83	96 (*S*)	47	98 (*S*)

[1–3]Reaction conditions, determination of conversion (Conv.) and *ee*: see Table 7. *Ec*MenD data were generated mainly by S. Waltzer (Albert-Ludwigs-University Freiburg, Germany).
[4]The 2-substituted benzaldehydes, 3-chloro- and 4-chlorobenzaldehyde recently reported in (Beigi *et al*., 2013).

Both wild-type MenDs are highly *R*-selective for the formation of 5-hydroxy-4-oxo-5-arylpentanoates. The reason for the lower *R*-selectivity in the case of *ortho*-substituted benzaldehydes is not known yet. As far as this can be decided from crystal structures obtained without substrates, the *S*-pocket should be blocked for any acceptor benzaldehyde in the wild-type enzymes (Figure 26 A). Owing to the flexibility of proteins however, especially *ortho*-substituted acceptor substrates might induce minor structural changes, which allow substrate binding in the antiparallel orientation to a certain extent.

In the case of *S*-selective MenD variants *ortho*- and *para*-substituted benzaldehydes resulted in higher ratios of the respective *R*-enantiomer relative to the reaction with benzaldehyde, whereas *meta*-substituted benzaldehydes were converted with higher *S*-selectivity. *Bs*MenD-I489A/F490A transformed all *meta*-substituted benzaldehydes, except 3-iodobenzaldehyde

(88 % *ee*), with excellent stereoselectivity (≥ 97 % *ee* (*S*)) and good conversions in most cases (Table 8). In the case of *Ec*MenD-I474A/F475G the higher *S*-selectivities (82–97 % *ee*) were accompanied by remarkable high conversions (43–94 %) relative to the reaction with benzaldehyde (15 %) under similar conditions. Obviously, *meta*-substituted benzaldehydes are optimal stabilized in the *S*-pocket of MenD, which results in higher *S*-selectivity. Moreover, in most cases higher *S*-selectivity correlated with highest conversion, indicating that good stabilization of the acceptor substrate in the *S*-pocket also positively affects the enzyme activity. Remarkably, the same preference for *meta*-substituted benzaldehydes could also be observed in the case of *Ap*PDC variants for the synthesis of (*S*)-benzoins starting from benzaldehyde derivatives (chapter 3.7.2, **Publication V**). However, the reasons for the preference of *meta*-substituted benzaldehydes as well as the switch in stereoselectivity in the case of *ortho*- and *para*-substituted benzaldehydes are yet not understood. Several aspects have to be considered including possible interactions of substrate and protein structure (flexibility), steric effects (size of the substituent), as well as electronic properties of the substituent (electronegativity). Indeed, a correlation could be found in the case of *para*-substituted benzaldehydes. *S*-Selectivity decreases and *R*-selectivity increases, respectively, with the size of the substituent in the series H < F < Cl < Br, probably due to steric effects in the *S*-pocket. Moreover, the *ee* seems to follow the Hammett correlation* in this case (Figure 27). Interestingly, a similar trend has been observed earlier in *Pp*BFD-catalyzed carboligations, but with substituted benzaldehydes as donors and acetaldehyde as acceptor yielding (*S*)-2-HPP derivatives (Liese, 2003).

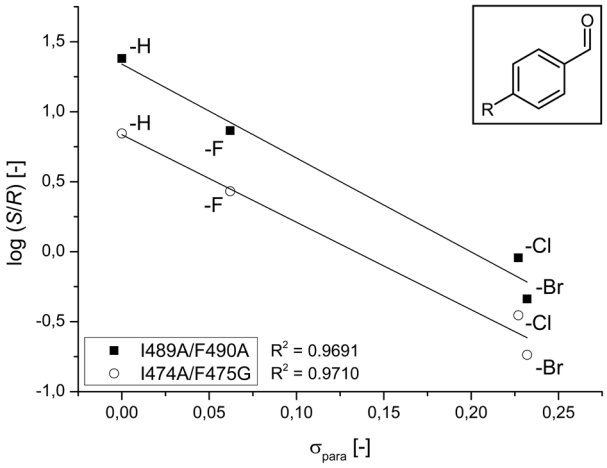


Figure 27: The logarithm of the molar ratio of the stereoisomers (*S*/*R*) of synthesized 5-hydro-4-oxo-5-arylpentanoates as function of the substituent constant σ_{para}, according to Hammett (Hammett, 1937). Log (*S*/*R*) is influenced by the different rates of formation of the *S* and *R* isomer, which in turn correlate with the electronic properties of the substituents in *para*-position according to the Hammett correlation (Hammett, 1937; Liese, 2003).

*The Hammett correlation considers the electronic properties of the different substituents of a benzene ring on the reaction rate (Hammett, 1937).

A possible approach to elucidate this phenomenon is the implementation of molecular dynamics (MD) simulations. Based on such simulations it is possible to calculate free enthalpies of the substrates in the active site, which can be subsequently compared to determine the energetically preferred substrate or mode of binding, *e.g.* *meta*-substituted versus *ortho*-substituted benzaldehyde in the *S*-pocket, or antiparallel-oriented versus parallel-oriented acceptors, respectively. To enable reliable MD simulations the structure of the respective enzyme variant must be known. Therefore, crystallization studies were started for *Ec*MenD-I474A/F475G and *Ec*MenD-I474A/F475G/R395Y (this variant is described in chapter 3.5.3) in cooperation with Prof. Dr. D. Dobritzsch (Department of Chemistry, Uppsala University, Sweden) during the course of this thesis. However, a structure could not yet be solved despite intensive efforts.

The strategy to improve *S*-selectivity by variation of the acceptor substrates was successful and additionally supported the existing *S*-pocket concept. In this context, the platform of α-hydroxy ketones could be enlarged by several new functionalized (*R*)- and (*S*)-5-hydroxy-4-oxo-5-arylpentanoates with excellent *ee*s.

The transfer of the principles of stereoselectivity to the new toolbox enzymes *Ec*AHAS, *Ec*MenD, and *Bs*MenD supported the general validity of the *S*-pocket concept for ThDP-dependent enzymes of the DC superfamily. Sufficient space for the respective side chain of the acceptor substrate in the *S*-pocket and its optimal stabilization are required for high *S*-selectivity. The current studies revealed that the prediction of stereoselectivities only based on *in silico* modeling is already quite good with respect to the mixed carboligation of aliphatic and aromatic substrates, which accelerates the development of new *S*-selective variants. Furthermore, the new toolbox enzyme MenD and *S*-selective variants thereof enable the enantiocomplementary synthesis of novel functionalized α-hydroxy ketones with up to > 99 % *ee*.

3.5 Refinement and extension of the *S*-selectivity concept

By introduction of the *S*-pocket concept, *S*-selective variants of *Pp*BFD, *Ap*PDC, and MenD could be rationally designed for the mixed carboligation reaction (Gocke *et al*., 2008; Rother *et al*., 2011, **Publication II–IV**). In all cases, bulky amino acids in standard position 477 (and 476 in the case of *Ec*MenD) were mutated to smaller ones in order to create space for the acceptor substrate in the *S*-pocket (*S*-pocket engineering). Although this approach was very successful in some cases, in other cases only stereoselectivities < 90 % *ee* could be reached under applied reaction conditions.

As already described before (chapter 3.4.2), one strategy to improve *S*-selectivity for selected reactions is the variation of the donor or acceptor substrate, respectively. Here, variation of the acceptor substrate as well as the alternative use of aldehydes or the corresponding α-keto acid as donors (chapter 1.7.1) (Gómez Baraibar *et al.*, 2013) can help to improve *S*-selectivity. Whereas the variation of the acceptor was successful in case of MenD yielding best results with *meta*-substituted benzaldehydes, a donor variation is not possible with MenD (chapter 3.3.2). However, substrate variation does not provide a general solution, if a biocatalyst is needed for a specific *S*-selective reaction starting from a "non-optimal" pair of substrates. Reaction optimization is one optimization tool (Gómez Baraibar *et al*., 2013), but also on the rational enzyme engineering site further improvement is possible. Therefore, new concepts were developed in this thesis to improve the *S*-selectivity of MenD variants and to provide new strategies based on protein engineering for other *S*-selective ThDP-dependent enzymes. At this point, the site-directed mutagenesis approach was extended by site-saturation mutagenesis*, which first required the development of a fast and reliable screening process to identify variants with reasonable carboligase activity and high *S*-selectivity.

3.5.1 Semi-automated screening system for ThDP-dependent enzymes

Site-saturation mutagenesis is a semi-rational approach, which enables the simultaneous examination of the impact of different amino acid residues on enzyme-specific parameters (here: activity and selectivity) in one specific position. To use this method for the investigation of ThDP-dependent enzymes, a semi-automated screening system was established, which enabled the fast and reliable screening for stereoselectivity of carboligation

*Saturation mutagenesis belongs to the method of directed evolution and is often applied when structural information is available. It facilitates the investigation of the influences of different amino acid residues on enzyme activity Moreover, this technique restricts randomization to a certain position in the protein and thus immensely reduces the screening effort compared to *e.g.* error-prone polymerase chain reaction (Reetz *et al*., 2008).

reactions using crude cell extracts in microtiter plates (Figure 28 A, for detailed experimental setup see **Publication III**). NDT codon degeneracy* (in which N = any nucleotide, D = adenine, thymine, or guanine) was applied to generate the mutant libraries. NDT libraries reduce the screening effort and are known to be of higher quality containing an increased number of positive hits (less "junk" transformants), therewith representing "smarter" and more focused libraries compared to other commonly used libraries (*e.g.* NKK) (Reetz *et al*., 2008; Bornscheuer & Kazlauskas, 2011). Moreover, saturation mutagenesis of one position# can be performed in a single 48 well FlowerPlate® (m2p-labs, Germany). These plates were chosen to achieve a higher cell density and good overexpression of the variants, which was facilitated by their unique form (Figure 28 B), enabling effective mixing, high oxygen transfer, and a big cultivation volume (2 mL) (Wenk *et al*., 2012). A good overexpression of the respective protein was a crucial step, especially for MenD, because the catalytic activity of new variants was not very high in many cases.

The cultivation was performed either in a BioLector® (m2p-labs, Germany), if only one FlowerPlate® was used, or in a microtiter plate incubator (*e.g.* INFORS HT, Switzerland), if more plates had to be screened. The BioLector® system (Figure 28 C) is a bench top micro cultivation system, which allows high-throughput cultivations and real time culture monitoring of common fermentation parameters, such as biomass, pH, and fluorescent molecules (m2p-labs, Germany; Funke *et al*., 2010). After cultivation, the plates were processed by a TECAN robot (Freedom EVO 200, TECAN, Switzerland) (Figure 28 A), which also performed the triphenyl tetrazolium chloride assay (Breuer *et al*., 2002) as initial screening for the presence of α-hydroxy ketones. Positive hits, indicated by red coloration, were finally analyzed on HPLC. Before use, the complete screening setup was tested with active *Ec*MenD variants as a positive control, which confirmed the general applicability of the system. Also the stereoselectivity of variants with low catalytic activity could be determined on HPLC. However, the measured *ee* was often lower compared to the results with purified enzyme, probably due to background reactions in the crude cell extract. For example, variant I474A/F475G exhibited an *ee* of 65 % for 5-hydroxy-4-oxo-5-phenylpentanoate in the screening, whereas the purified variant catalyzed the carboligation of α-KG and benzaldehyde with 75 % *ee*. Thus, to be able to detect improved variants (positive hits), the screening had to be always performed with a control (in this example *Ec*MenD variant I474A/F475G). Results

*NDT encodes for twelve amino acids (Arg, Asn, Asp, Cys, Gly, His, Ile, Leu, Phe, Ser, Tyr, and Val) representing a balanced mix of aliphatic and aromatic, polar and nonpolar, and negatively and positively charged amino acids.

#A screening of 46 colonies corresponds to a library completeness of 98 % (Patrick *et al*., 2003).

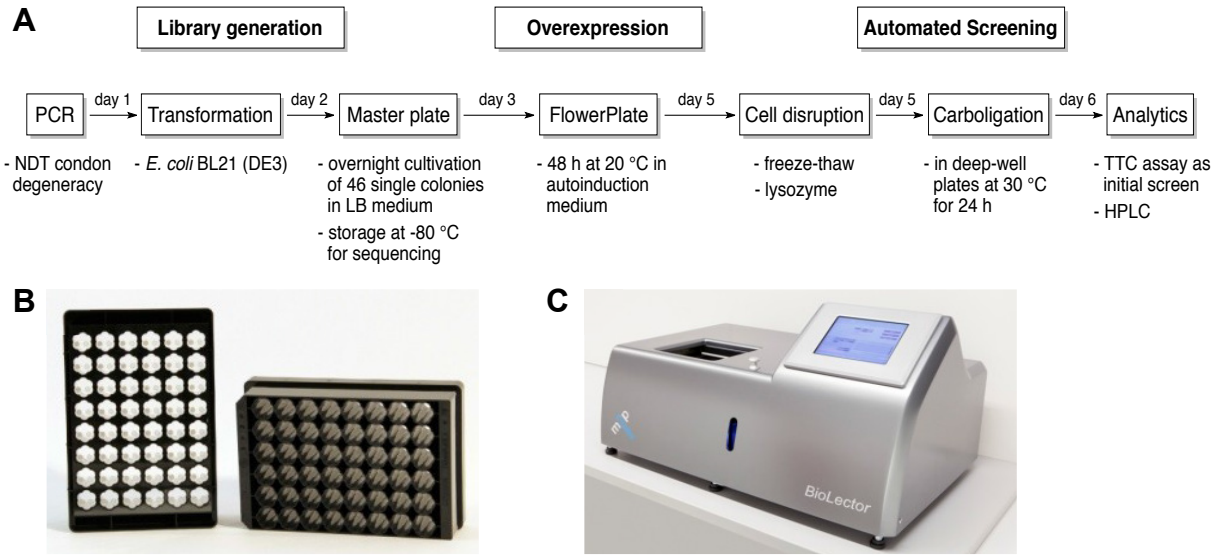


Figure 28: Setup for the fast and reliable screening for stereoselectivity of ThDP-dependent enzymes. **A:** The complete screening from the first polymerase chain reaction (PCR) to the determination of the *ee* by chiral-phase HPLC only takes 6 days. The automated screening was performed by the TECAN robot Freedom EVO 200 (TECAN, Switzerland). TTC assay: Triphenyl tetrazolium chloride assay. **B:** The characteristic FlowerPlate® design (m2p-labs, Germany; http://www.m2p-labs.com/products/microtiter-plates/flowerplates-mtp-48-bh). **C:** BioLector® cultivation unit (m2p-labs, Germany; http://www.m2p-labs.com/biolector-gbl100).

from promising screening hits were confirmed by more detailed analyses with purified enzymes.

It would be possible to fully automatized the developed semi-automated setup, except for the PCR and transformation, by integrating a picking robot and a cultivation unit (BioLector®, Rohe *et al*., 2012) to the existing robotic system.

3.5.2 Optimization of the *S*-pocket geometry

As already pointed out, the *S*-pocket geometry is essential for the stabilization of the antiparallel-oriented acceptor side chain. This geometry can be influenced by respective mutations of *S*-pocket residues, as shown for *Ec*MenD in positions 476 and 477 (standard numbers) (chapter 3.4.2, **Publication III**). In both positions, the mutations are yet limited to glycine and alanine, respectively (Gocke *et al*., 2008; Rother *et al*., 2011), because larger residues block the *S*-pocket again. But the *S*-pocket comprises of more than the up to now two varied amino acids. Thus, other residues in the *S*-pocket region were additionally investigated concerning their influence on stereoselectivity.

S-pocket residues

In addition to positions 476 and 477 (standard numbers), mainly three further amino acid residues in standard positions 26–28 define the *S*-pocket geometry of ThDP-dependent enzymes investigated in this thesis (Table 4). These positions are located on a highly

conserved flexible loop (exemplarily shown for *Ec*MenD in Figure 29), where mainly proline and exclusively glycine occupy positions 26 and 27, respectively. Both positions are not suitable for mutagenesis, since P(26) rather stabilizes the pyrimidine ring of ThDP than the acceptor substrate side chain and G(27) is located in close proximity to the acceptor side chain, in this example benzaldehyde (Figure 29). Most likely, a substitution of glycine to alanine in order to enhance stabilization of the acceptor in the *S*-pocket would already lead to unfavorable steric interactions. Thus, the only remaining interesting position in appropriate distance to the acceptor substrate is position 28, which is occupied either by aspartate, serine or glycine, respectively (Table 4).

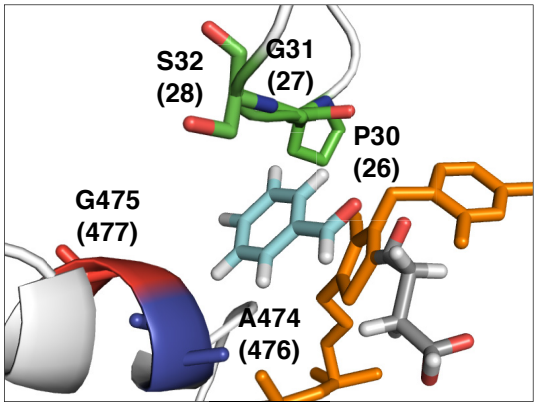


Figure 29: Important amino acid residues determining the *S*-pocket geometry of *Ec*MenD-I474A/F475G (adapted from **Publication III**). S32 (28) is located on a flexible loop and restricts the *S*-pocket from the top.

The influence of standard position 28 on stereoselectivity was investigated with *Ec*MenD-I474A/F475G using site-saturation mutagenesis (chapter 3.5.1, **Publication III**). The only variant obtained with still reasonable *S*-selectivity was I474A/F475G/S32C (*ee* 65 % with purified enzyme). In this case, the slightly larger van der Waals volume of cysteine relative to serine might already result in higher steric restrictions for benzaldehyde, which thus reduces stereoselectivity compared to I474A/F475G (*ee* 75 %). All other amino acid substitutions yielded either unselective or even *R*-selective variants. Even the separately prepared variant I474A/F475G/S32A, which was not included in the NDT codon library, revealed decreased *S*-selectivity (*ee* 24 %) accompanied by a drastically impaired conversion (< 5 % after 24 h), which could be due to a reduced catalytic activity and/or stability. Since S32 (28) is located in a flexible loop, the interpretation of these results based on a crystal structure without substrates is limited. However, the results suggest that the *S*-pocket geometry in variant I474A/F475G is the best of the possible options for benzaldehyde as an acceptor. Furthermore, standard position 28 is most likely not suitable for mutagenesis also in other ThDP-dependent enzymes.

Second-shell residues

As demonstrated by the studies of *Ec*MenD, *S*-pocket engineering is restricted to two positions: 476 and 477 (standard numbers). However, structure and sequence comparison of *S*-selective *Ec*MenD and *Bs*MenD variants, to explain significant differences in *S*-selectivity (chapter 3.4.2), revealed a second-shell residue*, which is crucial for the stereoselectivity of *Bs*MenD (**Publication IV**). This residue, G488 (475), is located in front of the main *S*-pocket residues I489 (476) and F490 (477), as part of the *S*-pocket α-helix (Figure 30 A). In *Ec*MenD, standard position 475 is occupied by a glutamine representing the main difference in the *S*-pocket regions of both enzymes (Figure 30 A, B). It was assumed that standard position 475 may influence the position of the *S*-pocket α-helix or its flexibility, which in turn affects the stereoselectivity. Mutation of G488 (475) to glutamine in *Bs*MenD-I489G/F490G, in order to mimic *Ec*MenD-I474G/F475G/Q473, supported this assumption. Whereas *Bs*MenD-I489G/F490G was highly *S*-selective for the formation of 5-hydroxy-4-oxo-5-phenylpentanoate, variant I489G/F490G/G488Q exhibited remarkably high *R*-selectivity for the same reaction (Figure 30 C).

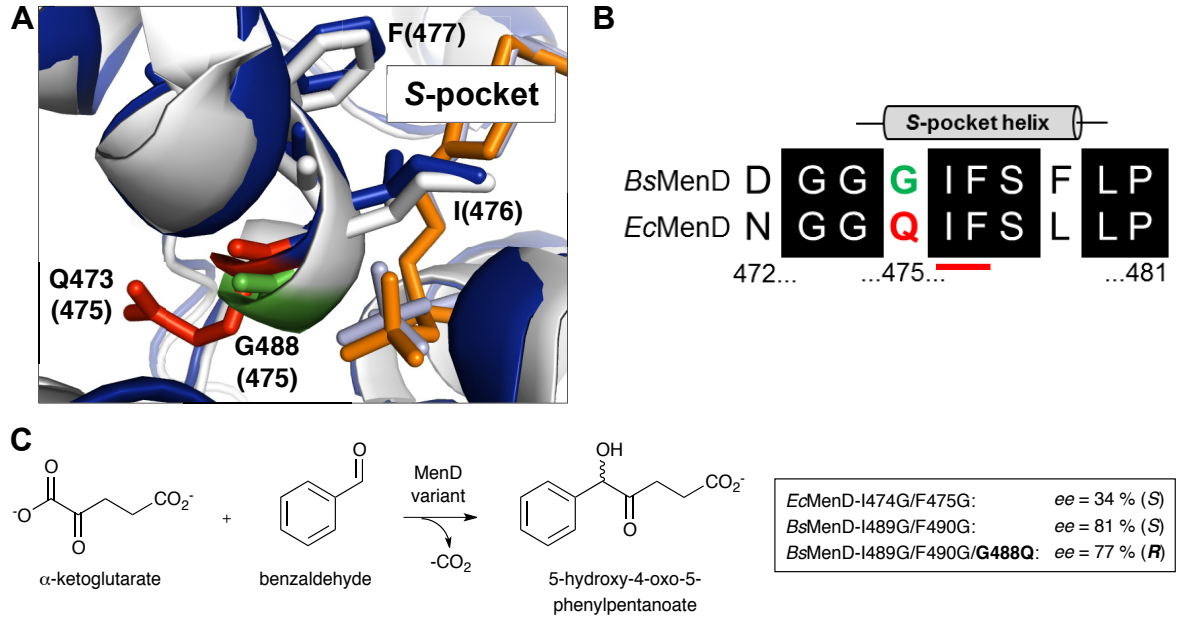


Figure 30: Influence of second-shell residue G488 (475) on the stereoselectivity of *Bs*MenD (**Publication IV**). The main difference in the strictly conserved *S*-pocket regions of *Bs*MenD and *Ec*MenD is standard position 475, directly in front of the most prominent *S*-pocket residues Ile (476) and Phe (477) (**A**, **B**). The amino acid side chain in standard position 475 is not directed towards the acceptor-binding site (*S*-pocket) (**A**), however mutation of G488 (475) in *Bs*MenD-I489G/F490G inverted the stereoselectivity (**C**). Thus, the higher *S*-selectivity of *Bs*MenD variants relative to *Ec*MenD variants might be due to a higher flexibility of the *S*-pocket induced by the three consecutive glycine residues adjacent to the *S*-pocket residues (underlined bold red, **B**).

*Second-shell residues are not in direct contact to the substrates in the active site, but may influence the catalytic activity as well as the selectivity of enzymes due to interaction with residues in the first shell that are in contact with the substrate.

There might be significant changes in the structure of the *S*-pocket region, such as a shift of the *S*-pocket α-helix that results in partial blockade of the *S*-pocket. Alternatively, substitution of one of the three consecutive glycine residues adjacent to the main *S*-pocket residues 476 and 477 (standard numbers, Figure 30 B) may affect the flexibility of the *S*-pocket. This potentially high flexibility of the *S*-pocket in *Bs*MenD variants (induced by the three consecutive glycine residues) might explain the higher *S*-selectivity compared to the corresponding *Ec*MenD variants, due to a better fitting of the benzaldehyde acceptor in the *S*-pocket (Table 7, 8, **Publication IV**).

Mutation of this newly identified residue in standard position 475 in other ThDP-dependent enzymes, *e.g*. *Ap*PDC with valine in this position, might be beneficial to increase *S*-selectivity by influencing the *S*-pocket backbone[*]. G488 (475) is the second example of a second-shell amino acid residue in ThDP-dependent enzymes that influences stereoselectivity. The first example was reported for *Pp*BFD, where the *S*-selectivity for the carboligation of benzaldehyde and acetaldehyde was improved by mutation of L476 (488) (Lingen *et al*., 2002). This residue is in close contact to the *S*-pocket residue L461 (477) allowing direct interactions, which might be influenced by mutation of L476.

3.5.3 Targeted destabilization of the "*R*-pathway"

As shown before, one strategy to enhance the *S*-selectivity of ThDP-dependent enzymes is to tailor the *S*-pocket geometry and thus enhance the stabilization of the acceptor substrate in the *S*-pocket, the so-called "*S*-pathway". As the set of proteinogenic amino acids limits the possibilities of *S*-pocket engineering, a viable alternative is the selective destabilization of the parallel acceptor orientation, the so-called "*R*-pathway" (**Publication III**). Because there is always a competition between antiparallel and parallel acceptor orientation in the active site, destabilization of the parallel orientation should in principle result in improved *S*-selectivity. This alternative strategy was investigated with *Ec*MenD-I474A/F475G, which only showed moderate *S*-selectivity in the formation of 5-hydroxy-4-oxo-5-phenylpentanoate (75 % *ee*).

Two amino acid residues, L478 (480) and R395 (393), were deduced to be important for the stabilization of the parallel benzaldehyde orientation either by nonpolar interactions in the case of L478, or by cation-π-interactions[#] in the case of R395, respectively (Figure 31 A). Site-saturation mutagenesis of L478 either resulted in reduced *S*-selectivity, *R*-selective

[*]Respective variants could not be prepared and tested in this thesis due to lack of time.

[#]Cation-π-interactions are noncovalent molecular interactions between electron-rich π-systems, *e.g.* a phenyl ring, and a cation, *e.g.* guanidinium cation (arginine side chain) (Blanco *et al*., 2011).

variants or in inactive variants. The reason for this might be the close proximity of this residue to the *S*-pocket (Figure 31 A), which might block the entrance to the *S*-pocket, influence the *S*-pocket geometry, or even influence the antiparallel benzaldehyde orientation (for further information see **Publication III**).

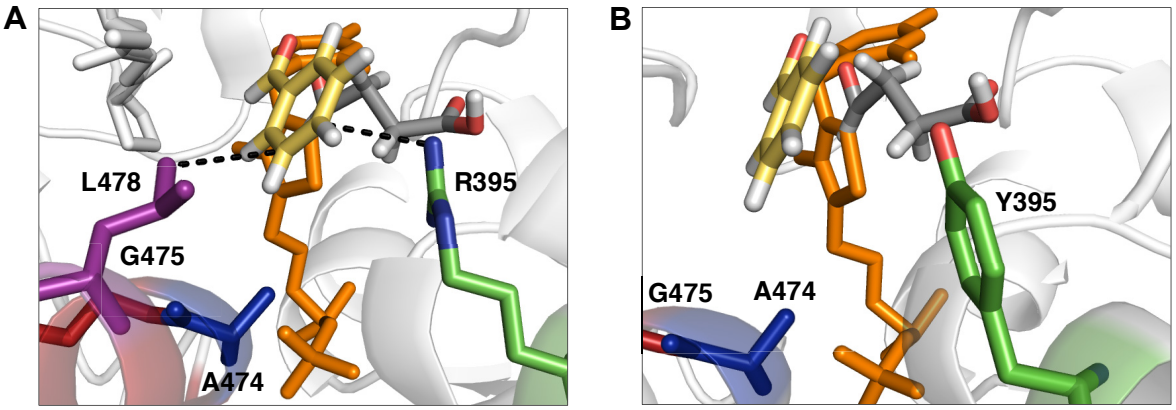
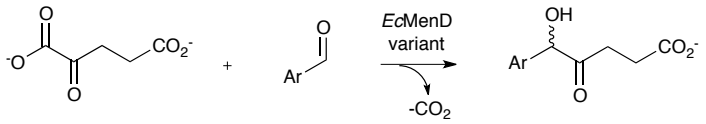


Figure 31: Potential stabilization of the parallel benzaldehyde orientation in the active site of *Ec*MenD-I474A/F475G (**Publication III**). **A:** L478 (purple) and R395 (green) are able to stabilize parallel-oriented benzaldehyde (yellow) by hydrophobic and cation-π-interactions, respectively. **B:** In contrast to R395, Y395 is not able to stabilize benzaldehyde by cation-π-interactions. Furthermore, the hydroxyl group of Y395 might additionally destabilize benzaldehyde in the parallel orientation. ThDP is colored in orange and succinylsemialdehyde in gray.

In contrast to L478, R395 is positioned more distantly from the *S*-pocket and neither influences the antiparallel benzaldehyde orientation nor the *S*-pocket itself. The mutant library in position 395 revealed one variant with significantly improved *S*-selectivity: I474A/F475G/R395Y (85 % *ee*). With tyrosine in this position, stabilizing cation-π-interactions are no longer possible. Moreover, the hydroxyl group of tyrosine might additionally destabilize parallel-oriented benzaldehyde (Figure 31 B), which thus improves *S*-selectivity. Also other substituted benzaldehyde derivatives as acceptor substrates were transformed with higher *S*-selectivity compared to I474A/F475G, supporting the high potential of this new concept (Table 9). Although the differences in *ee* were only between 2–10 %, a large activation free energy is required to achieve these changes (for > 80 % *ee*) according to equation: $\Delta G = -RT \ln[(100+ee) \times (100-ee)]$ (Phillips, 1992). In other words, the energy required for a few percent in *ee* costs much more energy at already high *ee*s compared to lower *ee*s, which highlights the importance of the additional mutation R395Y. The reduced catalytic activity of variant I474A/F475G/R395Y compared to variant I474A/F475G might probably result from the negative influences of mutation R395Y on the binding of α-KG. R395 with its positively charged side chain is able to stabilize the negatively charged γ-carboxylate group of the donor α-KG (Figure 31 A), whereas Y395 only supports the formation of hydrogen bonds.

Thus, the combination of *S*-pocket engineering and simultaneous suppression of the *R*-pathway provides a new strategy to improve the *S*-selectivity of ThDP-dependent enzymes. In addition to *Ec*MenD, the high potential of this novel strategy could be further demonstrated in this thesis for the optimization of *S*-selective *Ap*PDC variants (chapter 3.7.2).

Table 9: Carboligation of α-ketoglutarate and differently substituted benzaldehydes to 5-hydroxy-4-oxo-5-arylpentanoates catalyzed by *Ec*MenD-I474A/F475G and *Ec*MenD-I474A/F475G/R395Y (**Publication III**).[1]

	*Ec*MenD-I474A/F475G		*Ec*MenD-I474A/F475G/R395Y	
Ar	Conversion [%]	*ee* [%]	Conversion [%]	*ee* [%]
				
C_6H_5	15	75 (*S*)	7	85 (*S*)
2-FC_6H_4	55	48 (*S*)	9	57 (*S*)
2-ClC_6H_4	57	< 5	21	9 (*S*)
2-BrC_6H_4	52	< 5	24	< 5
2-IC_6H_4	69	< 5	20	15 (*R*)
2-$MeOC_6H_4$	13	< 5	2	n.d.[2]
3-FC_6H_4	43	82 (*S*)	13	87 (*S*)
3-ClC_6H_4	94	89 (*S*)	38	96 (*S*)
3-BrC_6H_4	87	97 (*S*)	29	> 99 (*S*)
3-IC_6H_4	86	93 (*S*)	14	96 (*S*)
3-$MeOC_6H_4$	64	93 (*S*)	21	96 (*S*)
4-FC_6H_4	14	46 (*S*)	3	n.d.
4-ClC_6H_4	15	48 (*R*)	4	n.d.
4-BrC_6H_4	17	69 (*R*)	4	n.d.
3,5-di-$MeOC_6H_3$	83	96 (*S*)	12	98 (*S*)

[1]For reaction conditions, determination of conversion, and *ee*: see Table 7. The data were generated mainly by S. Waltzer (Albert-Ludwigs-University Freiburg, Germany).
[2]n.d.: not determined

The design of *S*-selective ThDP-dependent enzymes was yet restricted to *S*-pocket engineering, which limits the access to highly *S*-selective enzyme variants. In this thesis, the *S*-selectivity concept could be refined and extended by identification of the role of a second-shell residue and targeted destabilization of the "*R*-pathway". The second-shell residue in front of the *S*-pocket influences its size and flexibility, and thus might be a new target for mutagenesis also in other enzymes in order to enhance *S*-selectivity. Furthermore, a new potent strategy to improve *S*-selectivity through selective destabilization of the "*R*-pathway" was demonstrated using *Ec*MenD. The establishment of a reliable semi-automated screening system for ThDP-dependent enzymes facilitates screenings of site-saturated mutant libraries.

3.6 The alternative *S*-pathway

Apart from the *S*-pocket and the antiparallel arrangement of the acceptor substrate relative to the donor, there is, theoretically, a second possibility to access (*S*)-α-hydroxy ketones: the parallel approach of the acceptor substrate from the backside of the donor (chapter 1.7.5). Previous studies with *Ap*PDC variant W388A indicated such a backside attack of the ThDP-bound donor (Rother *et al*., 2011) since reduced *R*-selectivity occurred while no *S*-pocket was accessible. This hypothesis was tested with the strictly *R*-selective *Pf*BAL. *Pf*BAL has no *S*-pocket (Figure 23 A), thus (*S*)-products may only originate *via* the alternative *S*-pathway.

The access to the donor backside of *Pf*BAL is blocked by mainly three amino acids: L112 (114), Q113 (115), and M421 (415) (Figure 32 A). All three residues might have important roles in the donor-binding site. L112 seems to be involved in stabilization of the donor, whereas Q113 is possibly involved in the proton transfer during the catalytic cycle (Table 4). M421 is crucial for stabilization of the conserved "V-conformation" of the cofactor ThDP (Lindqvist *et al*., 1992; Guo *et al*., 1998). As a consequence, mutations of these most likely essential amino acid residues might result in a loss of activity. However, *Pf*BAL turned out to be very robust towards mutations in the donor-binding region.

All prepared variants were analyzed concerning their carboligation potential using crude cell extracts (Table 10). Both single variants L112A and Q113A revealed decreased carboligation activity compared to wild-type *Pf*BAL accompanied by minor reduction of the *R*-selectivity for the formation of 2-HPP. However, this decrease could have also been caused by the use of crude cell extract (chapter 3.5.1). Even wild-type *Pf*BAL showed slightly reduced *R*-selectivity (*ee* 97 %) when applied as crude cell extract, compared to the excellent

stereoselectivity (*ee* > 99 %) obtained with purified enzyme (Demir *et al*., 2001). However, more interestingly, the chemoselectivity was significantly altered (chapter 1.7.4, Figure 16). Compared to wild-type *Pf*BAL, which exclusively catalyzes the formation of (*R*)-2-HPP and (*R*)-benzoin from benzaldehyde and acetaldehyde, variant L112A did not catalyze the synthesis of benzoin anymore. Instead, 2-HPP and small amounts of PAC were the only detectable products. In the case of variant Q113A, benzoin was the clearly preferred product and the activity towards the formation of 2-HPP was drastically reduced. The still reasonable activity of Q113A suggests that this glutamine is not involved in proton transfer reactions during the carboligation. Substitution of the corresponding glutamine in *Ec*AHAS-II and *Ec*MenD drastically impaired the catalytic activity supporting its importance for proton transfer (chapter 3.3.2, Fang *et al*., 2011; Vyazmensky *et al*., 2011). In *Pf*BAL, H29 (Table 4) most likely takes over this important function.

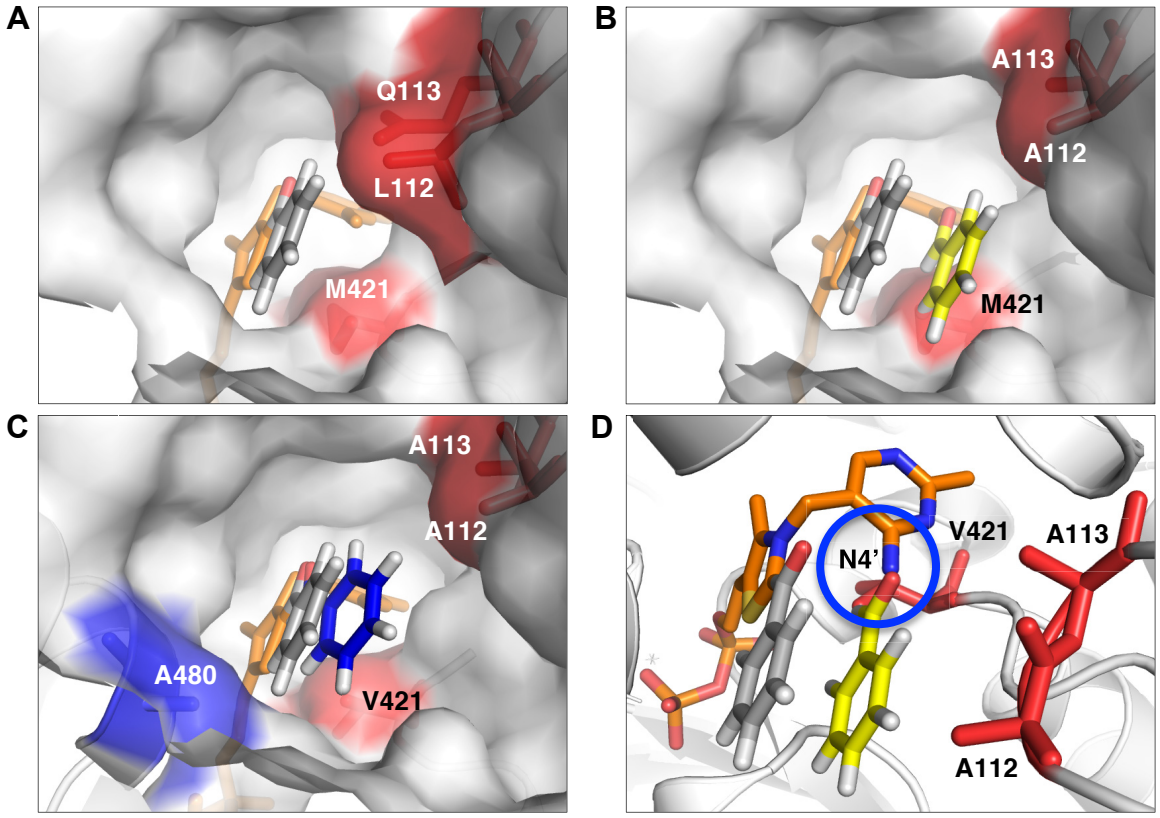
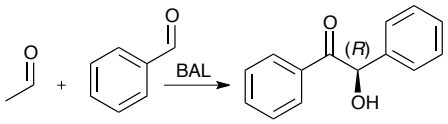
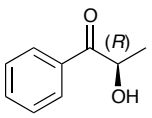
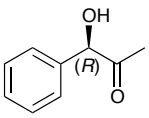


Figure 32: Investigation of the alternative *S*-pathway in *Pf*BAL. **A:** The backside of the ThDP-bound donor benzaldehyde (gray) is blocked mainly by L112, Q113, and M421 (red). **B:** Mutagenesis of L112 and Q113 to alanine open the backside of the donor revealing possible space for the acceptor benzaldehyde (yellow), which still seems to collide with M421. **C:** Although the donor-binding site was additionally opened, the stereoselectivity of L112A/Q113A/M421V was not significantly affected. Most likely, the missing stabilization of the ThDP-bound donor results in a movement of the donor (blue) towards the stabilizing amino acids A112 and A113, which in turn blocks the donor backside again. The introduction of large amino acids in position 480 (blue) might stabilize the original donor position of benzaldehyde (gray). **D:** A simple explanation for the inaccessible donor backside might be sterical hindrance of the acceptor benzaldehyde (yellow) with the 4'-imino group of the pyrimidine ring of ThDP (blue circle).

In the next step, both mutations were combined to increase the space for the acceptor at the backside of the donor (Figure 32 B). Additionally, variant L112A/Q113V was prepared. Valine in position 113 might better stabilize the acceptor at the backside of the donor relative to alanine. Both double variants, L112A/Q113A and L112A/Q113V, catalyzed moderate formation of (*R*)-benzoin and (*R*)-2-HPP. But again, stereoselectivity was not significantly affected for (*R*)-benzoin and only slightly reduced for (*R*)-2-HPP formation compared to wild-type *Pf*BAL (Table 10), although the donor-binding site was increased in both variants according to the model. PyMOL studies revealed that M421 might still limit the access to the donor backside, especially for the benzaldehyde acceptor. Mutation of M421 to valine additionally opened the donor backside (not shown). In this case valine was chosen, because mutation of M421 to glycine and alanine, respectively, were deleterious for *Pf*BAL activity in previous studies (personal communication with F. H. Andrews, Department of Chemistry and Chemical Biology, Indiana University-Purdue University Indianapolis, USA). Again, the resulting variants, L112A/Q113A/M421V and L112A/Q113V/M421V, were highly *R*-selective (Table 10). However, both triple variants showed remarkable catalytic activity. Whereas PAC and 2-HPP formation were strongly impaired, both triple variants showed high selectivity for the formation of (*R*)-benzoin accompanied by high conversions under the tested conditions. This is rather surprising, because M421, as bulky hydrophobic residue, is known to be important for stabilization of the conserved cofactor "V-conformation" (Guo *et al*., 1998; Andrews *et al*., 2013). Moreover, there are already three mutations in the donor-binding site of *Pf*BAL.

As a hypothesis, the still unaffected stereoselectivity may result from a movement of the ThDP-bound benzaldehyde (Figure 32 C) that might be induced by the decreased stabilization of the ThDP-bound donor in the large donor-binding site of the BAL variant. To obtain the required stabilization necessary for catalysis, the hydroxybenzyl-ThDP might move towards positions 112 and 113, which in turn would block the access to the donor backside again, even for the small acetaldehyde acceptor. In order to overcome this hypothetic movement, a large hydrophobic amino acid was introduced in position 480 (Figure 32 C) to create a kind of mirror image of the original donor-binding site. Therefore, A480 was substituted by leucine in order to stabilize the ThDP-bound benzaldehyde from the opposite site, which in turn might enable the desired backside access of an acceptor. However, also this approach did not affect the stereoselectivity of *Pf*BAL-L112A/Q112A/M421V/A480L, which is still highly *R*-selective for the formation of benzoin and 2-HPP (Table 10), and again surprisingly active, which highlights the robustness of *Pf*BAL.

Table 10: Mixed carboligation of acetaldehyde and benzaldehyde catalyzed by wild-type *Pf*BAL and variants thereof.[1]

	Benzoin		2-HPP		PAC	
BAL variant[4]	Conv. [%][2]	*ee* [%][3]	Conv. [%]	*ee* [%]	Conv. [%]	*ee* [%]
						
wild-type	72	> 99 (*R*)	19	97 (*R*)	< 1	n.d.[5]
L112A	not detected	n.d.	18	97 (*R*)	4	n.d.
Q113A	30	99 (*R*)	< 1	92 (*R*)	not detected	n.d.
L112A/Q113A	12	99 (*R*)	8	94 (*R*)	< 1	n.d.
L112A/Q113V	10	98 (*R*)	5	91 (*R*)	not detected	n.d.
L112A/Q113A/ M421V	90	> 99 (*R*)	2	93 (*R*)	2	n.d.
L112A/Q113V/ M421V	87	> 99 (*R*)	< 1	90 (*R*)	< 1	n.d.
L112A/Q113A/ M421V/A480L	43	99 (*R*)	4	98 (*R*)	not detected	n.d.

[1]Reaction conditions: 50 mM TEA buffer, pH 8.0, containing 2 mM $MgSO_4$, and 0.1 mM ThDP.; 1 mg mL^{-1} crude cell extract; 18 mM acetaldehyde and 18 mM benzaldehyde; 20 °C; 20 h. Analytics were performed mainly by U. Mackfeld (IBG-1, Forschungszentrum Jülich GmbH, Germany). Acetoin was not determined.
[2]Conversion (Conv.) determined by HPLC.
[3]Determined by chiral-phase HPLC.
[4]Variants were prepared mainly by D. Hahn (IBG-1, Forschungszentrum Jülich GmbH, Germany). *Pf*BAL sequence and primers see Appendix 5.2.
[5]n.d.: not determined

Accordingly, the high *R*-selectivity of *Pf*BAL could not be significantly altered yet. Thus, even the small acetaldehyde acceptor seemed not be able to properly bind at the backside of the donor, although a very large donor-binding site was already available according to the model (Figure 32 C). In addition to a potential movement of the ThDP-bound donor, which would block the donor backside again, another simple explanation for this result might be the sterical hindrance at the backside of the donor due to the pyrimidine ring of ThDP, especially the 4'-imino group (Figure 32 D). This unfavorable interaction would even prevent the arrangement of acetaldehyde. In addition, when the donor is already bound, it is less likely that the acceptor anneals from the backside, since chances are higher to be properly arranged in the substrate channel than to pass by the ThDP-bound donor.

In summary, although an alternative pathway could not be proven for the chosen reaction system, some interesting results and insights into the chemoselectivity of *Pf*BAL could be obtained, which have to be further investigated with purified enzyme.

3.6.1 *Ap*PDC-W388A and the alternative *S*-pathway

The main question still remains: why did variant *Ap*PDC-W388A exhibit a reduced *R*-selectivity although the *S*-pocket was closed? This was elucidated by a closer look into the equilibrated structure of *Ap*PDC-W388A*. Compared to wild-type *Ap*PDC (*ee* 96 %), *Ap*PDC-W388A revealed a significant drop in *R*-selectivity (*ee* 50 %) for the formation of phenylpropionylcarbinol (PPC) from propanal as the donor and benzaldehyde as the acceptor, indicating an alternative *S*-pathway, since the *S*-pocket is closed (chapter 1.7.5, Rother *et al*., 2011). However, the equilibrated structure of *Ap*PDC-W388A with ThDP-bound propanal revealed that a parallel arrangement of the benzaldehyde acceptor at the backside of the donor should not be possible at all (Figure 33). The access to the backside is clearly blocked by T384. In addition, the space (at the backside) is limited by the protein backbone of G409. Furthermore, the propanal side chain might also restrict the access to the donor backside.

As later on demonstrated, T384 mainly limits the size of the donor-binding site and not residue W388. T384 prevents the proper stabilization of benzaldehyde already in the donor position (chapter 3.7.2). Accordingly, binding of benzaldehyde as acceptor at the backside of the donor should be hardly possible. Therefore, the reasons for the reduced *R*-selectivity of *Ap*PDC-W388A can currently not be explained based on structural data of the wild-type enzyme and available biochemical data.

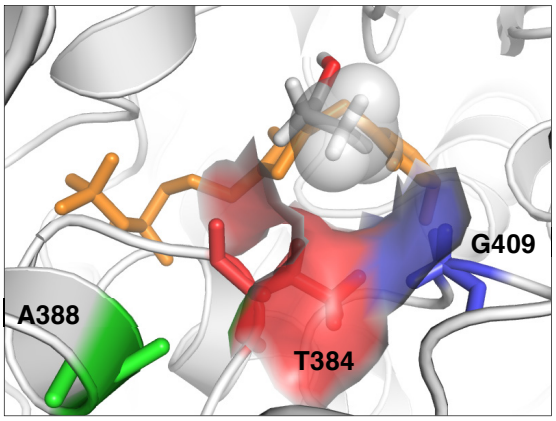


Figure 33: The donor-binding site of *Ap*PDC-W388A with propanal (gray) bound to ThDP (orange). The backside of the donor is mainly blocked by T384 (red). Additionally, the protein backbone of G409 (blue) limits the possible space. Moreover, the side chain of propanal (shown as sphere model) seems to block the access to the donor backside as well. The structure of *Ap*PDC (pdb: 2VBI) was equilibrated with ThDP-bound propanal for 0.5 ns in a water box with 2 Å distance to the protein using the YASARA software (YASARA Biosciences GmbH, Austria) and the standard energy minimization protocol (Krieger *et al*., 2002).

*The structure was kindly provided by C. Vogel (Institute of Technical Biochemistry, University of Stuttgart) and represents the wild-type structure of *Ap*PDC (pdb: 2VBI, Rother *et al*., 2011) including the substitution of W388 to alanine (introduced *in silico*).

A potential alternative *S*-pathway could not be proven by the actual data set. According to the available structural information, a possible arrangement of the acceptor at the backside of the donor is rather unlikely. Thus, the main question, why *Ap*PDC-W388A is less *R*-selective for the formation of PAC derivatives in the absence of a suitable *S*-pocket, still remains open.

3.7 Thiamine enzyme hybrids

The availability of *R*- and *S*-selective variants for enantiocomplementary syntheses is of major importance for enzyme toolboxes to become attractive tools in the chemical industry (chapter 1.5, Meyer *et al*., 2013). Based on the introduction of the *S*-pocket concept, *S*-pocket engineering facilitated the access to various (*S*)-α-hydroxy ketones with high *ee*s in the predominantly *R*-selective DC superfamily of ThDP-dependent enzymes, enabling the enantiocomplementary asymmetric synthesis of various PAC and 2-HPP derivatives (Gocke *et al*., 2008; Rother *et al*., 2011, **Publications II–IV**). However, the range of possible *S*-selective carboligations is still limited compared to *R*-selective reactions, due to the steric and chemical properties of the active sites in ThDP-dependent enzymes. *Inter alia*, no enzymes have been identified so far to catalyze the direct asymmetric (*S*)-benzoin* synthesis starting from benzaldehyde. Here, the potential alternative *S*-pathway might have closed this gap, however, this route to (*S*)-α-hydroxy ketones could not be proven so far (chapter 3.6).

The reasons for the limitations of ThDP-dependent enzyme-catalyzed (*S*)-benzoin synthesis are versatile (for detailed information see **Publication V**). The main problem is that the formation of (*S*)-benzoin requires both, a large and suitable donor-binding site for the benzaldehyde donor and a suitable *S*-pocket for the antiparallel-oriented benzaldehyde acceptor (Figure 34). Unfortunately, the combination of these structural features is not yet available among the known set of ThDP-dependent enzymes (Figure 23). Therefore, a novel strategy was followed in this thesis in order to get access to (*S*)-benzoin: the creation of so-called “thiamine enzyme hybrids”. Thiamine enzyme hybrids combine important structural features of different ThDP-dependent enzymes to enable new synthetic prospects. Figure 34 exemplarily demonstrates the combination (hybridization) of the large donor-binding site of *Pf*BAL and the large *S*-pocket of *Ap*PDC-E469G to get access to (*S*)-benzoin. In this case,

*Other routes to (*S*)-benzoin are available, *e.g*. kinetic resolution or asymmetric reduction (for more details see **Publication V**). However, they all share the same drawback that the starting material, if not commercially available, must be chemically synthesized.

two approaches are generally conceivable to combine the required features: (i) the introduction of the large *S*-pocket of *Ap*PDC into *Pf*BAL, or (ii) the extension of the donor-binding site of *Ap*PDC-E469G in order to mimic *Pf*BAL.

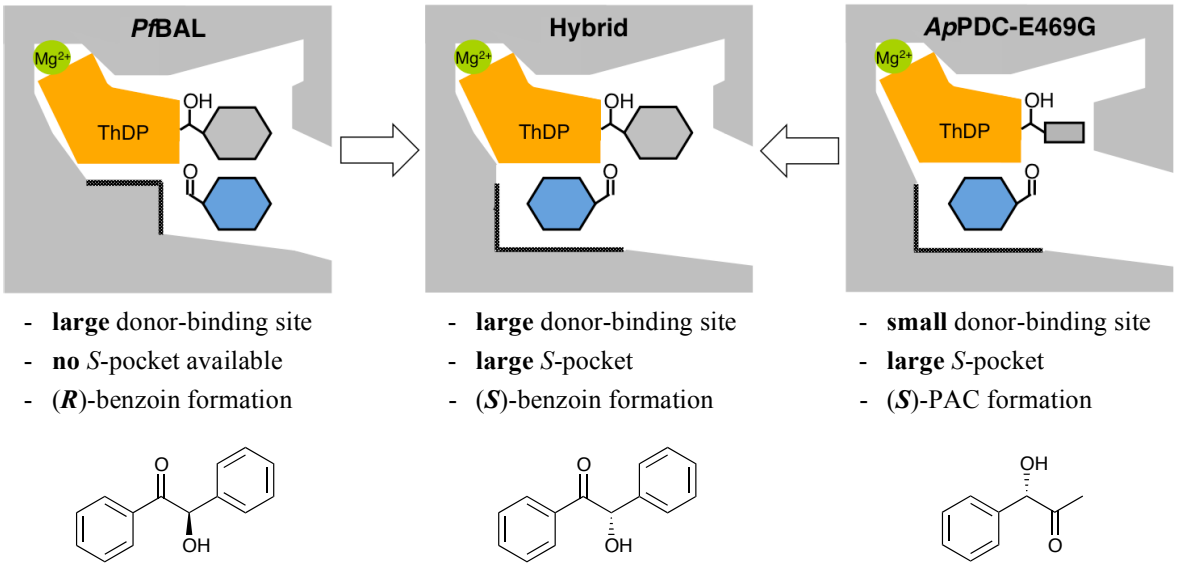


Figure 34: The combination of the features of *Pf*BAL and *Ap*PDC-E469G to a novel thiamine enzyme hybrid for the asymmetric synthesis of (*S*)-benzoin starting from benzaldehyde.

3.7.1 Remodeling the *S*-pocket region of *Pf*BAL*

The major limitation of *S*-selective carboligation is the requirement of a suitable *S*-pocket for the acceptor substrate. The protein backbone of the *S*-pocket α-helix restricts the potential *S*-pocket size and thus the acceptor substrate range for *S*-selective carboligation. In case of *Pf*BAL, no *S*-pocket can be found in the wild-type enzyme. A structural alignment of *Ap*PDC and *Pf*BAL revealed the differences of the *S*-pocket regions (Figure 35). The protein backbone of the respective α-helix in *Pf*BAL is situated in close distance to the ThDP cofactor, which prevents mutagenesis in this area in order to increase the space. Further, the α-helix in *Pf*BAL is four residues longer compared to *Ap*PDC, and *Pf*BAL has a larger loop following the α-helix (Figure 35 A). This extension might cause a small shift of the helix towards ThDP (relative to *Ap*PDC), removing the space for an *S*-pocket completely. The transfer of important structural properties of the *Ap*PDC *S*-pocket region into *Pf*BAL might help to overcome this problem. Based on a sequence alignment, seven amino acids were deleted to shorten the respective loop in *Pf*BAL and thus to introduce a backbone shift. In addition, four amino acids were substituted by the corresponding amino acid in *A*pPDC-E469G (Figure 35 B). These substitutions should help to adapt the respective α-helix in

*This project was performed in cooperation with C. Vogel (Institute of Technical Biochemistry, University of Stuttgart).

*Pf*BAL to that of *Ap*PDC and thus create a potential *S*-pocket in BAL. Specifically, the substitution of T381 (477) to glycine, which corresponds to E469G (477) in *Ap*PDC, should open the *S*-pocket. The resulting BAL/PDC hybrid was modeled based on the crystal structure of *Pf*BAL (pdb: 2AG0, Mosbacher *et al.*, 2005) using SWISS-MODEL (swissmodel.expasy.org). The deduced homology model revealed a new structure of the *S*-pocket region with a shortened *S*-pocket α-helix compared to *Pf*BAL. In addition, the new *S*-pocket region highly resembled the one of *Ap*PDC (Figure 35 D).

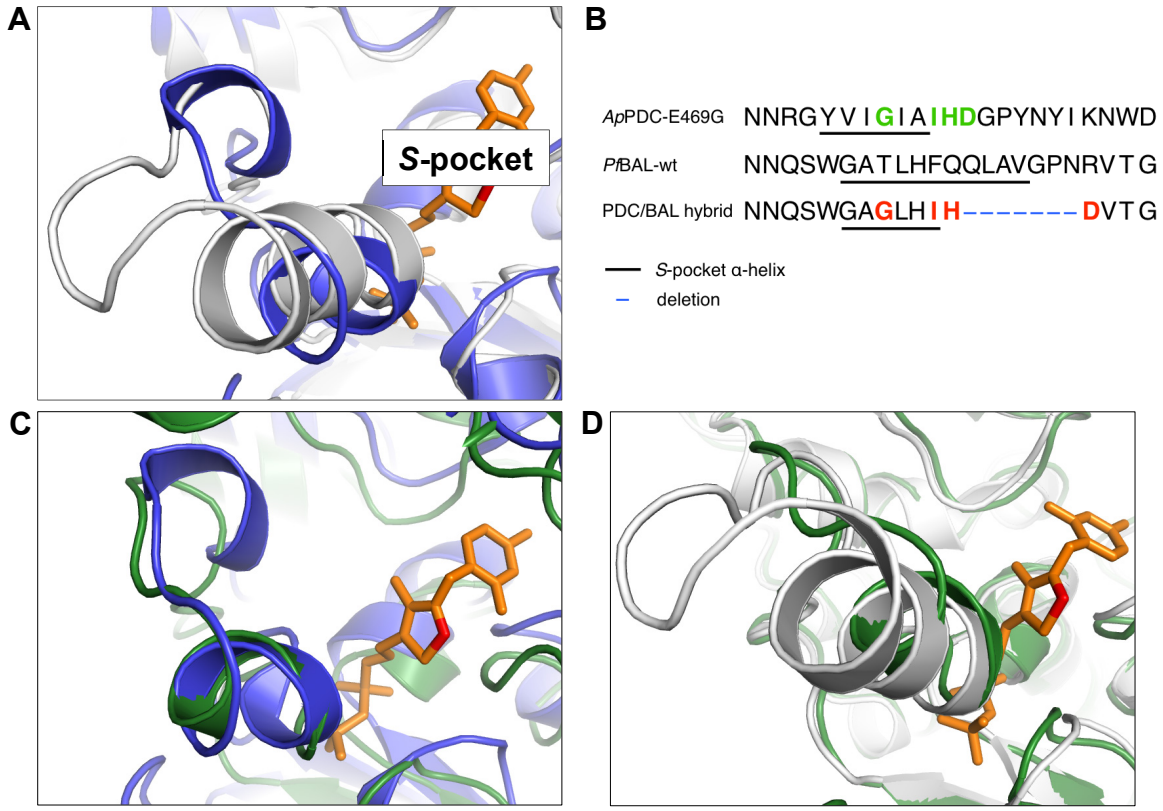


Figure 35: Comparison of the *S*-pocket regions of *Ap*PDC and *Pf*BAL (**A**, **B**) and modeling of a potential BAL/PDC hybrid enzyme (**C**, **D**). **A:** Structural alignment of *Ap*PDC (blue) and *Pf*BAL (white) showing the differences of the *S*-pocket regions: a longer *S*-pocket α-helix and an extended loop in *Pf*BAL compared to *Ap*PDC. **B:** Based on sequence alignment, four amino acids (red) were substituted by corresponding amino acids in *Ap*PDC-E469G (green) and seven amino acids were deleted in *Pf*BAL (for complete sequence see Appendix 5.3) in order to induce structural changes in its *S*-pocket region. **C:** The homology model of the new BAL/PDC hybrid (green) revealed a shortened *S*-pocket α-helix, which highly resembles the *S*-pocket region of *Ap*PDC (blue). **D:** MD simulations revealed that the *S*-pocket α-helix of the new BAL/PDC hybrid (green) slightly shifted towards the ThDP cofactor and thus blocks the *S*-pocket. ThDP, exemplarily shown for *Ap*PDC, is colored in orange and its catalytically important C2 atom in red.

The gene of the new *Pf*BAL variant was synthesized by GeneArt® Gene Synthesis (life technologies™, Carlsbad, USA), cloned into pET19b (Novagen, Merck Millipore, Darmstadt, Germany), and finally overexpressed in *E. coli* BL21 (DE3) (Novagen, Merck Millipore, Darmstadt, Germany). In spite of the high number of deletions and mutations, the new variant was overexpressed as soluble protein and, remarkably, still exhibited weak carboligation

activity (conversions < 1 % within 20 h). However, the stereoselectivity was not altered and the reaction of benzaldehyde and acetaldehyde resulted in the formation of (*R*)-2-HPP and (*R*)-benzoin (both with > 99 % *ee*), respectively. Later on, MD simulations* revealed the reason for the unaffected stereoselectivity: the *S*-pocket α-helix was indeed slightly shifted as proposed before, but unfortunately into the wrong direction towards the ThDP-cofactor (Figure 35 D), which blocked the *S*-pocket. Although this approach was yet not successful, it indicated the potential of this novel strategy. ThDP-dependent enzymes seem to be very robust towards mutations close to the active site and also to multiple mutations. Here, targeted variations of the *S*-pocket α-helix (and its position) seem to be possible, as was shown by the modeling studies, benefiting from the position of the *S*-pocket α-helix close to the protein surface. Meanwhile, other *Pf*BAL variants# were planned *in silico* and tested in MD simulations that revealed a shift of the *S*-pocket α-helix away from the ThDP cofactor, which in turn could open the *S*-pocket.

The same approach, if successful, might be also applied for other ThDP-dependent enzymes in order to overcome the limitations of *S*-selective carboligation, *e.g.* for *Ec*AHAS. Since *S*-selective *Ec*AHAS-I variants were not accessible *via* the conventional *S*-pocket engineering (chapter 3.4.1), the recycling cascade for the synthesis of nor(pseudo)ephedrine is yet limited to the production of (1*R*,2*R*)-norpseudoephedrine and (1*R*,2*S*)-norephedrine (chapter 3.3.2, Figure 24, **Publication I**). At this point, the adaption of the *Ec*AHAS-I *S*-pocket α-helix to that of *Ap*PDC might be a promising strategy to overcome the existing limitation and thus, give access to the missing (1*S*,2*R*)- and (1*S*,2*S*)-nor(pseudo)ephedrine.

3.7.2 Hybridisation of the active sites of *Ap*PDC-E469G and *Pf*BAL

The second approach to achieve *S*-selective benzoin formation aimed for the transfer of the large donor-binding site of *Pf*BAL either into *S*-selective *Ap*PDC-E469G or MenD variants, respectively, which both have large *S*-pockets. At this point, *Ap*PDC-E469G was chosen as a starting point, because wild-type *Ap*PDC showed already low benzoin-forming activity in the presence of benzaldehyde as sole substrate (Rother *et al*., 2011), whereas MenD did not accept benzaldehyde as donor due to the highly polar character of the donor-binding site (chapter 3.3.2).

*Performed by C. Vogel (Institute of Technical Biochemistry, University of Stuttgart) using GROMACS (Version 4.5.4, Pronk *et al*., 2013). The variant was parameterized in the AMBER force field and embedded in a water box with 30 Å distance to the protein containing Na^+ counterions for a balanced charge. After steepest decent energy minimization, the variants was equilibrated for 1 ns and subsequently simulated for 30 ns in triplicates.

#These variants are not discussed in the context of this thesis, because experimental data were not generated so far.

The combination of a comparison of the donor-binding sites of *Pf*BAL and *Ap*PDC (Table 4, Figure 23), a comprehensive analysis of the amino acid distribution in 43 sequences homologous to *Pf*BAL and 186 sequences homologous to *Ap*PDC, and modeling studies with ThDP-bound benzaldehyde, enabled the identification of the pivotal position that directly influences the chemoselectivity of *Ap*PDC (for detailed analysis see **Publication V**). In *Ap*PDC, T384 (388) limits the space for benzaldehyde in the donor-binding site (Figure 36 A), whereas in all BAL sequences standard position 388 is occupied by glycine. Indeed, mutation of T384 to glycine opened the donor-binding site in *Ap*PDC-E469G (Figure 36 B) and completely shifted its chemoselectivity. The new variant *Ap*PDC-E469G/T384G revealed typical BAL activity, catalyzing the formation of benzoin and 2-HPP as main products from benzaldehyde as donor and acetaldehyde as acceptor, whereas acetoin and PAC (from acetaldehyde as donor) were obtained only in traces (for an overview about possible carboligation products see chapter 1.7.4, Figure 16). Acetoin and PAC in turn were the only observed products of *Ap*PDC-E469G (Gerhards *et al*., 2012, **Publication V**).

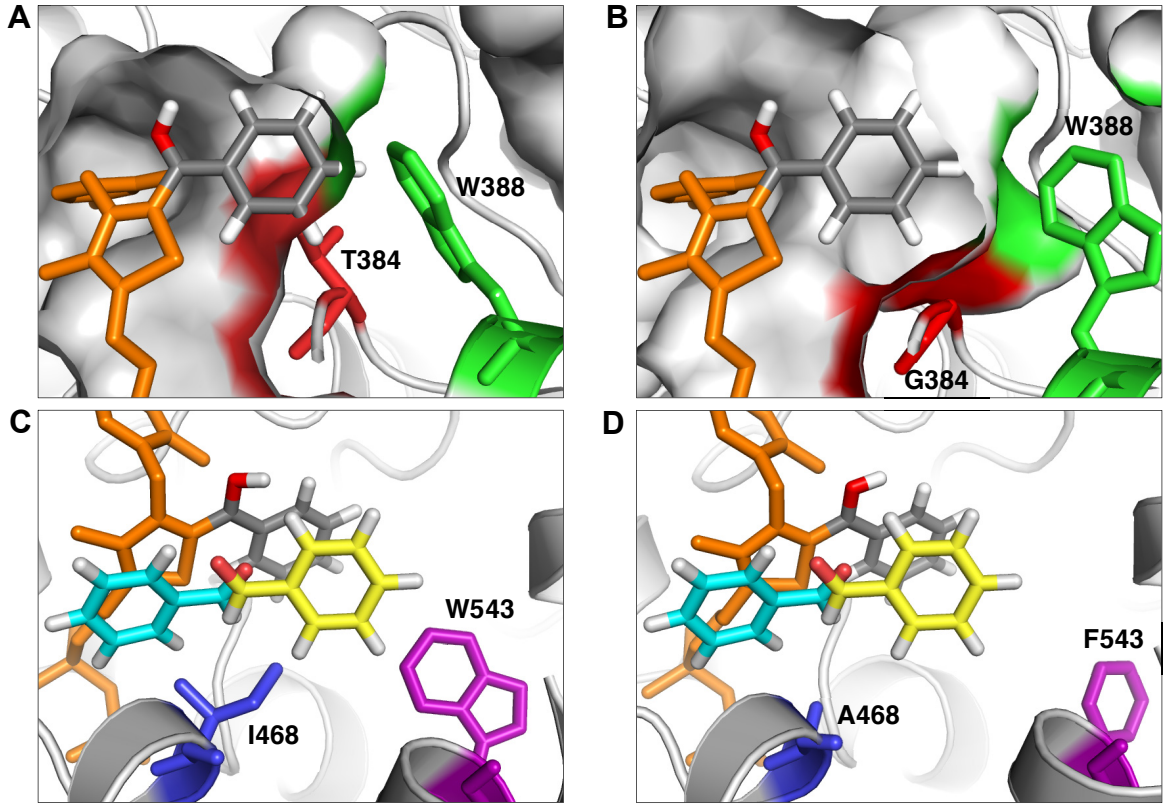


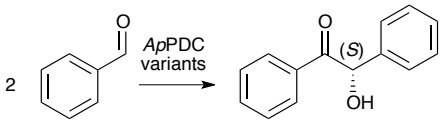
Figure 36: Tailoring the active site of *Ap*PDC-E469G (**Publication V**). **A:** The small donor-binding site of *Ap*PDC-E469G is restricted by T384 (red) and W388 (green), which prevents the binding of benzaldehyde as donor (gray). **B:** The donor-binding site could be opened for benzaldehyde by mutation of T384 to glycine (red). Furthermore, the equilibrated structure of E469G/T384G revealed that W388 (green) does not restrict the donor-binding site for benzaldehyde due to conformational changes, which was supported by experimental data. **C:** The parallel benzaldehyde orientation (yellow) is potentially stabilized in E469G/T384G by I468 (blue) and W543 (purple), which both do not influence the antiparallel orientation (cyan). **D:** Such a stabilization is not possible anymore in E469G/T384G/I468A/W543F due to the increased distance between the residues and benzaldehyde (yellow), resulting in improved *S*-selectivity (95 % *ee*) compared to E469G/T384G (59 % *ee*).

Furthermore, the hybridization approach was also successful with respect to stereoselectivity. The combination of a large donor-binding site and a large *S*-pocket enabled the first direct enzymatic asymmetric synthesis of (*S*)-benzoin (59 % *ee*) with good conversion (Table 11). Even 2-HPP was synthesized highly *S*-selective (91 % *ee*) by *Ap*PDC-E469G/T384G.

Optimization of S-selectivity

In a next step, the moderate stereoselectivity of *Ap*PDC-E469G/T384G should be increased by further rational design. As demonstrated in chapter 3.5.3, the selective destabilization of the *R*-pathway in *Ec*MenD variants resulted in improved *S*-selectivity. This approach was now also applied to improve the *S*-selectivity of the *Ap*PDC-variant (**Publication V**). Two positions in the active site of *Ap*PDC-E469G/T384G were identified that most probably stabilize the parallel orientation of acceptor benzaldehyde, but do not influence the binding of antiparallel-oriented benzaldehyde: I468 at the *S*-pocket entrance and W543 as part of the C-terminal α-helix (Figure 36 C). In both cases, mutation to smaller amino acids prevented stabilizing interactions of the residue and the parallel benzaldehyde orientation (Figure 36 D), and thus drastically enhanced *S*-selectivity. The triple variant E469G/T384G/I468A catalyzed the formation of (*S*)-benzoin with 87 % *ee*, accompanied by high conversion under the tested conditions (95 % within 6 h, Table 11).

Table 11: Formation of benzoin catalyzed by *Ap*PDC variants (**Publication V**).[1]

*Ap*PDC variant	Conversion [%][2]	*ee* [%][3]
		
E469G	< 1	n.d.[4]
E469G/T384G	52	59 (*S*)
E469G/T384G/I468G	23	66 (*S*)
E469G/T384G/I468A	95	87 (*S*)
E469G/T384G/I468V	40	76 (*S*)
E469G/T384G/I468A/W543F	36	95 (*S*)

[1]Reaction conditions: 50 mM TEA buffer, pH 8.0, containing 2 mM $MgSO_4$ and 0.1 mM ThDP, 1 mg mL^{-1} purified enzyme; 18 mM benzaldehyde or benzaldehyde derivative; 20 °C for 6 h.
[2]Conversion determined by HPLC based on benzaldehyde consumption.
[3]Determined by chiral-phase HPLC.
[4]n.d.: not determined

The already high *S*-selectivity could be even further improved by introduction of the fourth mutation in position 543. The new variant E469G/T384G/I468A/W543F was highly *S*-

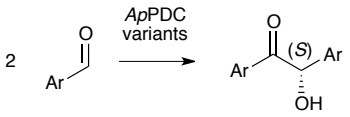
selective for the formation of (*S*)-benzoin (95 % *ee*) with a still reasonable conversion of 36 %, which again highlights the robustness of ThDP-dependent enzymes with respect to active site mutations. Furthermore, *S*-selectivity was also enhanced for the synthesis of 2-HPP (> 99 % *ee*) starting from acetaldehyde and benzaldehyde.

The results impressively confirm the high potential of suppressing the *R*-pathway in order to enhance *S*-selectivity. This concept, initially developed for *Ec*MenD (chapter 3.5.3), seems to be generally applicable for *S*-selective ThDP-dependent enzymes. Thus, the combination of *S*-pocket engineering and selective destabilization of the *R*-pathway paves the way for highly *S*-selective asymmetric C–C bond formations that start from readily available substrates.

Access to various stereochemically pure (S)-benzoins

As respective studies with MenD variants revealed *meta*-substituted benzaldehydes as optimal acceptor substrates for the *S*-pocket (chapter 3.4.2), the new *Ap*PDC variants were also tested with differently substituted benzaldehyde derivatives for the formation of respective benzoin derivatives (**Publication V**). Although the benzaldehyde derivative was now donor and acceptor at once, the trends were similar to the results obtained with MenD: *ortho*- and *para*-substituted benzaldehydes were converted either with lower *S*-selectivity compared to benzaldehyde, or *R*-selectively (not shown), whereas *meta*-substituted benzaldehydes were transformed highly *S*-selectively (Table 12).

Table 12: Formation of benzoin and benzoin derivatives catalyzed by *Ap*PDC variants (**Publication V**).[1]

Ar	E469G/T384G/I468A Conversion [%][3]	*ee* [%][4]	E469G/T384G/ I468A/W543F[2] Conversion [%]	*ee* [%]
		2  *Ap*PDC variants		
C_6H_5	92	89 (*S*)	26	98 (*S*)
3-FC_6H_4	80	87 (*S*)	36	93 (*S*)
3-ClC_6H_4	97	91 (*S*)	48	> 99 (*S*)
3-BrC_6H_4	85	95 (*S*)	30	> 99 (*S*)
3-IC_6H_4	30	96 (*S*)	11	> 99 (*S*)
3-$MeOC_6H_4$	93	98 (*S*)	58	> 99 (*S*)

[1]Reaction conditions: 50 mM TEA buffer, pH 8.0, containing 2 mM $MgSO_4$ and 0.1 mM ThDP, 1 mg mL^{-1} purified enzyme; 18 mM benzaldehyde or benzaldehyde derivative; 15 °C for 6 h.
[2]Obtained through site-saturation mutagenesis at position 543 using E469G/T384G/I468A as template (**Publication V**). Mutagenesis and screening were performed by D. Hahn and U. Mackfeld (IBG-1, Forschungszentrum Jülich GmbH, Germany).
[3]Determined by HPLC based on benzaldehyde consumption.
[4]Determined by chiral-phase HPLC.

Using *Ap*PDC variant E469G/T384G/ I468A/W543F, stereochemically pure (*S*)-benzoins (*ee* > 99 %) were obtained from all tested *meta*-substituted substrates, except for 3-fluorobenzaldehyde (93 % *ee*) under optimized conditions at reduced temperature (for further details see **Publication V**). As expected, variant E469G/T384G/I468A revealed lower *S*-selectivity but higher conversions relative to E469G/T384G/I468A/W543F (Table 12). Thus, the results confirm the findings from the MenD studies (**Publications II–IV**), but the reasons for the preference of *meta*-substituted benzaldehyde in the *S*-selective carboligation are still not clear and have to be further investigated.

Despite the exceptional catalytic potential of ThDP-dependent enzymes in asymmetric C–C bond formations and their exploration for almost 100 years now, *S*-selective benzoin synthesis starting from commercially available benzaldehydes as sole substrates was yet not possible. This long-standing problem was solved by the design of a hybrid substrate-binding site. Thereby, the platform of α-hydroxy ketones was complemented by (*S*)-benzoin and *meta*-substituted derivatives thereof with up to > 99 % *ee*. This example demonstrates the importance and potential of the detailed understanding of structure-function relationships in order to tailor ThDP-dependent enzymes for the desired activity, and again highlights their robustness with respect to active site mutations. Furthermore, this hybridization strategy might offer new perspectives on thiamine catalysis with respect to the combination of selectivity-determining modules of different ThDP-dependent enzymes to access completely new activities.

4 CONCLUSIONS AND FUTURE PERSPECTIVES

This thesis offers a new perspective on the rational design of highly *S*-selective ThDP-dependent enzymes for the asymmetric synthesis of α-hydroxy ketones. Based on a profound understanding of the principles of chemo- and stereoselectivity, new strategies were developed to tailor ThDP-dependent enzymes for *S*-selective carboligation. These strategies are not restricted only to *S*-pocket engineering anymore and thus allow overcoming the main limitations of moderate *S*-selectivity as well as limited product ranges.

As demonstrated for *S*-selective *Ec*MenD and *Ap*PDC variants, selective destabilization of the *R*-pathway in addition to basic *S*-pocket engineering provides a powerful strategy to improve *S*-selectivity and to access different (*S*)-α-hydroxy ketones with high stereoselectivity (*ee* 95–99%). There are now several potential setscrews in the active site to adjust *S*-selectivity.

The new concept of "thiamine enzyme hybrids" addresses the range of *S*-selective carboligations, which is limited by steric and chemical properties of the active site of the enzymes. The high potential of this approach could be demonstrated by design of a hybrid substrate-binding site of *Ap*PDC and *Pf*BAL, which solved the long-standing problem of *S*-selective benzoin formation starting from benzaldehyde and derivatives thereof. This combinatorial assembly of selectivity-determining modules might pave the way for the design of ThDP-dependent enzymes with new catalytic activities (selectivities).

Four new wild-type enzymes and 35 variants with characterized carboligation activity complement the toolbox of ThDP-dependent enzymes, which expand the platform of α-hydroxy ketones by novel functionalized mixed araliphatic (*S*)-α-hydroxy ketones as well as (*S*)-benzoins with excellent *ee*s. The results in this thesis highlight the tremendous catalytic potential of ThDP-dependent enzymes, their robustness with respect to active site mutations, as well as their applicability in valuable cascade reactions, which might enhance the attractiveness of the toolbox of ThDP-dependent enzymes for industrial application.

To expand the scope and utility of ThDP-dependent enzymes, the substrate spectrum should be expanded. Recently, the first enzyme accepting ketones as acceptors was identified, giving access to novel *tert*-α-hydroxy ketones. Further, careful optimization of the reaction conditions has to be considered as an additional tool to control chemo- and stereoselectivity and in order to make full use of the enzyme potential.

5 APPENDIX

5.1 Catalytic subunits of *Ec*AHAS-I and *Ec*AHAS-II (chapter 3.3.1)

For crystallizations studies, the catalytic subunits (CSU) of *Ec*AHAS-I and *Ec*AHAS-II were separately synthesized by GeneArt® Gene Synthesis (life technologies™, Carlsbad, USA) and provided in a pMA-RQ vector (start and end of the genes are highlighted in bold red).

EcAHAS-I_CSU in pMA-RQ

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAAT
CGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGGCCGCTACAGGGCGCTCCCATTCGCCATTCAG
GCTGCGCAACTGTTGGGAAGGGCGTTTCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAG
GCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGACGTAATACGACTC
ACTATAGGGCGAATTGGCGGAAGGCCGTCAAGGCCGCATTCTAGATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGG
CCTG**GCAAGT**TCGGGCACAACATCGACGCGTAAGCGCTTTACCGGCGCAGAATTTATCGTTCATTTCCTGGAACAGCAGGGC
ATTAAGATTGTGACAGGCATTCCGGGCGGTTCTATCCTGCCTGTTTACGATGCCTTAAGCCAAAGCACGCAAATCCGCCATAT
TCTGGCCCGTCATGAACAGGGCGCGGGCTTTATCGCTCAGGGAATGGCGCGCACCGACGGTAAACCGGCGGTCTGTATGGCC
TGTAGCGGACCGGGTGCGACTAACCTGGTGACCGCCATTGCCGATGCGCGGCTGGACTCCATCCCGCTGATTTGCATCACTGG
TCAGGTTCCCGCCTCGATGATCGGCACCGACGCCTTCCAGGAAGTGGACACCTACGGCATCTCTATCCCCATCACCAAACACA
ACTATCTGGTCAGACATATCGAAGAACTCCCGCAGGTCATGAGCGATGCCTTCCGCATTGCGCAATCAGGCCGCCCAGGCCC
GGTGTGGATAGACATTCCTAAGGATGTGCAAACGGCAGTTTTTGAGATTGAAACACAGCCCGCTATGGCAGAAAAAGCCGCC
GCCCCCGCCTTTAGCGAAGAAAGCATTCGTGACGCAGCGGCGATGATTAACGCTGCCAAACGCCCGGTGCTTTATCTGGGCG
GCGGTGTGATCAATGCGCCCGCACGGGTGCGTGAACTGGCGGAGAAAGCGCAACTGCCTACCACCATGACTTTAATGGCGCT
GGGCATGTTGCCAAAAGCGCATCCGTTGTCGCTGGGTATGCTGGGGATGCACGGCGTGCGCAGCACCAACTATATTTTGCAG
GAGGCGGATTTGTTGATAGTGCTCGGTGCGCGTTTTGATGACCGGGCGATTGGCAAAACCGAGCAGTTCTGTCCGAATGCCAA
AATCATTCATGTCGATATCGACCGTGCAGAGCTGGGTAAAATCAAGCAGCCGCACGTGGCGATTCAGGCGGATGTTGATGAC
GTGCTGGCGCAGTTGATCCCGCTGGTGGAAGCGCAACCGCGTGCAGAGTGGCACCAGTTGGTAGCGGATTTGCAGCGTGAGT
TTCCGTGTCCAATCCCGAAAGCGTGCGATCCGTTAAGCCATTACGGCCTGATCAACGCCGTTGCCGCCTGTGTCGATGACAAT
GCAATTATCACCACCGACGTTGGTCAGCATCAGATGTGGACCGCGCAAGCTTATCCGCTCAATCGCCCACGCCAGTGGCTGAC
CTCCGGTGGGCTGGGCACGATGGGTTTTGGCCTGCCTGCGGCGATTGGCGCTGCGCTGGCGAACCCGGATCGCAAAGTGTTGT
GTTTCTCCGGCGACGGCAGCCTGATGATGAATATTCAGGAGATGGCGACCGCCAGTGAAAATCAGCTGGATGTCAAAATCAT
TCTGATGAACAACGAAGCGCTGGGGCTGGTGCATCAGCAACAGAGTCTGTTCTACGAGCAAGGCGTTTTTGCCGCCACCTATC
CGGGCAAAATCAACTTTATGCAGATTGCCGCCGGATTCGGCCTCGAAACCTGTGATTTGAATAACGAAGCCGATCCGCAGGC
TTCATTGCAGGAAATCATCAATCGCCCTGGCCCGGCGCTGATCCATGTGCGCATTGATGCCGAAGAAAAAGTTTACCCGATGG
TGCCGCCAGGTGCGGCGAATACTGAAATGGTGGGG**GAATAA**GGATCCCTGGGCCTCATGGGCCTTCCGCTCACTGCCCGCTT
TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAACATGGTCATAGCTGTTTCCTTGCGTATTGGGCGCTCTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTCGTTCGGGTAAAGCCTGGGGTGCCTAATGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT
GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG
CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG
TAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTT
GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCA
GTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCA
GCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC
TCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTAT
TTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCA
ATGATACCGCGAGAACCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT
GGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTG
CGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGA
TCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTT
GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGAC
TGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA
CCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG
TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCA
AAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAAT
ATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCAC

EcAHAS-II_CSU in pMA-RQ

CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAAT
CGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGGCCGCTACAGGGCGCTCCCATTCGCCATTCAG
GCTGCGCAACTGTTGGGAAGGGCGTTTCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAG
GCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGACGTAATACGACTC
ACTATAGGGCGAATTGGCGGAAGGCCGTCAAGGCCGCATTCTAGATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGG
CCTG**AATGGC**GCACAGTGGGTGGTACATGCGTTGCGGGCACAGGGTGTGAACACCGTTTTCGGTTATCCGGGTGGCGCAATT
ATGCCGGTTTACGATGCATTGTATGACGGCGGCGTGGAGCACTTGCTATGCCGACATGAGCAGGGTGCGGCAATGGCGGCTA
TCGGTTATGCTCGTGCTACCGGCAAAACTGGCGTATGTATCGCCACGTCTGGTCCGGGCGCAACCAACCTGATAACCGGGCTT
GCGGACGCACTGTTAGATTCCATCCCTGTTGTTGCCATCACCGGTCAAGTGTCCGCACCGTTTATCGGCACTGACGCATTTCAG
GAAGTGGATGTCCTGGGATTGTCGTTAGCCTGTACCAAGCACAGCTTTCTGGTGCAGTCGCTGGAAGAGTTGCCGCGCATCAT
GGCTGAAGCATTCGACGTTGCCTGCTCAGGTCGTCCTGGTCCGGTTCTGGTCGATATCCCAAAAGATATCCAGTTAGCCAGCG
GTGACCTGGAACCGTGGTTCACCACCGTTGAAAACGAAGTGACTTTCCCACATGCCGAAGTTGAGCAAGCGCGCCAGATGCT
GGCAAAAGCGCAAAAACCGATGCTGTACGTTGGCGGTGGCGTGGGTATGGCGCAGGCAGTTCCGGCTTTGCGTGAATTTCTC
GCTGCCACAAAAATGCCTGCCACCTGTACGCTGAAAGGGCTGGGCGCAGTAGAAGCAGATTATCCGTACTATCTGGGCATGC
TGGGGATGCACGGCACCAAAGCGGCAAACTTCGCGGTGCAGGAGTGTGACCTGCTGATCGCCGTGGGCGCACGTTTTGATGA
CCGGGTGACCGGCAAACTGAACACCTTCGCGCCACACGCCAGTGTTATCCATATGGATATCGACCCGGCAGAAATGAACAAG
CTGCGTCAGGCACATGTGGCATTACAAGGTGATTTAAATGCTCTGTTACCAGCATTACAGCAGCCGTTAAATCAATATGACTG
GCAGCAACACTGCGCGCAGCTGCGTGATGAACATTCCTGGCGTTACGACCATCCCGGTGACGCTATCTACGCGCCGTTGTTGT
TAAAACAACTGTCGGATCGTAAACCTGCGGATTGCGTCGTGACCACAGATGTGGGGCAGCACCAGATGTGGGCTGCGCAGCA
CATCGCCCACACTCGCCCGGAAAATTTCATCACCTCCAGCGGTTTAGGTACCATGGGTTTTGGTTTACCGGCGGCGGTTGGCG
CACAAGTCGCGCGACCGAACGATACCGTTGTCTGTATCTCCGGTGACGGCTCTTTCATGATGAATGTGCAAGAGCTGGGCACC
GTAAAACGCAAGCAGTTACCGTTGAAAATCGTCTTACTCGATAACCAACGGTTAGGGATGGTTCGACAATGGCAGCAACTGT
TTTTTCAGGAACGATACAGCGAAACCACCCTTACTGATAACCCCGATTTCCTCATGTTAGCCAGCGCCTTCGGCATCCATGGC
CAACACATCACCCGGAAAGACCAGGTTGAAGCGGCACTCGACACCATGCTGAACAGTGATGGGCCATACCTGCTTCATGTCT
CAATCGACGAACTTGAGAACGTCTGGCCGCTGGTGCCGCCTGGCGCCAGTAATTCAGAAATGTTGGAGAAATTA**TCATGA**GG
ATCCCTGGGCCTCATGGGCCTTCCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAACATGGTCAT
AGCTGTTTCCTTGCGTATTGGGCGCTCTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGGTAAAGCCTGGGGTG
CCTAATGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCC
TGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCT
GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCG
CTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGC
CACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA
AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT
CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTA
CGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAACCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCG
TCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGT
TAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC
TCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCG
ACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAA
CGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATC
TTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCG
ACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC
ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCAC

Both genes were subcloned by Dr. J. Kulig (IBG-1, Forschungszentrum Jülich GmbH, Germany) into the plasmid pET-YSBLIC3C using the following forward (fw) and reverse (rv) primers:

AHAS-I_fw:	5'-CCAGGGACCAGCAATGGCAAGTTCGGGCACAACATCGACG-3'
AHAS-I_rv:	5'-GAGGAGAAGGCGCGTTATTATTCCCCCACCATTTCAGTATTCGCCGC-3'
AHAS-II_fw:	5'-CCAGGGACCAGCAATGAATGGCGCACAGTGGGTGGTACATG-3'
AHAS-II_rv:	5'-GAGGAGAAGGCGCGTTATCATGATAATTTCTCCAACATTTCTGAATTAC-3'.

Resulting constructs were applied for transformation of *E. coli* (BL21 (DE3) or B834 (DE3)) and for subsequent overexpression of the respective CSU for crystallization.

5.2 *Pf*BAL and primer sequences (chapter 3.6)

Wild-type (wt) *Pf*BAL gene was cloned from pKK233-2 into pET28a by D. Hahn (IBG-1, Forschungszentrum Jülich GmbH, Germany).

PfBAL wild-type sequence

ATGGCGATGATTACAGGCGGCGAACTGGTTGTTCGCACCCTAATAAAGGCTGGGGTCGAACATCTGTTCGGCCTGCACGGCG
CGCATATCGATACGATTTTTCAAGCCTGTCTCGATCATGATGTGCCGATCATCGACACCCGCCATGAGGCCGCCGCAGGGCAT
GCGGCCGAGGGCTATGCCCGCGCTGGCGCCAAGCTGGGCGTGGCGCTGGTCACGGCGGGCGGGGGATTTACCAATGCGGTCA
CGCCCATTGCCAACGCTTGGCTGGATCGCACGCCGGTGCTCTTCCTCACCGGATCGGGCGCGCTGCGTGATGATGAAACCAAC
ACGTTGCAGGCGGGGATTGATCAGGTCGCGATGGCGGCGCCCATTACCAAATGGGCGCATCGGGTGATGGCAACCGAGCATA
TCCCACGGCTGGTGATGCAGGCGATCCGCGCCGCGTTGAGCGCGCCACGCGGGCCGGTGTTGCTGGATCTGCCGTGGGATATT
CTGATGAACCAGATTGATGAGGATAGCGTCATTATCCCCGATCTGGTCTTGTCCGCGCATGGGGCCAGACCCGACCCTGCCGA
TCTGGATCAGGCTCTCGCGCTTTTGCGCAAGGCGGAGCGGCCGGTCATCGTGCTCGGCTCAGAAGCCTCGCGGACAGCGCGC
AAGACGGCGCTTAGCGCCTTCGTGGCGGCGACTGGCGTGCCGGTGTTTGCCGATTATGAAGGGCTAAGCATGCTCTCGGGGCT
GCCCGATGCTATGCGGGGCGGGCTGGTGCAAAACCTCTATTCTTTTGCCAAAGCCGATGCCGCGCCAGATCTCGTGCTGATGC
TGGGGGCGCGCTTTGGCCTTAACACCGGGCATGGATCTGGGCAGTTGATCCCCCATAGCGCGCAGGTCATTCAGGTCGACCCT
GATGCCTGCGAGCTGGGACGCCTGCAGGGCATCGCTCTGGGCATTGTGGCCGATGTGGGTGGGACCATCGAGGCTTTGGCGC
AGGCCACCGCGCAAGATGCGGCTTGGCCGGATCGCGGCGACTGGTGCGCCAAAGTGACGGATCTGGCGCAAGAGCGCTATGC
CAGCATCGCTGCGAAATCGAGCAGCGAGCATGCGCTCCACCCCTTTCACGCCTCGCAGGTCATTGCCAAACACGTCGATGCA
GGGGTGACGGTGGTAGCGGATGGTGCGCTGACCTATCTCTGGCTGTCCGAAGTGATGAGCCGCGTGAAACCCGGCGGTTTTCT
CTGCCACGGCTATCTAGGCTCGATGGGCGTGGGCTTCGGCACGGCGCTGGGCGCGCAAGTGGCCGATCTTGAAGCAGGCCGC
CGCACGATCCTTGTGACCGGCGATGGCTCGGTGGGCTATAGCATCGGTGAATTTGATACGCTGGTGCGCAAACAATTGCCGCT
GATCGTCATCATCATGAACAACCAAAGCTGGGGGGCGACATTGCATTTCCAGCAATTGGCCGTCGGCCCCAATCGCGTGACG
GGCACCCGTTTGGAAAATGGCTCCTATCACGGGGTGGCCGCCGCCTTTGGCGCGGATGGCTATCATGTCGACAGTGTGGAGA
GCTTTTCTGCGGCTCTGGCCCAAGCGCTCGCCCATAATCGCCCCGCCTGCATCAATGTCGCGGTCGCGCTCGATCCGATCCCG
CCCGAAGAACTCATTCTGATCGGCATGGACCCCTTCGCACTCGAGCACCACCACCACCACCACTGA

The following fw and rv primers were applied to generate *Pf*BAL-L112A and Q113A, respectively, by standard the Quikchange® mutagenesis protocol using pET28a_*Pf*BAL-wt as template (the mutated codon is underlined):

BAL-L112A: 5'-GATGAAACCAACACGGCGCAGGCGGGGATTG-3' (fw)

5'-CAATCCCCGCCTGCGCCGTGTTGGTTTCATC-3' (rv)

BAL-Q113A: 5'-GAAACCAACACGTGGGCGGCGGGGATTGATC-3' (fw)

5'-GATCAATCCCCGCCGCCAACGTGTTGGTTTC-3' (rv)

*Pf*BAL-L112A was used as template for the generation of L112A/Q113A and L112A/Q113V, respectively, using the following primers:

BAL-L112A/Q113A: 5'-GAAACCAACACGGCGGCGGCGGGGATTGATC-3' (fw)

5'-GATCAATCCCCGCCGCCGCCGTGTTGGTTTC-3' (rv)

BAL-L112A/Q113V: 5'-GAAACCAACACGGCGGTGGCGGGGATTGATCAG-3' (fw)

5'-CTGATCAATCCCCGCCACCGCCGTGTTGGTTTC-3' (rv)

*Pf*BAL-L112A/Q113A and L112A/Q113V, respectively were used as template for the generation of L112A/Q113A/M421V and L112A/Q113V/M421V, respectively, using fw primer 5'-CTATCTAGGCTCGGTGGGCGTGGGCTTC-3' and rv primer 5'-GAAGCCCACGCCCACCGAGCCTAGATAG-3'. *Pf*BAL-L112A/Q113A/M421V/A480L

was prepared based on L112A/Q113A/M421V using fw primer 5'-CAACCAAAGCTGGGGGCTGACATTGCATTTCCAG-3' and rv primer 5'-CTGGAAATGCAATGTCAGCCCCCAGCTTTGGTTG-3'.

E. coli BL21 (DE3) was transformed with the resulting PCR products and used for overexpression of respective variants. Gene sequences were confirmed by DNA sequencing (LGC Genomics, Berlin).

5.3 *Pf*BAL hybrid sequence (chapter 3.7.1)

The *Pf*BAL/*Ap*PDC hybrid was synthesized by GeneArt® Gene Synthesis (life technologies™, Carlsbad, USA) and provided in pMA-RQ vector. The gene was cloned into pET28a by D. Hahn (IBG-1, Forschungszentrum Jülich GmbH, Germany) and subsequently used for the transformation of *E. coli* BL21 (DE3).

PfBAL/ApPDC hybrid sequence

ATGGCGATGATTACAGGCGGCGAACTGGTTGTTCGCACCCTAATAAAGGCTGGGGTCGAACATCTGTTCGGCCTGCACGGCG
CGCATATCGATACGATTTTTCAAGCCTGTCTCGATCATGATGTGCCGATCATCGACACCCGCCATGAGGCCGCCGCAGGGCAT
GCGGCCGAGGGCTATGCCCGCGCTGGCGCCAAGCTGGGCGTGGCGCTGGTCACGGCGGGCGGGGGATTTACCAATGCGGTCA
CGCCCATTGCCAACGCTTGGCTGGATCGCACGCCGGTGCTCTTCCTCACCGGATCGGGCGCGCTGCGTGATGATGAAACCAAC
ACGTTGCAGGCGGGGATTGATCAGGTCGCGATGGCGGCGCCCATTACCAAATGGGCGCATCGGGTGATGGCAACCGAGCATA
TCCCACGGCTGGTGATGCAGGCGATCCGCGCCGCGTTGAGCGCGCCACGCGGGCCGGTGTTGCTGGATCTGCCGTGGGATATT
CTGATGAACCAGATTGATGAGGATAGCGTCATTATCCCCGATCTGGTCTTGTCCGCGCATGGGGCCAGACCCGACCCTGCCGA
TCTGGATCAGGCTCTCGCGCTTTTGCGCAAGGCGGAGCGGCCGGTCATCGTGCTCGGCTCAGAAGCCTCGCGGACAGCGCGC
AAGACGGCGCTTAGCGCCTTCGTGGCGGCGACTGGCGTGCCGGTGTTTGCCGATTATGAAGGGCTAAGCATGCTCTCGGGGCT
GCCCGATGCTATGCGGGGCGGGCTGGTGCAAAACCTCTATTCTTTTGCCAAAGCCGATGCCGCGCCAGATCTCGTGCTGATGC
TGGGGGCGCGCTTTGGCCTTAACACCGGGCATGGATCTGGGCAGTTGATCCCCCATAGCGCGCAGGTCATTCAGGTCGACCCT
GATGCCTGCGAGCTGGGACGCCTGCAGGGCATCGCTCTGGGCATTGTGGCCGATGTGGGTGGGACCATCGAGGCTTTGGCGC
AGGCCACCGCGCAAGATGCGGCTTGGCCGGATCGCGGCGACTGGTGCGCCAAAGTGACGGATCTGGCGCAAGAGCGCTATGC
CAGCATCGCTGCGAAATCGAGCAGCGAGCATGCGCTCCACCCCTTTCACGCCTCGCAGGTCATTGCCAAACACGTCGATGCA
GGGGTGACGGTGGTAGCGGATGGTGCGCTGACCTATCTCTGGCTGTCCGAAGTGATGAGCCGCGTGAAACCCGGCGGTTTTCT
CTGCCACGGCTATCTAGGCTCGATGGGCGTGGGCTTCGGCACGGCGCTGGGCGCGCAAGTGGCCGATCTTGAAGCAGGCCGC
CGCACGATCCTTGTGACCGGCGATGGCTCGGTGGGCTATAGCATCGGTGAATTTGATACGCTGGTGCGCAAACAATTGCCGCT
GATCGTCATCATCATGAACAACCAAAGCTGGGGGGCGGGCTTGCATATCCACGATGTGACGGGCACCCGTTTGGAAAATGGC
TCCTATCACGGGGTGGCCGCCGCCTTTGGCGCGGATGGCTATCATGTCGACAGTGTGGAGAGCTTTTCTGCGGCTCTGGCCCA
AGCGCTCGCCCATAATCGCCCCGCCTGCATCAATGTCGCGGTCGCGCTCGATCCGATCCCGCCCGAAGAACTCATTCTGATCG
GCATGGACCCCTTCGCACTCGAG

6 REFERENCES

AGRANAT, I; CANER, H. & CALDWELL, J. (2002) Putting chirality to work: the strategy of chiral switches. *Nat. Rev. Drug Discovery*, 1, 753–768.

ALVAREZ, M.E.; ROSA, A.L.; TEMPORINI, E.D.; WOLSTENHOLME, A.; PANZETTA, G.; PATRITO, L. & MACCIONI, H.J. (1993) The 59-kDa polypeptide constituent of 8-10 nm cytoplasmic filaments in *Neurospora crassa* is a pyruvate decarboxylase. *Gene*, 130, 253–258.

AMÁRITA, F.; FERNÁNDEZ-ESPLÁ, D.; REQUENA, T. & PELAEZ, C. (2001) Conversion of methionine to methional by *Lactococcus lactis*. *FEMS Microbiol. Lett.*, 204, 189–195.

ANASTAS, P.T. & WARNER, J.C. (1998) Green chemistry: theory and practice. Oxford University Press, New York.

ANDREWS, F.H.; TOM, A.R.; GUNDERMAN, P.R.; NOVAK, W.R.P. & MCLEISH, M.J. (2013) A bulky hydrophobic residue is not required to maintain the V-conformation of enzyme-bound thiamin diphosphate. *Biochemistry*, 52, 3028–3030.

BARROWMAN, M.M. & FEWSON, C.A. (1985) Phenylglyoxylate decarboxylase and phenylpyruvate decarboxylase from *acinetobacter calcoaceticus*. *Curr. Microbiol.*, 12, 235–239.

BASF AG (2008) http://m.pressebox.de/boxid/182938. *Pressrelease*.

BEIGI, M.; WALTZER, S.; FRIES, A.; EGGELING, L.; SPRENGER, G.A. & MÜLLER, M. (2013) TCA cycle involved enzymes SucA and Kgd , as well as MenD : Efficient biocatalysts for asymmetric C-C bond formation. *Org. Lett.*, 15, 452–455.

BERGMEIER, S.C. (2000) The synthesis of vicinal amino alcohols. *Tetrahedron*, 56, 2561–2576.

BERTHOLD, C.L.; GOCKE, D.; WOOD, M.D.; LEEPER, F.J.; POHL, M. & SCHNEIDER, G. (2007) Structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemoselective and enantioselective carboligation reaction. *Acta Cryst. D*, 63, 1217–1224.

BLANCO, F.; KELLY, B.; ALKORTA, I.; ROZAS, I. & ELGUERO, J. (2011) Cation-pi interactions: Complexes of guanidinium and simple aromatic systems. *Chem. Phys. Lett.*, 511, 129–134.

BORNEMANN, S.; CROUT, D.H.G.; DALTON, H.; KREN, V.; LOBELL, M.; DEAN, G.; THOMSON, N. & TURNER, M.M. (1996) Stereospecific formation of R-aromatic acyloins by *Zymomonas mobilis* pyruvate decarboxylase. *J. Chem. Soc. Perkin Trans. 1*, 1, 425–430.

BORNSCHEUER, U. & KAZLAUSKAS, R.J. (2011) Survey of protein engineering strategies. *Curr. Protoc. Protein Sci.*, Chapter 26, Unit26.7.

BORNSCHEUER, U.T.; HUISMAN, G.W.; KAZLAUSKAS, R.J.; LUTZ, S.; MOORE, J.C. & ROBINS, K. (2012) Engineering the third wave of biocatalysis. *Nature*, 485, 185–194.

BRESLOW, R. (1957) The mechanism of thiamine action. 2. Rapid deuterium exchange in thiazolium salts. *J. Am. Chem. Soc.*, 79, 1762–1763.

BREUER, M.; POHL, M.; HAUER, B. & LINGEN, B. (2002) High-throughput assay of (*R*)-phenylacetylcarbinol synthesized by pyruvate decarboxylase. *Anal. Bioanal. Chem.*, 374, 1069–1073.

BRINGER-MEYER, S. & SAHM, H. (1988) Acetoin and phenylacetylcarbinol formation by pyruvate decarboxylase of *Zymomonas mobilis* and *Saccharomyces carlsbergensis*. *Biocatalysis*, 1, 321–331.

BROVETTO, M.; GAMENARA, D.; SAENZ MÉNDEZ, P. & SEOANE, G.A. (2011) C–C Bond-forming lyases in organic synthesis. *Chem. Rev.*, 111, 4346–4403.

CAVINATO, G.; TONIOLO, L. & VAVASORI, A. (1994) Synthesis of gamma-ketocarboxylic acids via reduction of gamma-keto-alpha-hydroxycarboxylic acids with carbon-monoxide catalyzed by a pd-hcl system. *J. Mol. Catal.*, 89, 93–100.

CHANDRA RAJ, K., INGRAM, L.O. & MAUPIN-FURLOW, J.A. (2001) Pyruvate decarboxylase: a key enzyme for the oxidative metabolism of lactic acid by *Acetobacter pasteurianus*. *Archives of Microbiology*, 176, 443–451.

CHEN, M.; JIANG, M. & GUO, Z. (2012) The thiamine-dependent enzyme of the vitamin K biosynthesis catalyzes reductive C-N bond ligation between nitroarenes and α-ketoacids. *Science China Chemistry*, 56, 312–320.

CHIPMAN, D.; BARAK, Z. & SCHLOSS, J. V (1998) Biosynthesis of 2-aceto-2-hydroxy acids: acetolactate synthases and acetohydroxyacid synthases. *Biochim. Biophys. Acta*, 1385, 401–419.

CHIR TU-BERLIN (2011). www.chir.tu-berlin.de.

COSTELLOE, S.J.; WARD, J.M. & DALBY, P.A. (2008) Evolutionary analysis of the TPP-dependent enzyme family. *Journal of Molecular Evolution*, 66, 36–49.

CROSSLEY, R. (1992) The Relevance of Chirality to the Study of Biological Activity. *Tetrahedron*, 48, 8155–8178.

DAWSON, A.; CHEN, M.; FYFE, P.K.; GUO, Z. & HUNTER, W.N. (2010) Structure and Reactivity of *Bacillus subtilis* MenD Catalyzing the First Committed Step in Menaquinone Biosynthesis. *J. Mol. Biol.* 401, 253–264.

DAWSON, A.; FYFE, P.K. & HUNTER, W.N. (2008) Specificity and Reactivity in Menaquinone Biosynthesis: The Structure of *Escherichia coli* MenD (2-Succinyl-5-Enolpyruvyl-6-Hydroxy-3-Cyclohexadiene-1-Carboxylate Synthase). *J. Mol. Biol.*, 384, 1353–1368.

DEMIR, A.S.; PERUZE, A.; UMUT, D. & UMUT, E. (2008) *Fusarium roseum* and *Aspergillus oryzae*-mediated enantioselective reduction of benzils to benzoin. *J. Mol. Catal. B: Enzym.*, 55, 164–168.

DEMIR, A.S.; POHL, M.; JANZEN, E. & MÜLLER, M. (2001) Enantioselective synthesis of hydroxy ketones through cleavage and formation of acyloin linkage. Enzymatic kinetic resolution via C–C bond cleavage. *J. Chem. Soc. Perkin Trans. 1*, 633–635.

DEMIR, A.S.; SEŞENOGLU, O.; DÜNKELMANN, P. & MÜLLER, M. (2003) Benzaldehyde lyase-catalyzed enantioselective carboligation of aromatic aldehydes with mono- and dimethoxy acetaldehyde. *Org. Lett.*, 5, 2047–2050.

DEMIR, A.S.; SESENOGLU, O.; EREN, E.; HOSRIK, B.; POHL, M.; JANZEN, E., KOLTER, D.; FELDMANN, R.; DÜNKELMANN, P. & MÜLLER, M. (2002) Enantioselective synthesis of alpha-hydroxy ketones via benzaldehyde lyase-catalyzed C-C bond formation reaction. *Adv. Synth. Catal.*, 344, 96–103.

DOBRITZSCH, D.; KÖNIG, S.; SCHNEIDER, G. & LU, G. (1998) High resolution crystal structure of pyruvate decarboxylase from *Zymomonas mobilis*. Implications for substrate activation in pyruvate decarboxylases. *J. Biol. Chem.*, 273, 20196–20204.

DOMINGUEZ DE MARIA, P.; POHL, M.; GOCKE, D.; GRÖGER, H.; TRAUTWEIN, H.; STILLGER, T. & MÜLLER, M. (2007) Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes. *Eur. J. Org. Chem.*, 2940–2944.

DUGGLEBY, R.G. (2006) Domain relationships in thiamine diphosphate-dependent enzymes. *Acc. Chem. Res.*, 39, 550–557.

DÜNKELMANN, P.; KOLTER-JUNG, D.; NITSCHE, A.; DEMIR, A.S.; SIEGERT, P.; LINGEN, B.; BAUMANN, M.; POHL, M. & MÜLLER M. (2002) Development of a donor-acceptor concept for enzymatic cross-coupling reactions of aldehydes: the first asymmetric cross-benzoin condensation. *J. Am. Chem. Soc.*, 124, 12084–12085.

DÜNNWALD, T.; DEMIR, A.S.; SIEGERT, P.; POHL, M. & MÜLLER, M. (2000) Enantioselective synthesis of (*S*)-2-hydroxypropanone derivatives by benzoylformate decarboxylase catalyzed C–C bond formation. *Eur. J. Org. Chem.*, 2161–2170.

ENDERS, D. & KALLFASS, U. (2002) An efficient nucleophilic carbene catalyst for the asymmetric benzoin condensation. *Angew. Chem. Int. Ed.*, 41, 1743–1745.

ENGEL, S.; VYAZMENSKY, M.; BERKOVICH, D.; BARAK, Z. & CHIPMAN, D.M. (2004a) Substrate range of acetohydroxy acid synthase I from *Escherichia coli* in the stereoselective synthesis of alpha-hydroxy ketones. *Biotechnol. Bioeng.*, 88, 825–831.

ENGEL, S.; VYAZMENSKY, M.,; GERESH, S.; BARAK, Z. & CHIPMAN, D.M. (2003) Acetohydroxyacid synthase: a new enzyme for chiral synthesis of *R*-phenylacetylcarbinol. *Biotechnol. Bioeng.*, 83, 833–840.

ENGEL, S.; VYAZMENSKY, M.; VINOGRADOV, M.; BERKOVICH, D.; BAR-ILAN, A.; QIMRON, U.; ROSIANSKY, Y.; BARAK, Z. & CHIPMAN, D.M. (2004b) Role of a conserved arginine in the mechanism of acetohydroxyacid synthase: catalysis of condensation with a specific ketoacid substrate. *J. Biol. Chem.*, 279, 24803–24812.

ENZYME NOMENCLATURE (1992) Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes. Academic press, San Diego, California.

FABER, K. (2011) Biotransformations in organic chemistry. Springer-Verlag, Berlin.

FABER, K. & PATEL, R. (2000) Chemical biotechnology. A happy marriage between chemistry and biotechnology: asymmetric synthesis via green chemistry. *Curr. Opin. Biotechnol.*, 11, 517–519.

FANG, M.; MACOVA, A.; HANSON, K.L.; KOS, J. & PALMER, D.R.J. (2011) Using substrate analogues to probe the kinetic mechanism and active site of *Escherichia coli* MenD. *Biochemistry*, 50, 8712–8721.

FRANK, R.A.W.; LEEPER, F.J. & LUISI, B.F. (2007) Structure, mechanism and catalytic duality of thiamine-dependent enzymes. *Cell. Mol. Life Sci.*, 64, 892–905.

FUNKE, M.; BUCHENAUER, A.; SCHNAKENBERG, U.; MOKWA, W.; DIEDERICHS, S.; MERTENS, A.; MÜLLER, C.; KENSY, F. & BÜCHS, J. (2010) Microfluidic biolector-microfluidic bioprocess control in microtiter plates. *Biotechnol. Bioeng.*, 107, 497–505.

GAWLEY, R.E. & AUBÉ, J. (2012) Principles of Asymmetric Synthesis, 2nd Ed., Elsevier Limited, Oxford.

GEIBLER, J.; LIEDTKE, C.; ECKERMANN, A.; KAISER, C.; RITTHOFF, M.; HARTMANN, D.; ROHN, H.; BRENTEL, M.; KUHNDT, M & WALLBAUM, H. (2005) BioBeN – Assesing social sustainability of biotech products. BMBF, Wuppertal.

GERHARDS, T. (2012) Einfluss unkonventioneller Medien auf die Selektivität ThDP-abhängiger Enzyme. Dissertation. Heinrich-Heine-Universität Düsseldorf.

GERHARDS, T.; MACKFELD, U.; BOCOLA, M.; VON LIERES, E.; WIECHERT, W.; POHL, M. & ROTHER, D. (2012) Influence of organic solvents on enzymatic asymmetric carboligations. *Adv. Synth. Catal.*, 354, 2805–2820.

GOCKE, D. (2007) New and optimised thiamine diphosphate (ThDP)-dependent enzymes for carboligation. Dissertation. Heinrich-Heine-Universität Düsseldorf.

GOCKE, D.; GRAF, T.; BROSI, H.; FRINDI-WOSCH, I.; WALTER, L.; MÜLLER, M. & POHL, M. (2009) Comparative characterisation of thiamin diphosphate- dependent decarboxylases. *J. Mol. Catal. B: Enzym.*, 61, 30–35.

GOCKE, D.; KOLTER, G.; MÜLLER, M. & POHL, M. (2010) Entwicklung einer Enzym-Plattform für die biokatalytische C–C Verknüpfung. *Chem. Ing. Tech.*, 82, 81–86.

GOCKE, D.; NGUYEN, C.L.; POHL, M.; STILLGER, T.; WALTER, L. & MÜLLER, M. (2007) Branched-chain keto acid decarboxylase from *Lactococcus lactis* (KdcA), a valuable thiamine diphosphate-dependent enzyme for asymmetric C–C bond formation. *Adv. Synth. Catal.*, 349, 1425–1435.

GOCKE, D.; WALTER, L.; GAUCHENOVA, E.; KOLTER, G.; KNOLL, M.; BERTHOLD, C.L.; SCHNEIDER, G.; PLEISS, J.; MÜLLER, M. & POHL, M. (2008) Rational protein design of ThDP-dependent enzymes-engineering stereoselectivity. *ChemBioChem*, 9, 406–412.

GÓMEZ BARAIBAR, A.; VON LIERES, E.; WIECHERT, W.; POHL, M. & ROTHER, D. (2013) Effective production of (*S*)-a-hydroxy ketones: an reaction engineering approach. *Top. Catal*. DOI: 10.1007/s11244-013-0237-5.

GONZÁLEZ, B. & VICUÑA, R. (1989) Benzaldehyde lyase, a novel thiamine PPi-requiring enzyme, from *Pseudomonas fluorescens* biovar I. *J. Bacteriol.*, 171, 2401–2405.

GRÖGER, H. & ASANO, Y. (2012) Potential of enzymes as catalysts in organic synthesis. In *Enzyme catalysis in organic synthesis*. 3rd Ed. (Eds.: May, O.; Gröger, H.; Drauz, W.). pp. 14–17. Wiley-VCH, Weinheim.

GUO, F.S.; ZHANG, D.Q.; KAHYAOGLU, A.; FARID, R.S. & JORDAN, F. (1998) Is a hydrophobic amino acid required to maintain the reactive V conformation of thiamin at the active center of thiamin diphosphate-requiring enzymes? Experimental and computational studies of isoleucine 415 of yeast pyruvate decarboxylase. *Biochemistry*, 37, 13379–13391.

HAILES, H.C.; ROTHER, D.; MÜLLER, M.; WESTPHAL, R.; WARD, J.M.; PLEISS, J., VOGEL, C. & POHL, M. (2013) Engineering stereoselectivity of ThDP-dependent enzymes. *FEBS J.*, 280, 6374–6394.

HAMMETT, L.P. (1937) Effect of structure upon the reactions of organic compounds. Benzene derivatives. *J. Am. Chem. Soc.*, 59, 96–103.

HASSON, M.S.; MUSCATE, A.; MCLEISH, M.J.; POLOVNIKOVA, L.S.; GERLT, J.A.; KENYON, G.L.; PETSKO, G.A. & RINGE, D. (1998) The crystal structure of benzoylformate decarboxylase at 1.6 A resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. *Biochemistry*, 37, 9918–9930.

HAWKINS, C.F.; BORGES, A. & PERHAM, R.N. (1989) A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS Lett.*, 255, 77–82.

HEGEMAN, G.D. (1966) Synthesis of the enzymes of the mandelate pathway by *Pseudomonas putida* I. Synthesis of enzymes by the wild type. *J. Bacteriol.*, 91, 1140–1154.

HEGEMAN, G.D. (1970) Benzoylformate decarboxylase (*Pseudomonas putida*). *Meth. Enzymol.*, 17A, 674–678.

HERNÁIZ, M.J.; ALCÁNTARA, A.R.; GARCÍA, J.I. & SINISTERRA, J. V (2010) Applied biotransformations in green solvents. *Chem. Eur. J.*, 16, 9422–9437.

HINRICHSEN, P.; GOMEZ, I. & VICUNA, R. (1994) Cloning and sequencing of the gene encoding benzaldehyde lyase from *Pseudomonas-fluorescens* Biovar-I. *Gene*, 144, 137–138

HÖHNE, M.; SCHÄTZLE, S.; JOCHENS, H.; ROBINS, K. & BORNSCHEUER, U.T. (2010) Rational assignment of key motifs for function guides *in silico* enzyme identification. *Nat. Chem. Biol.*, 6, 807–813.

HOLTON, R.A.; KIM, H.-B.; SOMOZA, C.; LIANG, F.; BIEDIGER, R.J.; BOATMAN, P.D.; SHINDO, M.; SMITH, C.C.; KIM, S.; NADIZADEH, H.; SUZUKI, Y.; TAO, C.; VU, P.; TANG, S.; ZHANG, P.; MURTHI, K.K.; GENTILE, L.N. & LIU, J.H. (1994) First total synthesis of Taxol. 2. Completion of the C-ring and D-ring. *J Am Chem Soc*, 116, 1599–1600.

HOLTON, R.A.; KIM, H.-B.; SOMOZA, C.; LIANG, F.; BIEDIGER, R.J.; BOATMAN, P.D.; SHINDO, M.; SMITH, C.C.; KIM, S.; NADIZADEH, H.; SUZUKI, Y.; TAO, C.; VU, P.; TANG, S.; ZHANG, P.; MURTHI, K.K.; GENTILE, L.N. & LIU, J.H. (1994) First total synthesis of Taxol. 1. Functionalization of the B-ring. *J Am Chem Soc*, 116, 1597–1598.

HOLZER, H.; SCHULTZ, G.; VILLARPALASI, C. & JUNTGENSELL, J. (1956) Isolierung der Hefecarboxylase und Untersuchungen über die Aktivität des Enzyms in lebenden Zellen. *Biochem. Z.*, 5, 331–344.

HOSSAIN, M.A.; HUQ, E.; GROVER, A.; DENNIS, E.S.; PEACOCK, W.J. & HODGES, T.K. (1996) Characterization of pyruvate decarboxylase genes from rice. *Plant Molecular Biology*, 31, 761–770.

HOYOS, P.; SINISTERRA, J.-V.; MOLINARI, F.; ALCÁNTARA, A.R. & DOMÍNGUEZ DE MARÍA, P. (2010) Biocatalytic strategies for the asymmetric synthesis of alpha-hydroxy ketones. *Acc. Chem. Res.*, 43, 288–299.

IDING, H.; DÜNNWALD, T.; GREINER, L.; LIESE, A.; MÜLLER, M.; SIEGERT, P., GRÖTZINGER J.; DEMIR, A.S. & POHL, M. (2000) Benzoylformate decarboxylase from *Pseudomonas putida* as stable catalyst for the synthesis of chiral 2-hydroxy ketones. *Chem. Eur. J.*, 6, 1483–1495.

IUPAC (1997) Compendium of chemical terminology, 2nd edition. Blackwell Scientific Publications, Oxford.

JACOBSEN, E.N.; MARKO, I.; MUNGALL, W.S.; SCHROEDER, G. & SHARPLESS, K.B. (1988) Asymmetric dihydroxylation via ligand-accelerated catalysis. *J. Am. Chem. Soc.*, 110, 1968–1970.

JAÑA, G.; JIMÉNEZ, V.; VILLÀ-FREIXA, J.; PRAT-RESINA, X.; DELGADO, E. & ALDERETE, J. (2010) Computational study on the carboligation reaction of acetohydroxyacid synthase: new approach on the role of the HEThDP-intermediate. *Proteins*, 78, 1774–1788.

JANSEN, S. (2013) Erweiterung der Toolbox Thiamindiphosphat-abhängiger Enzyme für die Synthese enantiokomplementärer α-Hydroxyketone: Charakterisierung von MenD aus *Bacillus subtilis*. Bachelor thesis. Fachhochschule Aachen, Campus Jülich.

JANZEN, E.; MÜLLER, M.; KOLTER-JUNG, D.; KNEEN, M.M.; MCLEISH, M.J. & POHL, M. (2006) Characterization of benzaldehyde lyase from *Pseudomonas fluorescens*: A versatile enzyme for asymmetric C-C bond formation. *Bioorg. Chem.*, 34, 345–361.

JIANG, M.; CAO, Y.; GUO, Z.-F.; CHEN, M.; CHEN, X. & GUO, Z. (2007) Menaquinone biosynthesis in *Escherichia coli*: identification of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate as a novel intermediate and re-evaluation of MenD activity. *Biochemistry*, 46, 10979–10989.

JORDAN, F. (2003) Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. *Nat. Prod. Rep.*, 20, 184–201.

NICOLAOU, K.C.; YANG, Z.; LIU, J.-J.; UENO, H.; NANTERMET, P.G.; GUY, R.K.; CLAIBORNE, C.F.; RENAUD, J.; COULADOUROS, E.A.; PAULVANNAN, K. & SORENSEN, E.J. (1994) Total synthesis of Taxol. *Nature*, 367, 630–634.

KAZLAUSKAS, R.J. (2005) Enhancing catalytic promiscuity for biocatalysis. *Curr. Opin. Chem. Biol.*, 9, 195–201.

KERN, D.; KERN, G.; NEEF, H.; TITTMANN, K.; KILLENBERG-JABS, M.; WIKNER, C.; SCHNEIDER, G. & HUBNER, G. (1997) How thiamine diphosphate is activated in enzymes. *Sciences*, 275, 67–70.

KLAWEHN, W. & HILDERBRANDT, G. (1934) Manufacture of laevo-1-phenyl-2-methylaminopropanol-1. US Patent 1,956,950.

KLUGER, R. & TITTMANN, K. (2008) Thiamin diphosphate catalysis: enzymic and nonenzymic covalent intermediates. *Chem. Rev.*, 108, 1797–1833.

KNOLL, M.; MÜLLER, M.; PLEISS, J. & POHL, M. (2006) Factors mediating activity, selectivity, and substrate specificity for the thiamin diphosphate-dependent enzymes benzaldehyde lyase and benzoylformate decarboxylase. *ChemBioChem*, 7, 1928–1934.

KOLTER, G. (2010) Neue Thiamindiphosphat-abhängige Enzyme für die Synthese enantiokomplementärer 2-Hydroxyketone. Dissertation. Heinrich-Heine-Universität Düsseldorf.

KÖNIG, S. (1998) Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. *Biochim. Biophys. Acta*, 1385, 271–286.

KRIEGER, E.; KORAIMANN, G. & VRIEND, G. (2002) Increasing the precision of comparative models with YASARA NOVA - a self-parameterizing force field. *Proteins*, 47, 393–402.

KULIG, J.; SIMON, R.C.; ROSE, C.A.; HUSAIN, S.M.; HÄCKH, M.; LÜDEKE, S.; ZEITLER, K.; KROUTIL, W.; POHL, M. & ROTHER, D. (2012) Stereoselective synthesis of bulky 1,2-diols with alcohol dehydrogenases. *Catal. Sci. Technol.*, 2, 1580–1589.

KURUTSCH, A. (2009) Shikimatbiosyntheseweg: Mikrobielle Produktgewinnung der Metabolite Chorismat und SHCHC und Einsatz des Enzyms MenD als Biokatalysator für asymmetrische C-C-Bindungsknüpfung. Dissertation. Albert-Ludwigs-Universität Freiburg.

KURUTSCH, A.; RICHTER, M.; BRECHT, V.; SPRENGER, G.A. & MÜLLER, M. (2009) MenD as a versatile catalyst for asymmetric synthesis. *J. Mol. Catal. B: Enzym.*, 61, 56–66.

LEUTHARDT, F. & NIELSEN, H. (1952) Phosphorylation biologique de la thiamine. *Helv. Chim. Acta*, 35, 1196–1209.

LIESE, A. (2003) Biological principles applied to technical asymmetric catalysis. Schriften des Forschungszentrums Jülich, Reihe Lebenswissenschaften, Band 7, Forschungszentrum Jülich GmbH.

LIESE, A.; SEELBACH, K. & WANDREY, C. (2006) Industrial Biotransformations, 2nd edition. John Wiley & Sons.

LINDQVIST, Y.; SCHNEIDER, G.; ERMLER, U. & SUNDSTRÖM, M. (1992) Three-dimensional structure of transketolase, a thiamine diphosphate dependent enzyme, at 2.5 A resolution. *EMBO J.*, 11, 2373–2379.

LINGEN, B.; GRÖTZINGER, J.; KOLTER, D.; KULA, M.-R. & POHL, M. (2002) Improving the carboligase activity of benzoylformate decarboxylase from *Pseudomonas putida* by a combination of directed evolution and site-directed mutagenesis. *Protein Eng.*, 15, 585–593.

LORENZ, P. & ECK, J. (2005) Metagenomics and industrial applications. *Nat. Rev. Microbiol.*, 3, 510–516.

MEYER, D.; NEUMANN, P.; FICNER, R. & TITTMANN, K. (2013b) Observation of a stable carbene at the active site of a thiamin enzyme. *Nat. Chem. Biol.*, 9, 488–490.

MEYER, D.; WALTER, L.; KOLTER, G.; POHL, M.; MÜLLER, M. & TITTMANN, K. (2011) Conversion of pyruvate decarboxylase into an enantioselective carboligase with biosynthetic potential. *J. Am. Chem. Soc.*, 133, 3609–3616.

MEYER, H.-P.; EICHHORN, E.; HANLON, S.; LÜTZ, S., SCHÜRMANN, M.; WOHLGEMUTH, R. & COPPOLECCHIA, R. (2013a) The use of enzymes in organic synthesis and the life sciences: perspectives from the Swiss Industrial Biocatalysis Consortium (SIBC). *Catal. Sci. Technol.*, 3, 29–40.

MITRA, S. & CHOPRA, P. (2011) Chirality and anaesthetic drugs: A review and an update. *Indian Journal of Anaesthesia*, 55, 556–562.

MOSBACHER, T.G.; MÜLLER, M. & SCHULZ, G.E. (2005) Structure and mechanism of the ThDP-dependent benzaldehyde lyase from *Pseudomonas fluorescens*. *FEBS J.*, 272, 6067–6076.

MÜLLER, M.; GOCKE, D. & POHL, M. (2009) Thiamin diphosphate in biological chemistry: exploitation of diverse thiamin diphosphate-dependent enzymes for asymmetric chemoenzymatic synthesis. *FEBS J.*, 276, 2894–2904.

MÜLLER, M.; SPRENGER, G.A. & POHL, M. (2013) C-C bond formation using ThDP-dependent lyases. *Curr. Opin. Chem. Biol.*, 17, 261–270.

MULLER, Y.A.; LINDQVIST, Y.; FUREY, W.; SCHULZ, G.E.; JORDAN, F. & SCHNEIDER, G. (1993) A thiamin diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase. *Structure London England 1993*, 1, 95–103.

NEUBERG, C. & HIRSCH, J. (1921) Über ein Kohlenstoffketten knüpfendes Ferment (Carboligase). *Biochem. Z.*, 115, 282.

NICOLAOU, K.; VOURLOUMIS, D.; WINSSINGER, N. & BARAN, P. (2000) The art and science of total synthesis at the dawn of the twenty-first century. *Angew. Chem. Int. Ed.*, 39, 44–122.

NICOLAOU, K.C. & SORENSEN, E.J. (1996) Classics in Total Synthesis. VCH, New York.

O'TOOLE, S.E.; ROSE, C.A.; GUNDALA, S.; ZEITLER, K. & CONNON, S.J. (2011) Highly chemoselective direct crossed aliphatic-aromatic acyloin condensations with triazolium-derived carbene catalysts. *J. Org. Chem.*, 76, 347–357.

OKU, H. & KANEDA, T. (1988) Biosynthesis of branched-chain fatty acids in *Bacillus subtilis*. A decarboxylase is essential for branched-chain fatty acid synthetase. *J. Biol. Chem.*, 263, 18386–18396.

PANG, S.S.; DUGGLEBY, R.G. & GUDDAT, L.W. (2002) Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors. *J. Mol. Biol.*, 317, 249–262.

PATEL, R.;HANSON, R.; GOSWAMI, A.; NANDURI, V.; BANERJEE, A.; DONOVAN, M. –J., GOLDBERG, S.; JOHNSTON, R.; BRZOZOWSKI, D.; TULLY, T., HOWELL, J.;CAZZULINO, D.; & KO, R. (2003) Enzymatic synthesis of chiral intermediates for pharmaceuticals. *J. Ind. Microbiol. and Biotechnol.*, 30, 252–259.

PATEL, R.N. (2008) Synthesis of chiral pharmaceutical intermediates by biocatalysis. *Coord. Chem. Rev.*, 252, 659–701.

PATEL, R.N. (2011) Biocatalysis: Synthesis of key intermediates for development of pharmaceuticals. *ACS Catal.*, 1, 1056–1074.

PATRICK, W.M.; FIRTH, A.E. & BLACKBURN, J.M. (2003) User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. *Protein Eng.*, 16, 451–457.

PHILLIPS, R.S. (1992) Temperature effects on stereochemistry of enzymatic reactions. *Enzyme Microb. Technol.*, 14, 417–419.

POHL, M.; DRESEN, C.; BEIGI, M. & MÜLLER, M. (2012) Enzymatic acyloin and benzoin condensations. In *Enzyme Catalysis in Organic Synthesis* (Eds. Drauz, K. Gröger, H. & May, O.), pp. 919–945. Wiley-VCH, Weinheim.

POHL, M.; GOCKE, D. & MÜLLER, M. (2009) Thiamine-based enzymes for biotransformations. In *Handbook of Green Chemistry–Green Catalysis Vol. 3* (Ed. Anastas, P.T.), pp. 75–114. Wiley-VCH, Weinheim.

POHL, M.; GRÖTZINGER, J.; WOLLMER, A. & KULA, M.R. (1994) Reversible dissociation and unfolding of pyruvate decarboxylase from *Zymomonas mobilis*. *Eur. J. Biochem.*, 224, 651–661.

POLLAK, P. (2011) Fine Chemicals – The Industry and the Business, 2nd Edition. J Wiley & Sons.

PRONK, S.; PÁLL, S.; SCHULZ, R.; LARSSON, P.; BJELKMAR, P.; APOSTOLOV, R.; SHIRTS, M.R.; SMITH, J.C.; KASSON, P.M.; VAN DER SPOEL, D.; HESS, B. & LINDAHL, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29, 845–854.

PSCHYREMBEL, W. (2013) Pschyrembel Klinisches Wörterbuch (265. A.). Gruyter, Walter de GmbH.

PUTMAN, D.G.; HOGENKAMP, D.J.; DASSE, O.A.; WHITTEMORE, E.R. & JENSEN, M.S. (2007) Enantiomerically pure (*R*)-etifoxine pharmaceutical compositions thereof and methods of their use. *WO 2007/109288 A2.*

WOODWARD, R.B.; CAVA, M.P.; OLLIS, W.D.; HUNGER, A.; DAENIKER, H.U. & SCHENKER, K. (1954) The total synthesis of strychnine. *J. Am. Chem. Soc.*, 76, 4749–4751.

REETZ, M.T.; KAHAKEAW, D. & LOHMER, R. (2008) Addressing the numbers problem in directed evolution. *ChemBioChem*, 9, 1797–1804.

ROHE, P.; VENKANNA, D.; KLEINE, B.; FREUDL, R. & OLDIGES, M. (2012) An automated workflow for enhancing microbial bioprocess optimization on a novel microbioreactor platform. *Microb. Cell Fact.*, 11:144.

ROTHER, D.; KOLTER, G.; GERHARDS, T.; BERTHOLD, C.L.; GAUCHENOVA, E.; KNOLL, M., PLEISS, J.; MÜLLER, M.; SCHNEIDER, G. & POHL, M. (2011) *S*-selective mixed carboligation by structure-based design of the pyruvate decarboxylase from *Acetobacter pasteurianus*. *ChemCatChem*, 3, 1587–1596.

SCHELLENBERGER, A. (1998) Sixty years of thiamin diphosphate biochemistry. *Biochim. Biophys. Acta*, 1385, 177–186.

SCHLOSS, J. V; VAN DYK, D.E.; VASTA, J.F. & KUTNY, R.M. (1985) Purification and properties of *Salmonella typhimurium* acetolactate synthase isozyme II from *Escherichia coli* HB101/pDU9. *Biochemistry*, 24, 4952–4959.

SCHMITZ, C. (2012) Untersuchung der Stereoselektivität und des Substratspektrums von AHASI & II aus *Escherichia coli*. Bachelor thesis. Fachhochschule Aachen, Campus Jülich.

SEHL, T.; KULIG, J.; WESTPHAL, R. & ROTHER, D. (2013) Synthetic enzyme cascades for valuable diols and amino alcohols – smart composition and optimization strategies. In *Industrial Biocatalysis* (Ed.: Grunwald, P.), Pan Stanford Publishing, Singapore, *accepted for publication*.

SEHL, T.; SIMON, R.; HAILES, H.; WARD, J.; SCHELL, U.; POHL, M. & ROTHER, D. (2012) TTC-based screening assay for ω-transaminases: a rapid method to detect reduction of 2-hydroxy ketones. *J. Biotechnol.*, 159, 1–7.

SHAANAN, B. & CHIPMAN, D.M. (2009) Reaction mechanisms of thiamin diphosphate enzymes: new insights into the role of a conserved glutamate residue. *FEBS J.*, 276, 2447–2453.

SHUKLA, V.B. & KULKARNI, P.R. (2000) L-Phenylacetylcarbinol (L-PAC): biosynthesis and industrial application. *World J. Microbiol. Biotechnol.*, 16, 499–506.

SIJBESMA, F. & SCHEPENS, H. (2003) White Biotechnology: Gateway to a more sustainable future. *EuropaBio.* Lyon, 10.04.2003.

SMIT, B.A.;VAN HYLCKAMA VLIEG, J.E.T.; ENGELS, W.J.M.; MEIJER, L.; WOUTERS, J.T.M. & SMIT, G. (2005) Identification, cloning, and characterization of a *Lactococcus lactis* branched-chain α-keto acid decarboxylase involved in flavor formation. *Appl. Environ. Microbiol.*, 71, 303–311.

SMITH, S.W. (2009) Chiral toxicology: it's the same thing...only different. *Toxicological sciences : an official journal of the Society of Toxicology*, 110, 4–30.

STANIER, R.Y. & ORNSTON, L.N. (1973) The beta-ketoadipate pathway. *Adv. Microbiol. Physiol.*, 9, 89–151.

STEINMETZ, A. (2010) Untersuchungen zum Katalyse- und Regulationsmechanismus Thiamin- und Flavin-abhängiger Carboligasen aus *Escherichia coli*. Dissertation. Martin-Luther-Universität Halle-Wittenberg.

SUKUMARAN, J. & HANEFELD, U. (2005) Enantioselective C-C bond synthesis catalysed by enzymes. *Chem. Soc. Rev.*, 34, 530–542.

TANG, S.L.Y., SMITH, R.L. & POLIAKOFF, M. (2005) Principles of green chemistry: PRODUCTIVELY. *Green Chemistry*, 7, 761–762.

THAYER, A.M. (2006) Enzymes at work. *Chem. Eng. News*, 84, 15–25.

UMBARGER, H.E. (1978) Amino acid Biosynthesis and its Regulation. *Ann. Rev. Biochem*, 47, 533–606.

UMBARGER, H.E. & BROWN, B. (1958) Isoleucine and valine metabolism in *Escherichia coli*. VIII. The formation of acetolactate. *J. Biol. Chem.*, 233, 1156–1160.

VINOGRADOV, V.; VYAZMENSKY, M.; ENGEL, S.; BELENKY, I.; KAPLUN, A.; KRYUKOV, O.; BARAK, Z. & CHIPMAN, D.M. (2006) Acetohydroxyacid synthase isozyme I from *Escherichia coli* has unique catalytic and regulatory properties. *Biochim. Biophys. Acta*, 1760, 356–363.

VOGEL, C.; WIDMANN, M.; POHL, M. & PLEISS, J. (2012) A standard numbering scheme for thiamine diphosphate-dependent decarboxylases. *BMC Biochem.*, 13:24.

VYAZMENSKY, M.; STEINMETZ, A.; MEYER, D.; GOLBIK, R.; BARAK, Z.; TITTMANN, K. & CHIPMAN, D.M. (2011) Significant catalytic roles for Glu47 and Gln 110 in all four of the C–C bond-making and -breaking steps of the reactions of acetohydroxyacid synthase II. *Biochemistry*, 50, 3250–3260.

VYAZMENSKY, M.; ZHERDEV, Y.; SLUTZKER, A.; BELENKY, I.; KRYUKOV, O.; BARAK, Z. & CHIPMAN, D.M. (2009) Interactions between large and small subunits of different acetohydroxyacid synthase isozymes of *Escherichia coli*. *Biochemistry*, 48, 8731–8737.

WARD, O.P. & SINGH, A. (2000) Enzymatic asymmetric synthesis by decarboxylases. *Curr. Opin. Biotechnol.*, 11, 520–526.

WEINSTOCK, O.; SELLA, C.; CHIPMAN, D.M. & BARAK, Z. (1992) Properties of subcloned subunits of bacterial acetohydroxy acid synthases. *J. Bacteriol.*, 174, 5560–5566.

WENDA, S.; ILLNER, S.; MELL, A. & KRAGL, U. (2011) Industrial biotechnology—the future of green chemistry? *Green Chemistry*, 13, 3007–3047.

WENDORFF, M. (2006) Neue Benzoylformiatdecarboxylasen für die Biokatalyse. Dissertation Heinrich-Heine-Universität Düsseldorf.

WENK, P.; HEMMERICH, J.; MULLER, C. & KENSY, F. (2012) High-throughput bioprocess development in shaken microbioreactors. *Chem. Ing. Tech.*, 84, 704–714.

WIDMANN, M.; RADLOFF, R. & PLEISS, J. (2010) The thiamine diphosphate dependent enzyme engineering database: A tool for the systematic analysis of sequence and structure relations. *BMC Biochem.*, 11:9.

WILCOCKS, R. & WARD, O.P. (1992) Factors affecting 2-hydroxypropiophenone formation by benzoylformate decarboxylase from *Pseudomonas putida*. *Biotechnol. Bioeng.*, 39, 1058–1063.

WILCOCKS, R.; WARD, O.P.; COLLINS, S.; DEWDNEY, N.J.; HONG, Y. & PROSEN, E. (1992) Acyloin formation by benzoylformate decarboxylase from *Pseudomonas putida*. *Appl. Environ. Microbiol.*, 58, 1699–1704.

WÖHLER, F. (1828) Untersuchungen über das Radikal der Benzoesäure. *Ann. Phys. Chem.*, 253.

WOLFENDEN, R. & SNIDER, M.J. (2001) The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.*, 34, 938–945.

YADAV, J.S. & MANDAL, S.S. (2011) First stereoselective total synthesis of pectinolide A. *Tetrahedron Lett.*, 52, 5747–5749.

YEP, A.; KENYON, G.L. & MCLEISH, M.J. (2006) Determinants of substrate specificity in KdcA, a thiamin diphosphate-dependent decarboxylase. *Bioorg. Chem.*, 34, 325–336.

ZHOU, P.; CHEN, B.A. & DAVIS, F.A. (2001) Asymmetric oxidation reactions. In pp. 128–145. Oxford University Press, Oxford.