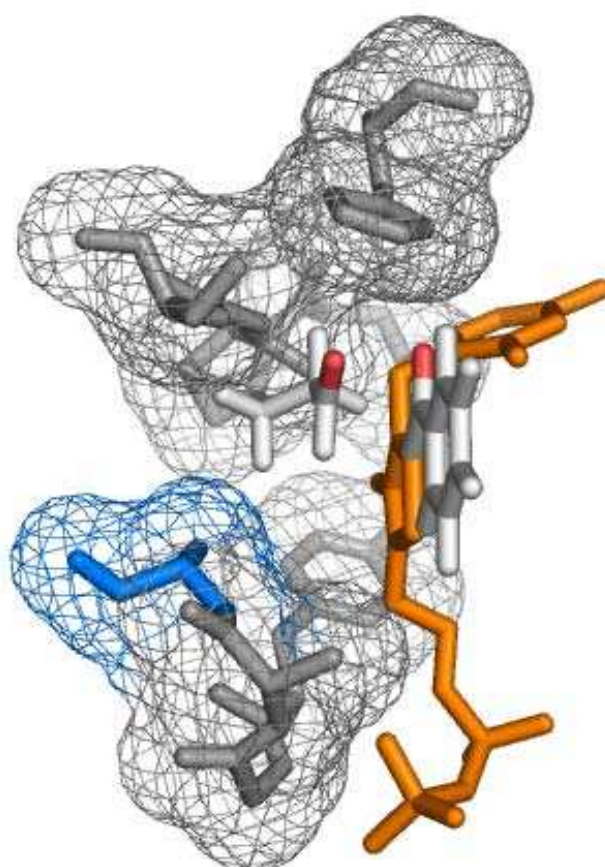**Dörte Gocke**

New and optimised thiamine diphosphate (ThDP)-dependent enzymes for carboligation

Creation of a toolbox for chiral 2-hydroxy ketones

Thesis for doctoral degree (Dr. rer. nat.)

New and optimised thiamine diphosphate (ThDP)-dependent enzymes for carboligation

- Creation of a toolbox for chiral 2-hydroxy ketones

Inaugural-Dissertation

zur Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Dörte Gocke

aus Dorsten

Köln, November 2007

Aus dem Institut für Molekulare Enzymtechnologie

der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent: PD Dr. Martina Pohl

Korreferent: Prof. Dr. Dieter Willbold

Tag der mündlichen Prüfung: 14. April 2008

"Das Schönste, was wir erleben können, ist das Geheimnisvolle. Es ist das Grundgefühl, das an der Wiege von wahrer Kunst und Wissenschaft steht. Wer es nicht kennt und sich nicht mehr wundern, nicht mehr staunen kann, der ist sozusagen tot und sein Auge erloschen."

Albert Einstein, aus ‚Mein Weltbild'

Für Jörn, David und Valentin

ABSTRACT

Various thiamine diphosphate (ThDP)-dependent enzymes catalyse the carboligation of aldehydes to chiral 2-hydroxy ketones with high stereoselectivity. Since 2-hydroxy ketones are versatile building blocks for organic syntheses, the aim of the thesis was the creation of a ThDP-dependent enzyme toolbox in order to access a diversely substituted and enantio-complementary 2-hydroxy ketone platform. Several goals could be achieved:

1. Development of a **quick and reliable high-throughput assay** to screen large toolboxes of purified enzymes or whole cell catalysts for the product formation of diversely substituted 2-hydroxy ketones.
2. **Enlargement of the enzyme toolbox** by six wild type enzymes and four variants showing carboligase activity. The platform now encompasses 15 wild type 2-keto acid decarboxylases and two benzaldehyde lyases as well as 70. For almost all newly added enzymes a biochemical characterisation of the substrate spectrum for the decarboxylase and carboligase activity including the determination of optimal reaction conditions was conducted.
3. With the new catalysts and the high-throughput assay an effective **enlargement of the 2-hydroxy ketone platform** was achieved. Branched-chain aliphatic 2-hydroxy ketones, substituted (*R*)-phenylacetylcarbinol-analogue products as well as several (*S*)-2-hydroxy ketones are now for the first time biocatalytically available. By combining decarboxylase activity and carboligase activity the ligation of instable aldehydes is also possible.

(*S*)-selective C-C bond formation was possible due to an understanding of the underlying **structure-function relationship** of the active site. By gaining two further wild type **crystal structures** the pool of structure information was broadened. Superimposition of all known structures combined with biochemical information and modelling studies enabled the mapping of catalytically important regions in the active site. It was concluded that the substrate ranges of both, the decarboxylase and the ligase activity of ThDP-dependent enzymes as well as the chemo- and enantioselectivity of the carboligase reaction is a result of optimal stabilisation of the substrates in the active site. This knowledge paves the way for product prediction by modelling studies and therewith the design of new catalysts with desired ligase activity.

KURZFASSUNG

Die Carboligation von Aldehyden zu chiralen 2-Hydroxyketonen wird von einer Vielzahl Thiamindiphosphat (ThDP)-abhängiger Enzyme mit hoher Stereoselektivität katalysiert. Da diese chiralen Verbindungen als Ausgangsstoffe zahlreicher organischer Synthesen dienen, war der Aufbau einer Plattform vielfältig substituierter und enantiokomplementärer 2-Hydroxyketone durch Bereitstellung eines Baukastens ThDP-abhängiger Enzyme Ziel dieser Doktorarbeit. Folgende Ergebnisse führten zur erfolgreichen Umsetzung der Aufgabenstellung:

1. Entwicklung eines **schnellen, reproduzierbaren Hochdurchsatz-Tests** zur Untersuchung der Produktbildung in Ganzzelltransformationen sowie einer Vielzahl gereinigter Enzyme.
2. **Vergrößerung des Enzymbaukastens** durch Hinzufügen von sechs Wildtypenzymen und vier Varianten mit Carboligaseaktivität. Die biochemische Charakterisierung des Substratspektrums der Decarboxylase- sowie Carboligaseaktivität inklusive Bestimmung optimaler Reaktionsbedingungen wurde für fast alle neuen Enzyme durchgeführt.
3. Durch die neugewonnenen Biokatalysatoren und den Einsatz des Hochdurchsatz-Tests konnte die **Plattform an diversen 2-Hydroxyketonen wesentlich erweitert** werden. Verzweigtkettige aliphatische 2-Hydroxyketone, substituierte (*R*)-Phenylacetyl-carbinol-analoge Produkte sowie verschiedene (*S*)-2-Hydroxyketone sind nun erstmalig enzymatisch zugänglich. Durch die Kombination der Decarboxylase- und Ligaseaktivität ist eine Ligation instabiler Aldehyde mit ThDP-abhängigen Decarboxylasen möglich.

Die (*S*)-selektive Carboligation wurde erst durch die Aufklärung der **Struktur-Funktionsbeziehungen** im aktiven Zentrum ermöglicht. Zusätzlich zu den bereits bekannten Strukturen konnten zwei weitere Wildtyp-**Kristallstrukturen** gelöst werden. Durch Übereinanderlagerungen aller Strukturen, Modellierungen und Berücksichtigung entstehender Produkte konnten katalytisch wichtige Regionen im aktiven Zentrum kartiert werden. Das Substratspektrum der Decarboxylase- als auch Carboligaseaktivität der Enzyme sowie ihre Chemo- und Enantioselektivität ist ein Ergebnis optimaler Substratstabilisierung im aktiven Zentrum. Diese Erkenntnis ermöglicht den Zugang zur Vorhersage gebildeter Produkte durch Modellierungen und somit zur Entwicklung von Katalysatoren mit gewünschter Ligaseaktivität.

LIST OF PUBLICATIONS

1. Gocke, D., Wendorff, M. & Pohl, M. A colorimetric assay for the detection of 2-hydroxy ketones. *submitted to Analytical Biochemistry*

2. Domínguez de María, P., Pohl, M., Gocke, D., Gröger, H., Trauthwein, H., Walter, L. & Müller, M. (2007). Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes. *European Journal of Organic Chemistry*, 2940-2944

3. Spaepen, S., Versées, W., Gocke, D., Pohl, M., Steyaert, J. & Vanderleyden, J. (2007). Characterization of phenylpyruvate decarboxylase, involved in auxin production of *Azospirillum brasilense. Journal of Bacteriology*, 189 (21), 7626-7633

4. Gocke, D., Berthold, C. L., Graf, T., Brosi, H., Frindi-Wosch, I., Knoll, M., Stillger, T., Walter, L., Müller, M., Pleiss, J., Schneider, G. & Pohl, M. Pyruvate decarboxylase from *Acetobacter pasteurianus*: biochemical and structural characterisation. *submitted to Biochimica et Biophysica Acta – Proteins and Proteomics*

5. 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. *Advanced Synthesis & Catalysis,* 349, 1425-1435

6. Berthold, C. L., Gocke, D., Wood, M. D., Leeper, F. J., Pohl, M. & Schneider, G. (2007). Crystal structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemo- and enantioselective carboligation reaction. *Acta Crystallographica, Section D: Biological Crystallography*, 63, 1217-1224

7. 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

LIST OF POSTER PRESENTATIONS

1. Gocke, D., Stillger, T., Knoll, M., Berthold, C., Schneider, G., Pleiss, J., Müller, M., Pohl, M. (2007). New *(S)*-2-hydroxy ketones by engineering the active site of benzoylformate decarboxylase. *8th International Symposium on Biocatalysis and Biotransformations,* Oviedo (Spain)

2. Gocke, D., Nguyen, C. L., Stillger, T., Müller, M., Pohl, M. (2007). A new enzyme for asymmetric carboligation. *8th International Symposium on Biocatalysis and Biotransformations,* Oviedo (Spain)

3. Kolter, G., Gocke, D., Knoll, M., Pleiss, J., Müller, M., Pohl, M. (2007). A new variant of the ThDP-dependent enzyme benzoylformate decarboxylase. *8th International Symposium on Biocatalysis and Biotransformations,* Oviedo (Spain)

4. Gocke, D., Stillger, T., Müller, M., Pohl, M. (2006). Carboligation of branched-chain aliphatic aldehydes with thiamine diphosphate (ThDP)-dependent enzymes. *3rd International Congress on Biocatalysis,* Hamburg-Harburg

5. Gocke, D., Graf, T., Wendorff, M., Eggert, T., Müller, M., Pohl, M. (2005). Isolation, cloning and characterisation of the novel thiamine diphosphate-dependent pyruvate decarboxylase from *Acetobacter pasteurianus*. *9th International SFB-Symposium*, Aachen

6. Wendorff, M., Dresen, C., Gocke, D., Henning, H., Jaeger, K.E., Kühl, S., Lütz, S., Müller, M., Pohl, M., Eggert, T. (2005). Benzoylformate decarboxylase from *Pseudomonas putida*: improvement of substrate specificity by directed evolution. *9th International SFB-Symposium*, Aachen

7. Dresen, C., Cosp, A., Walter, L., Gocke, D., Wendorff, M., Pohl, M., Müller, M. (2005). α,β-Unsaturated aldehydes as substrates for asymmetric C-C bond forming reactions with thiamine diphosphate (ThDP)-dependent biocatalysts. *9th International SFB-Symposium*, Aachen

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviations for enzymes

*Ab*PhPDC	phenylpyruvate decarboxylase from *Azospirillum brasilense*
*Ap*PDC	pyruvate decarboxylase from *Acetobacter pasteurianus*
*Bh*PDC	pyruvate decarboxylase from brewers yeast
*BjB*FD	benzoylformate decarboxylase from *Bradyrhizobium japonicum*
*Ll*KdcA	branched-chain ketoacid from *Lactococcus lactis*
*Pa*BFD	benzoylformate decarboxylase from *Pseudomonas aeruginosa*
*Pf*BAL	benzaldehyde lyase from *Pseudomonas fluorescens*
*Pp*BFD	benzoylformate decarboxylase from *Pseudomonas putida*
*Rp*BAL	benzaldehyde lyase from *Rhodopseudomonas palustris*
*Sc*PDC	pyruvate decarboxylase from *Saccharomyces cerevisiae*
*Sc*PhPDC	phenylpyruvate decarboxylase from *Saccharomyces cerevisiae*
*Sv*PDC	pyruvate decarboxylase from *Sarcina ventriculi*
yeast ADH	yeast alcohol dehydrogenase from yeast
*Zm*PDC	pyruvate decarboxylase from *Zymomonas mobilis*
*Zp*PDC	pyruvate decarboxylase from *Zymobacter palmae*

Further abbreviations

ATCC	american type culture collection
CAST	combinatorial active site mutagenesis
CD	circular dichroism
CLIB	cluster of industrial biotechnology
3,5-DChlBA	3,5-dichlorobenzaldehyde
DMH	4-hydroxy-(2,5-dimethyl)-hexan-3-one
DMO	5-hydroxy-(2,7-dimethyl)-octan-4-one
DMSO	dimethylsufoxide
DNA	desoxyribonucleic acid
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
E.C.	enzyme commission
ee	enantiomeric excess

ExPASy	expert protein analysis system
GC	gas chromatography
HPLC	high performance liquid chromatography
2-HPP	2-hydroxypropiophenone
IBA	isobutyraldehyde
INT	iodonitrotetrazolium chloride
IPTG	isopropyl-β-D-thiogalactopyranoside
IVA	isovaleraldehyde
LB-medium	Luria-Bertani medium
MTBE	methyl tertiary-butyl ether
MTT	thiazolylblue tetrazolium bromide
NADH	nicotinamide adenine dinucleotide
NMR	nuclear magnetic resonance
NTB	nitrotetrazolium blue
NTC	neotetrazolium chloride
PAC	phenylacetylcarbinol
pdb-code	RCSB protein data bank-code
PP-domain	pyrophosphate domain
PYR-domain	pyrimidine domain
R-domain	regulatory domain
STY	space time yield
ThDP	thiamine diphosphate
TTC	2,3,5-triphenyltetrazolium chloride
U	units

1 INTRODUCTION

1.1 Enzymes for industrial applications

1.1.1   Ancient and modern applications of biocatalysts

Already 6000 years ago the old Egyptians used microorganisms as yeast to bake bread and brew beer, although not being aware of the biological procedures occurring. It was only 150 years ago when the scientific discussion about the origin of enzymatic activity started and when Louis Pasteur and other scientists discovered microorganisms as the origin for several conversions taking place during food production. At the same time Pasteur pioneered the field of asymmetric synthesis by detecting optical activity as a consequence of molecular asymmetry (Fessner & Walter, 1997; Pasteur, 1848) and the ability of microorganisms to degrade just one enantiomer of a racemate while the other stays untouched (Turner, 1998). Emil Fischer could finally prove that not just whole cells but also enzymes are able to distinguish asymmetry of molecules (Fischer, 1894). This, as well as the successful application of cell free yeast extract for alcoholic fermentation by Buchner (Buchner, 1897), paved the way for the application of enzymes for biotransformations. In 1926 Sumner crystallised the first enzyme, urease, demonstrating the chemical origin of enzymes (Sumner, 1926).

With the discovery of the structure and the chemical composition of DNA and RNA by Watson and Crick in 1953 (Watson & Crick, 1953) the era of modern biotechnology started. Due to the development of recombinant gene technologies, enzymes became easily accessible for commercialisation, leading to a reduction of the time scales for industrial developments from 5-10 years to 1-2 years (Hodgson, 1994; Kirk *et al.*, 2002). Today most of the estimated 1.5 billion $ worldwide spent for industrial enzymes go into the detergents, starch, textile and fuel alcohol industries (Kirk *et al.*, 2002). But further applications in the sector of red biotechnology (medical sector), green biotechnology (focussing on genetically modified crops) and industrial biotechnology are strongly developing. Industrial biotechnology, nowadays called *white biotechnology* according to a decision of the EU industrial association EuropaBio (Sijbesma & Schepens, 2003), is based on biocatalysis and fermentation technologies combined with molecular genetics, enzyme engineering and metabolic engineering to produce chemicals, materials and energy (Soetaert & Vandamme, 2006). Ideally the chemical substances - like fine and bulk chemicals, pharmaceuticals, biocolourants, solvents, biopolymers, antibiotics, agro-chemicals, vitamins, food additives and biofuels - should be derived from renewable resources (BMBF, 2007; Dale, 2003).

The future prospects for the development of mild and more sustainable syntheses via biocatalysis are promising, primarily in those countries in which technical installations for alternative chemical processes do not yet exist (Festel *et al.*, 2004; Schulze & Wubbolts, 1999). According to a survey of FESTEL CAPITAL, interviewing 20 European chemical- and life science companies, the consulting- and investment agency expects that in 2010 approximately 20% of the chemical products will be gained via biotechnological processes which corresponds to 310 billion US$ of the current market value (Fig. 1-1) (Festel *et al.*, 2004). This development is the consequence of a decrease of production costs due to process simplification - one biotechnological step replaces several chemical reactions - and lower expenses for raw products, incidentals and disposals. Particularly the production of enantiomerically pure pharmaceuticals is currently emerging to a key technology (Panke *et al.*, 2004). 89% of all industrially applied biotransformations lead to chiral products used as fine chemicals (Straathof *et al.*, 2002).
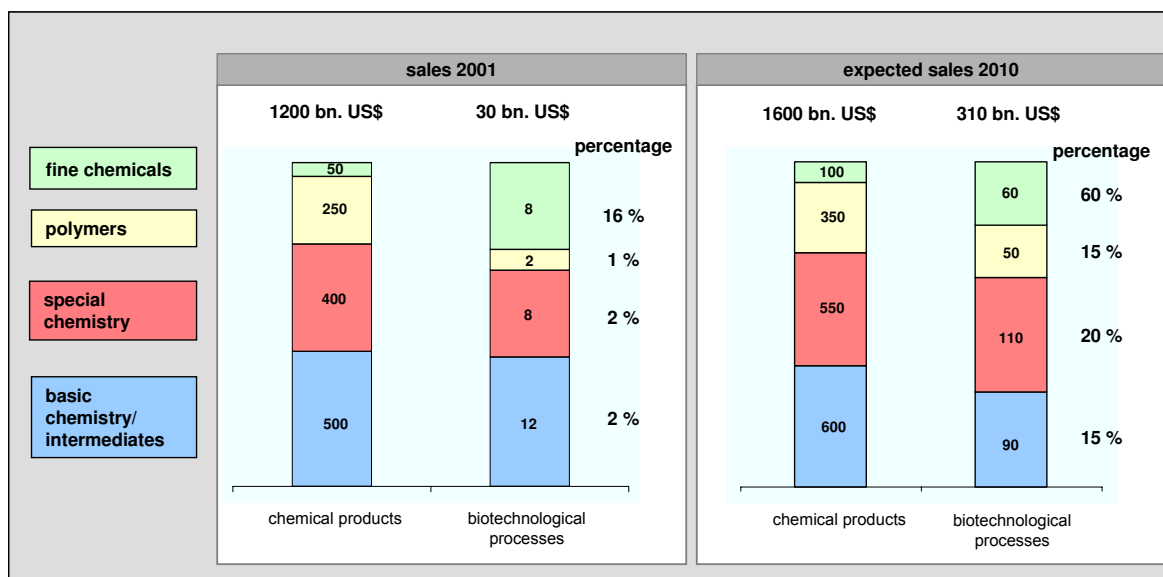
Fig. 1-1: Expected sales of chemical products produced via a biotechnological process (scheme according to Festel *et al.*, 2004).

In an overview about the current development in biotransformation Faber and Patel point out two focuses for new developments (Faber & Patel, 2000). One is the catalytic de-racemisation in which a combination of different enzymes should lead to 100 theoretical yield (Strauss *et al.*, 1999). The second focus is the development of biocatalysts for reactions which are very difficult to perform by classical chemical means, including e.g. catalytic decarboxylation, stereoselective oxidation as well as asymmetric C-C bond formation in aqueous solutions. Asymmetric carboligations are the key steps in many syntheses (Sukumaran & Hanefeld, 2005) and the target of this thesis.

1.1.2 Enantiomerically pure drugs

Although enantiomers possess identical chemical and physical properties, their biological effects as flavour compounds, crop protection products, pheromones and drugs can differ immensely (Rouhi, 2003). As in many cases just one enantiomer is biologically active or exhibits the desired effect, the production of racemic compounds is connected with the formation of 50% waste. Especially in pharmacology, negative consequences of the "wrong enantiomer" make the implementation of optically pure substances in the production of drugs indispensable. In 1992 the US-Food and Drug Administration (FDA) enacted a law to further permit racemates on the market but not without a primarily testing phase of both enantiomers as well as the racemate (Breuer *et al.*, 2004). As a consequence of this costly and time consuming test procedure, most companies decided to implement new chiral products in optically pure form (Faber, 2000; Margolin, 1993). Hence, since 2001 no new racemic drugs have been placed on the market (Agranat *et al.*, 2002; Rouhi, 2003). There are principally three possible ways to obtain chiral substances:

1. Isolation of chiral precursors from natural sources
2. Separation of racemates (e.g. by chromatography, kinetic resolution, deracemisation)
3. Asymmetric synthesis

Asymmetric synthesis is one of the most effective methods since chiral molecules are formed selectively from usually cheap achiral precursors. A further advantage is the possible yield of 100%, whereas e.g. the separation of racemates are usually limited to 50% yield.

1.1.3 Advantages of enzymes

Compared to classical organic synthesis the application of enzymes provides many advantages:

1. *Effective catalysis:* In contrast to non-catalysed reactions turnover numbers can be increased up to 10^8-10^{12} times, in some cases even up to 10^{19} times (Frey & Hegeman, 2006; Sukumaran & Hanefeld, 2005). Therefore the results are much higher than the acceleration mediated by most of the chemical catalysts (Azerad, 1995).
2. *Mild reaction conditions:* Using biotransformation, extreme pressures, high temperatures and the employment of organic solvents can be avoided. This is not only an advantage for the transformation by minimising side reactions like isomerisation, racemerisation, epimerisation and rearrangement (Patel, 2001) but also diminishes risks for the employees.

3. *Versatile applications:* There is probably an enzyme-catalysed analogue for almost every existing or applied organic reaction (Azerad, 1995). As more than 35 000 enzyme-catalysed reactions are known, the biological variety and the potential of biocatalytic applications are immense (Straathof *et al.*, 2002). Furthermore, multi-enzyme-reactions can be implemented since many enzymes are active under similar conditions. This can either simplify a process or switch an unfavourable equilibration towards the desired product (Faber, 2000).

4. *Environmental compatibility:* In contrast to e.g. heavy metals used in chemical syntheses, enzymes are completely bio-degradable (Faber, 2000).

5. *Selectivity*: The predominant advantage of many enzymes is their high selectivity, including chemo-, regio- and enantioselectivity (Azerad, 1995). These parameters can further be adjusted by reaction engineering and molecular biological methods. Especially for the synthesis of chiral fine chemicals enzymes are superior to chemical catalysts (Thayer, 2006). Enantioselective synthesis avoids the purification of the desired enantiomer and therewith the production of waste.

1.1.4 Enzyme classes

The term *enzyme* ("in yeast") was first described in 1876 by W. Kühne (Kühne, 1876). Enzymes are also called biocatalysts due to their ability to form an enzyme-substrate-complex thereby decreasing the activation energy of a reaction (Schlee & Kleber, 1991). The activity of enzymes is given in *Units*. According to an agreement of the *International Union of Biochemistry* enzymes are divided into six classes due to their reaction specificity (Tab. 1-1). Every enzyme possesses its own four-digit E.C. (enzyme commission) number starting with the number of the enzyme class, followed by three digits sub-classifying the enzyme according to its type of chemical reaction, the substrate and, if necessary, the required coenzyme.

Tab. 1-1: Enzyme classes and their reaction specificities (Held *et al.*, 2000; Pohl, 2000).

enzyme class	catalysed reaction (examples of subclasses)
1. oxidoreductases	redox reactions (dehydrogenases, reductases, oxidases, monooxygenases)
2. transferases	transfer of functional groups (transketolases, transaminases, acyltransferases)
3. hydrolases	hydrolysis of functional groups, e.g. C-O, C-N (proteases, lipases, esterases, amidases, glycosidases, acylases, amylases)
4. lyases	non-hydrolytic cleavage of e.g. C-C, C-O and C-N under double bound formation; binding of groups to double bonds (decarboxylases, aldolases, dehydratases)
5. isomerases	structural or geometrical rearrangement (racemases, epimerases, isomerases, tautomerases, mutases)
6. ligases	ligation of two substrates under simultaneous cleavage of a pyrophosphate of a nucleoside triphosphate (carboxylases)

1.1.5   Bottlenecks for the application of enzymes and possible solutions

A big challenge to increase the use of enzymes in chemo-enzymatic syntheses is changing the mind-set of organic chemists not to be worried about using enzymes and try them as a matter of choice. Since successful chemistry has been done for decades, an alternative approach is often just tested if all chemical syntheses failed and biotransformation is not immediately included in the set up of new synthetic pathways (Thayer, 2006). To enable the easy use of enzymes, some companies sell ready to use biocatalysts (e.g. Biocatalytics, Pasadena, CA) which can be applied easily also by unskilled customers.

A further limitation for the broader application of biocatalysts is the fact that new developments in biocatalysis might take several months, which is much too long for companies who have to answer customers' demands usually within a few weeks. As a consequence, the time consuming development of products up to their readiness for marketing led to a loss of investors especially in Biotech companies in 2006 (Hoffritz, 2007). Time- and money consuming fundamental research will always be important in order to find and understand fundamental mechanisms as a basis for later applications. Besides, less time consuming product developments are imperative. This can be achieved e.g. by a close cooperation between academic research and industry having the power to further develop the catalysts up to readiness for marketing. Currently there are several efforts pointing into this direction, e.g. the Cluster of Industrial Biotechnology CLIB 2021, a German consortium of big pharma- and start up companies and universities (CLIB2021,

2007), or the European Technology Platform for Sustainable Chemistry (SusChem Solutions, 2007).

Rather than adapting a process to a known enzyme, the screening for an appropriate catalyst or the enzyme's optimisation towards the process by genetic modifications is often much more efficient. Therefore many pharmaceutical companies have started the buildup of strain- and enzyme-collections for in-house screening and development. This is often realised in cooperation with biotechnology companies as seen between BASF (Ludwigshafen) and Verenium (Cambridge, GB; formerly Diversa, San Diego, CA), Evonik Industries (Essen; formerly Degussa, Hanau) and BRAIN AG (Zwingenberg) and DSM (Geleen, NL) and the IEP GmbH (Wiesbaden) (Thayer, 2006).

1.2   Technology platform for chiral 2-hydroxy ketones

1.2.1   Toolbox concept

The establishment of biocatalyst platforms allows fine chemical producers to react much more quickly to customer's requests with an efficient and economic performance. Although the implementation of such a toolbox can be expensive, variable applications make it profitable. Due to their high selectivity, bioprocesses often are the matter of choice if enantiomerically pure agents of 99.5% are required (Thayer, 2006). Beside the implementation of a strain and enzyme collection for the effective screening of desired products, many companies started to build up additional platforms of chiral building blocks and intermediates to provide rapid access for the production of complex chiral compounds. DSM follows this strategy by combining chemically as well as biocatalytically synthesised intermediates under the ChiralTree[®] label. With its ChiPros platform BASF built up a toolbox of intermediates with high optical purity exclusively obtained via biocatalysis (Thayer, 2006).

A similar strategy is followed in this thesis by the development of an enzymatic toolbox encompassing thiamine diphosphate (ThDP)-dependent enzymes which catalyse a benzoin-condensation like carboligase reaction of aldehydes yielding chiral 2-hydroxy ketones. Using the overlapping substrate ranges of various wild type enzymes and variants thereof, a toolbox of differently substituted aromatic, aliphatic and araliphatic 2-hydroxy ketones should be accessible (Fig. 1-2). The relevance of 2-hydroxy ketones as building blocks for the organic synthesis and pharmaceutical chemistry will be discussed in chapter 1.3.1.
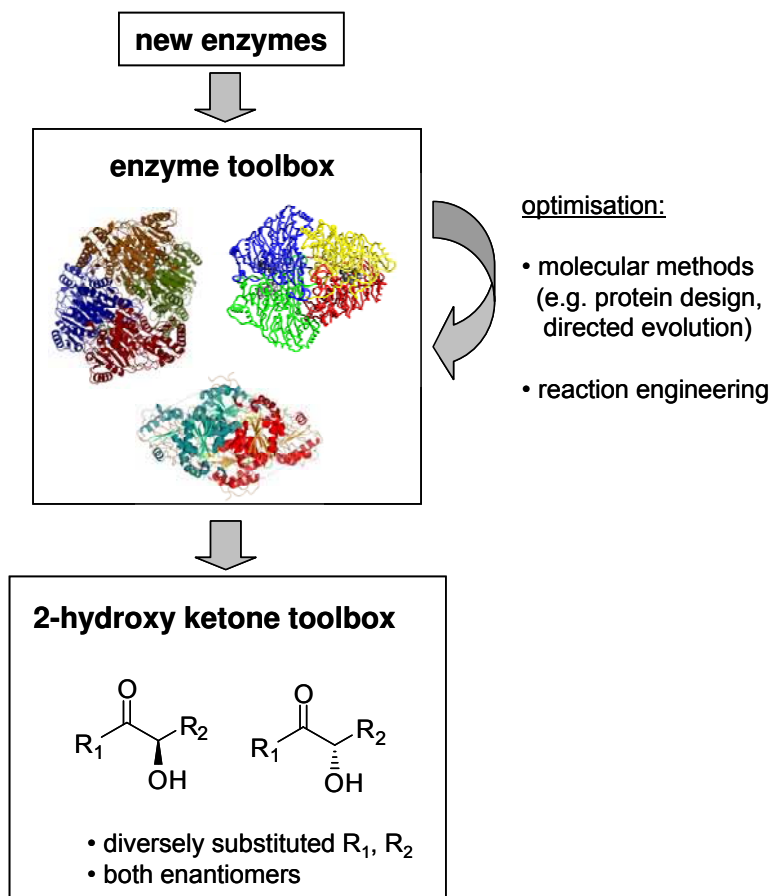
Fig. 1-2: Design of a 2-hydroxy ketone toolbox by implementation of an enzyme toolbox with C-C coupling activity. The enzyme toolbox should comprise new wild type enzymes and variants thereof.

1.2.2   Extension of the enzymatic toolbox

Multiple methods are nowadays available to identify and engineer suitable biocatalysts with high catalytic activities and stabilities (Otten & Quax, 2004):

1. *Database mining*: The so-called database mining approach (Wackett, 2004) is based upon the alignment of amino acid sequences of enzymes with known activity and sequences of unknown function from databases. Since the search tools are easily applicable and effective (BLAST, 2007), it is a powerful tool to identify sequences with putative similar activity, especially when a sequence motif for the identification of certain enzymes is highly conserved, e.g. the G-D-G-(X)-N-motif in the binding site of ThDP-dependent enzymes (Hawkins *et al.*, 1989). This approach gives first hints to novel enzymes with a desired function. However, care has to be taken as enzymes with similar sequences do not necessarily exhibit similar activities (Toscano *et al.*, 2007). Therefore putative activities must always be proven biochemically. Furthermore, if studies aim to identify enzymes with a diverse substrate range, larger sequence variability might be important in order to gain the required diversity.

2. *Promiscous activity of known enzymes:* Catalytic promiscuity, defined as the ability of a single active site to catalyse more than one reaction, is more widespread than often recognised. It has a natural role in evolution and is occasionally used in the biosynthesis of secondary metabolites (Kazlauskas, 2005). For several ThDP-dependent 2-keto acid decarboxylases a carboligase side activity is known beside their physiological decarboxylase activity (chapter 1.4). Therefore, the investigation of promiscuous activity of enzymes with known decarboxylase activity is most promising. Additionally even weak promiscuity can be enhanced immensely by small modifications in the active site (e.g. by rational design) (Bornscheuer & Kazlauskas, 2004; Toscano et al., 2007).

3. *Metagenome search:* Probably up to 99.8% of the existing microbes are not readily cultivable and therefore cannot be exploited. The *metagenome technology* offers a solution by developing culture-independent methods to isolate, clone and express environmental DNA. This enables the mining of microbial diversity, gives access to genomes, allows identification of protein coding sequences and might even allow biochemical pathway construction (Ferrer *et al.*, 2005; Lorenz & Eck, 2005; Streit *et al.*, 2004). The approach was successfully used for the detection of benzoylformate decarboxylases (Henning *et al.*, 2006).

4. *Directed evolution:* Various techniques have been developed to introduce random mutations and screen designed libraries for desired activities, like mimicking evolution on a laboratory scale (Dalby, 2003). Directed evolution is a powerful tool especially if e.g. no structural information of the catalyst is available or in case of uncertainty which enzyme trait has to be changed in order to gain a desired ability (Arnold, 1998). The technique is especially valuable to adapt enzymes to industrial processes (Reetz, 2006). This approach has successfully been applied to adapt the substrate range and the stability of some ThDP-dependent enzymes (Lingen *et al.*, 2002; Lingen *et al.*, 2003; Wendorff, 2006).

5. *Rational protein design:* Rational protein design is the oldest but still up to date way to improve enzyme properties. It requires detailed knowledge of the protein structure and the underlying molecular principles of the properties to be altered. If certain amino acid side chains related to the targeted effect can be identified, this method is a very important tool to understand structure-function relationships, especially in combination with structural and biochemical investigations. Several new crystal structures and the possibility of molecular modelling to test a possible substrate range in advance (e.g. via docking experiments) or the simulation of variants simplify the application of site-directed mutagenesis (Kazlauskas, 2000).

Moreover the given techniques can be combined to increase the effects. A much discussed combination of single-site saturation mutagenesis and simultaneous randomisation at multiple sites is the combinatorial active site mutagenesis (CAST) (Reetz *et al.*, 2005).

1.2.3   Reaction engineering

Beside alteration of enzymatic properties by sequence diversity, the adjustment of optimal reaction conditions can significantly influence the amount and the purity of accessible products. Reaction engineering approaches are especially useful for two-substrate reactions, as the variation of the relative substrate concentrations might influence selectivities and therewith the product range of biotransformations. As an example the enantioselectivity of benzoylformate decarboxylase from *Pseudomonas putida* was found to be a function of the temperature and the concentration of benzaldehyde in mixed carboligations of benzaldehyde and acetaldehyde (Iding *et al.*, 2000). Several biotransformations using ThDP-dependent enzymes were investigated in batch- and continuous reaction systems (Goetz *et al.*, 2001; Hildebrand *et al.*, 2007; Stillger *et al.*, 2006).

1.3   2-Hydroxy ketones – versatile precursors

1.3.1   Chiral 2-hydroxy ketones as building blocks

2-Hydroxy ketones are versatile building blocks for the pharmaceutical and chemical industry (Ward & Singh, 2000). Their bifunctional nature and the presence of one stereo centre as well as one prochiral carbonyl group makes them amenable to further synthetic transformations (Sukumaran & Hanefeld, 2005). By enantioselective reduction or reductive amination of the keto group chiral diols and amino alcohols are accessible. Both products find versatile applications for the synthesis of chiral ligands and chiral auxiliaries (Fig. 1-4) (Kihumbu *et al.*, 2002; Stillger, 2004).

The first and most famous commercial application of a 2-hydroxy ketone used as a building block for a pharmaceutical agent was already dated in the early 1930's, when (*R*)-phenylacetylcarbinol ((*R*)-PAC) was produced by fermenting yeast in the presence of benzaldehyde (Neuberg & Hirsch, 1921). In a subsequent reductive amination (*R*)-PAC was further converted to *L*-ephedrine (Fig. 1-3), once widely used as a bronchodilator for asthma treatment. Nowadays, *L*-ephedrine is mostly replaced by pseudoephedrine, showing less adverse affects, but still misused as a party drug and for athlete's doping due to its performance enhancing effects (DSHS Köln, 2007).
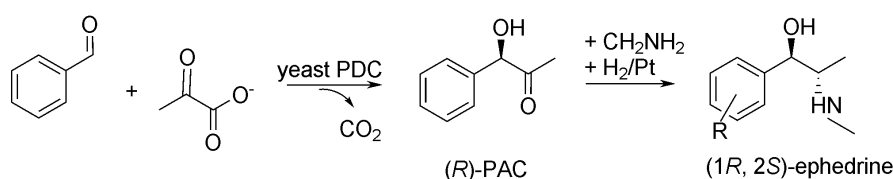
Fig. 1-3: Chemo-enzymatic synthesis of *L*-ephedrine by yeast fermentation in the presence of benzaldehyde. Pyruvate is generated from glucose and decarboxylated to acetaldehyde by yeast PDC, which also catalyses the chiral carboligation step towards (*R*)-PAC (according to Neuberg & Hirsch, 1921; Stillger, 2004).

Furthermore (*R*)-PAC is a key intermediate for the synthesis of norephedrine, adrenaline, amphetamine, methamphetamine, phenylpropanolamine and phenylamine (Shukla & Kulkarni, 2000). It is also used as a chiral ligand for asymmetrical chemical catalysis, e.g. the addition of dialkylzine reagents to aldehydes (Trost, 2004).

Figure 1-4 gives an overview about several interesting target molecules which can be gained from 2-hydroxy ketones. Nitidanine (**a**), up to now still derived from the bark of the medical plant *Silybum marianum,* is effective in the treatment of liver diseases (Morazzoni & Bombardelli, 1995), (-)-cytoxazone (**b**) is applied as an inhibitor of the type I cytokine production in immunotherapy (Kakeya *et al.*, 1999), while 5-methoxyhydnocarpin (**c**) inhibits the multi drug pump in microorganisms (Stermitz *et al.*, 2000). Especially from substituted (*R*)-hydroxypropiophenones ((*R*)-HPP), accessible by some ThDP-dependent enzymes (chapter 1.5), further active pharmaceutical ingredients can be synthesised. (*R*)-1-(3-chlorophenyl)-2-hydroxypropanone is a precursor of bupropion (**d**) in the drugs Wellbutin® and Zyban® (Glaxo Wellcome) used as antidepressants (Fang *et al.*, 2000). Another antidepressant agent is 1555U88 derived from (*R*)-1-(3,5-difluoro-phenyl)-2-hydroxypropanone (**e**) (Stillger, 2004). Finally from (*R*)-1-(2,4-difluorophenyl)-2-hydroxypropanone the fungicide Ro09-3355 (**f**) can be obtained (Leuenberger *et al.*, 1999).
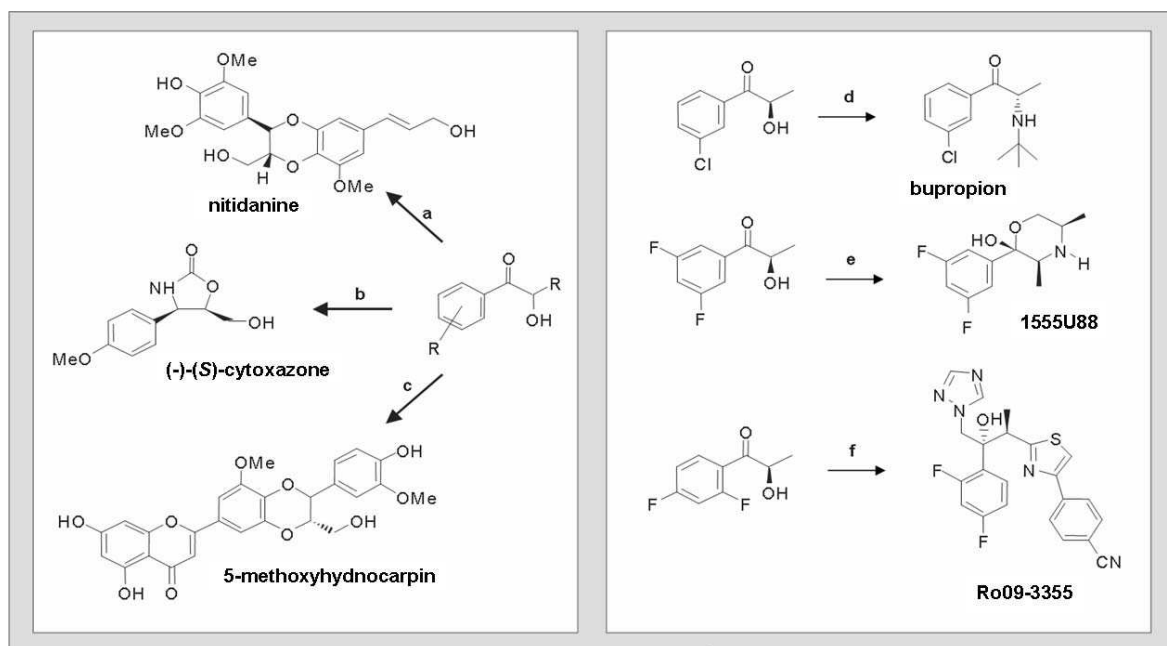
Fig. 1-4: Pharmaceutically interesting compounds which can be deduced from chiral 2-hydroxy ketones as building blocks (scheme according to Stillger, 2004).

1.3.2   Synthesis of chiral 2-hydroxy ketones

Several chemical syntheses have been described to obtain chiral 2-hydroxy ketones. Among these are the enantioselective oxidation of enolethers (Curci *et al.*, 1996), the

regio- and enantioselective reduction of 1,2-diketones (Toukoniitty *et al.*, 2001) and the ligation of aldehydes in a benzoin-condensation-like reaction (Wöhler & Liebig, 1832) using thiazolium salts as catalysts (Ukai *et al.*, 1943). Whereas these methods result in low enantioselectivities (*ee*s < 90%), methods employing chiral, bicyclic thiazolium salts for C-C coupling have been described yielding higher *ee*s up to 95% (Enders & Kallfass, 2002; Knight & Leeper, 1998).

Thus most of the chemical methods are not competitive with biotransformations. By choosing appropriate biocatalysts and reaction conditions, an enantio- and regioselective synthesis of 2-hydroxy ketones via biotransformation yielding enantiomeric excesses (*ee*s) up to 99% is possible. Several ways are currently known:

1. *Reduction* of 1,2-diketones by whole yeast cell catalysis (*ee*s up to 98%) (Nakamura *et al.*, 1996).
2. *Stereoselective oxidation* of cis-1,2-diols by NAD-dependent glycerol dehydrogenases (*ee* 99%) (Lee & Whitesides, 1986).
3. *Kinetic resolution* by either (*R*)-enantioselective reduction of racemic 2-hydroxy ketones with glycerol dehydrogenases (Lee & Whitesides, 1986), or by transforming an enantioselective acylation or a selective hydrolysis of the 2-hydroxy ketone with lipases (*ee* 66-96%) (Adam *et al.*, 1999). Furthermore benzaldehyde lyase (chapter 1.5.5) catalyses the kinetic resolution of racemic benzoin in the presence of acetaldehyde by transforming only the (*R*)-enantiomer to (*R*)-2-hydroxy-propiophenone, leaving the (*S*)-product with an *ee* > 99% (Demir *et al.*, 2001).
4. *Asymmetric synthesis* of chiral 2-hydroxy ketones by C-C coupling of aldehydes. This type of biochemical synthesis can be catalysed with high chemo- and stereoselectivity (*ee* > 99%) by several ThDP-dependent enzymes in a single step from cheap aldehyde substrates and is further exploited in this thesis (chapter 1.4.5).

1.4   ThDP-dependent enzymes

1.4.1   Overview

Thiamine diphosphate (ThDP)-dependent enzymes are widely spread among several central points of anabolism and catabolism (e.g. pentose-phosphate pathway, tricarboxylic acid pathway) in which ThDP mediates the formation and cleavage of C-C, C-S, C-N, C-O and C-P-bonds (Frank *et al.*, 2007; Pohl *et al.*, 2004). This group of enzymes can be roughly categorised into enzymes with decarboxylase activity and those with transferase activity. Both reaction types rely on the same mechanism with ThDP as the reactive agent

(chapter 1.4.4). Among those enzymes are transketolase (E.C. 2.2.1.3), 2-oxo acid dehydrogenase (E.C. 1.2.4.X), pyruvate oxidase (E.C. 1.2.3.3), pyruvate ferredoxin oxido-reductase (E.C. 1.2.7.1), deoxyxylulose 5-phosphate synthase (E.C. 2.2.1.7), aceto-hydroxyacid/acetolactate synthase (E.C. 2.2.1.6), N^2-2-carboxyethylarginine synthase (E.C. 2.5.1.66), sulfoacetaldehyde transferase (E.C. 2.3.3.15), glyoxylate carboligase (E.C. 4.1.1.47) and most important for this work, benzaldehyde lyase (E.C. 4.1.2.38) and 2-keto acid decarboxylases (E.C. 4.1.1.X) (Frank *et al.*, 2007; Pohl *et al.*, 2004). Currently ten entries of ThDP-dependent 2-keto acid decarboxylases can be found in the ExPASy (Expert Protein Analysis System) proteomic server (Gasteiger *et al.*, 2003). The most important ones for this work are given in Tab. 1-2.

Tab. 1-2: ThDP-dependent 2-keto acid decarboxylases important for this work.

enzyme (abbreviation)	E.C.-number	physiological activity	first references
pyruvate decarboxylase (PDC)	4.1.1.1	alcoholic fermentation	(Holzer *et al.*, 1956)
benzoylformate decarboxylase (BFD)	4.1.1.7	metabolism of mandelate	(Hegeman, 1966)
phenylpyruvate decarboxylase (PhPDC)	4.1.1.43	metabolism of phenylalanine	(Costacurta *et al.*, 1994)
indole-3-pyruvate decarboxylase (In3PDC)	4.1.1.74	metabolism of tryptophane	(Asakawa *et al.*, 1968)
branched-chain 2-keto acid decarboxylase (KdcA)	4.1.1.72	metabolism of leucine	(Smit *et al.*, 2005a)

Further entries in the ExPASy proteomic server are as follows: oxalyl-CoA decarboxylase (E.C. 4.1.1.8), 2-oxoglutarate decarboxylase (E.C. 4.1.1.71), 5-guanidino-2-oxopentanoate decarboxylase (E.C. 4.1.1.75), sulfopyruvate decarboxylase (E.C. 4.1.1.79), phosphono-pyruvate decarboxylase (E.C. 4.1.1.82).

1.4.2   Thiamine diphosphate

As the name already indicates ThDP-dependent enzymes require the cofactor thiamine diphosphate (ThDP) (Fig. 1-5), the biologically active form of vitamin B1 (thiamine), for their biological activity. In contrast to yeasts, plants and bacteria, mammals are not able to synthesise thiamine and thus have to assimilate it by a balanced diet to produce the vitally important ThDP (Bettendorff & Wins, 2004; Sauermost, 1994). Vitamin B_1 is essential for the metabolism of carbohydrates (to produce energy) and for nerve and heart functions. A consequence of thiamine deficiency in humans is called beriberi.
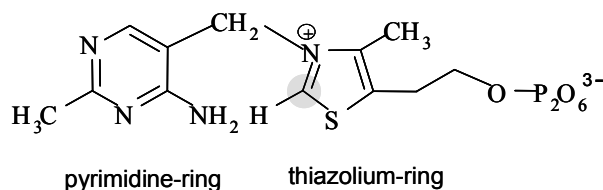
Fig. 1-5: Structure of ThDP. It comprises of a five-membered thiazolium ring with a diphosphate-terminated side chain and a six-membered 4'-amino pyrimidine ring. The catalytically important C2-atom of the thiazolium ring is highlighted grey.

Furthermore, ThDP-dependent enzymes require divalent cations, mainly Mg^{2+}, which function as an anchor for the diphosphate group of ThDP in the active sites (Dobritzsch *et al.*, 1998).

1.4.3   Overall Structure

Although sequence similarities among ThDP-dependent enzymes are usually less than 20%, many of the tertiary structures are remarkably similar. Most likely the common binding fold emerged by divergent evolution with a selective pressure on the geometric position of the small number of residues binding ThDP, but tolerating other substitutions elsewhere (Frank *et al.*, 2007).

Secondary structure
As demonstrated in Tab. 3-11 the crystal structures of several 2-keto acid decarboxylases as well as the benzaldehyde lyase from *Pseudomonas fluorescens* (*Pf*BAL) have been solved. All of them show a general fold consisting of three domains of α/β-topology (Muller *et al.*, 1993). Among these the PYR- (pyrimidine or α) and PP- (pyrophosphate or γ) domains are topologically equivalent while the so-called R- (regulatory or ß) domain has a slightly different fold. Each domain consists of a central β-sheet surrounded by a varying number of α-helices. As an example Fig. 1-6 shows the secondary structure of the benzoylformate decarboxylase from *Pseudomonas putida* (*Pp*BFD).
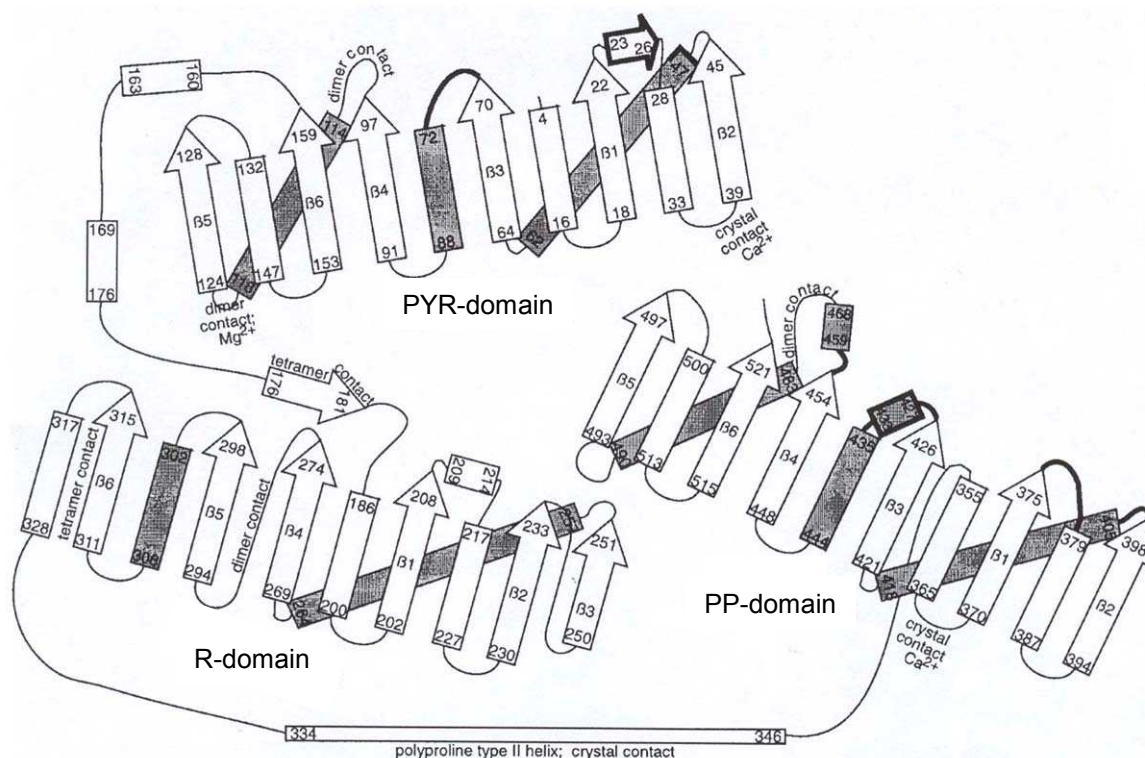
Fig. 1-6: Schematic presentation of the secondary structure of benzoylformate decarboxylase from *Pseudomonas putida* (*Pp*BFD). β-Strands are indicated as arrows, α-helices as rectangles; white helices are situated in front of β-sheets, shaded helices lay behind. Contacts between dimers and tetramers are labelled (according to Hasson *et al.*, 1998).

Quaternary structure

Two monomers form a dimer with two ThDPs bound at the dimer's interface. The pyrimidine ring of the cofactor interacts with the PYR-domain of one subunit, whereas the residual part interacts with the PP-domain of the neighbouring subunit (Dobritzsch *et al.*, 1998; Hasson *et al.*, 1998; Mosbacher *et al.*, 2005). ThDP is bound in a conserved V-shape (Fig. 1-7, **publication VI**), forced by a large hydrophobic residue next to the thiazolium ring (Lindqvist *et al.*, 1992).

ThDP-dependent enzymes are mostly active either as dimers, e.g. acetohydroxyacid synthases (Pang *et al.*, 2002) and transketolases (Lindqvist *et al.*, 1992), or tetramers, e.g. most 2-keto acid decarboxylases (Arjunan *et al.*, 1996; Dobritzsch *et al.*, 1998; Dyda *et al.*, 1993; Hasson *et al.*, 1998; Mosbacher *et al.*, 2005; Versées *et al.*, 2007). The tetramer can be best described as a dimer of dimers (Duggleby, 2006). Beside the tetrameric structure some PDCs have the tendency to form higher association states such as octamers (Pohl *et al.*, 1994); in some plants oligomers up to 300-500 kDa have been observed (König, 1998).

The distortion angle of two dimers forming one tetramer has been related to the kinetic behaviour of the enzymes. In the case of yeast PDCs, which are activated by their substrate pyruvate, the dimers change their relative position upon binding of the activator (König,

1998). By contrast *Zm*PDC and *Pp*BFD, which originate from bacterial sources, do not show such cooperative effects in the presence or absence of substrates.

1.4.4 Lyase and decarboxylase activity

ThDP-dependent enzyme reactions start with a chemically challenging breaking of a C-C or C-H bond adjacent to a carboxyl group in the substrate (Fig. 1-8 b/c) (Frank *et al.*, 2007). To mediate this activity the C2-carbon of the thiazolium ring ThDP first needs to be activated by deprotonation to form a potent nucleophilic *ylide* (Breslow, 1957; Kern *et al.*, 1997) (Fig. 1-7 B/C). This deprotonation is supported by a conserved glutamate residue present in all ThDP-dependent enzymes, which induces the formation of an 1',4'-imino tautomer in the pyrimidine ring. The V-conformation of the cofactor is essential for this step since the C2-proton is positioned at a reactive distance to the 4'-imino group of the neighbouring pyrimidine ring (Fig. 1-7) (Frank *et al.*, 2007; Lindqvist *et al.*, 1992).
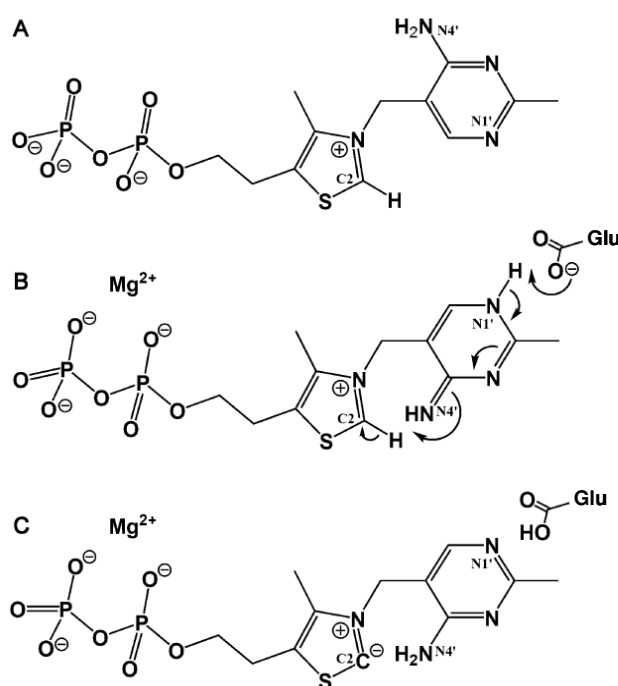
Fig. 1-7: Chemical structure of ThDP in solution (**A**) and in its activated form (**B+C**) (according to Frank *et al.*, 2007). In the active sites ThDP is bound in a highly conserved V-conformation, placing the N4' of the pyrimidine ring next to C2 of the thiazolium ring. The rare imino-form of the pyrimidine ring occurs (**B**) before the C2-proton is abstracted by the N4'-of the pyrimidine ring and relayed to the conserved glutamate yielding the activated ylide form (**C**).

As the first step of a **decarboxylase reaction** (Fig. 1-8, left cycle) the carbonyl group of the 2-keto acid substrate reacts with the ylide-form of ThDP (**b**) yielding the corresponding 2-hydroxy acid adduct. After CO_2 is released, a carbanion-enamine, the so called *active aldehyde,* is formed as a reactive intermediate (**c**). Subsequently the enamine intermediate is protonated (**d**) and the corresponding aldehyde is released, thereby reconstituting the ylide (Frank *et al.*, 2007; Kluger, 1987).

Carboligation reactions result from the binding of a carbonyl compound such as an aldehyde as the second substrate (acceptor aldehyde), leading to the formation of 2-hydroxy ketones (**e**). In this case, again a proton donor is required to neutralise the resulting negative charge. Therefore all ThDP-dependent enzymes need a proton acceptor during the first part of the reaction cycle and a proton donor during the second part. The proton acceptor for the first step is the 1',4'-imino tautomer of ThDP in combination with the invariant glutamate residue (Fig. 1-7 B). The proton relay system for the second step is different in the enzymes. As was shown for the pyruvate dehydrogenase E1 component (Frank *et al.*, 2004), two histidine residues are operating as proton relays: one situated next to the diphosphate moiety of ThDP and the other close to the C2-atom. In PDCs (Frank *et al.*, 2007), *Pp*BFD (Kneen *et al.*, 2005; Polovnikova *et al.*, 2003; Siegert, 2000) and *Pf*BAL (Kneen *et al.*, 2005) the thiazolium-proximal histidines are positional conserved while other ThDP-dependent enzymes contain a polar residue (Arg or Gln) in these positions (Frank *et al.*, 2007). Since binding of an aldehyde as the donor substrate without a preliminary decarboxylation step is also possible, the decarboxylation of a 2-keto acid is not mandatory for carboligation.
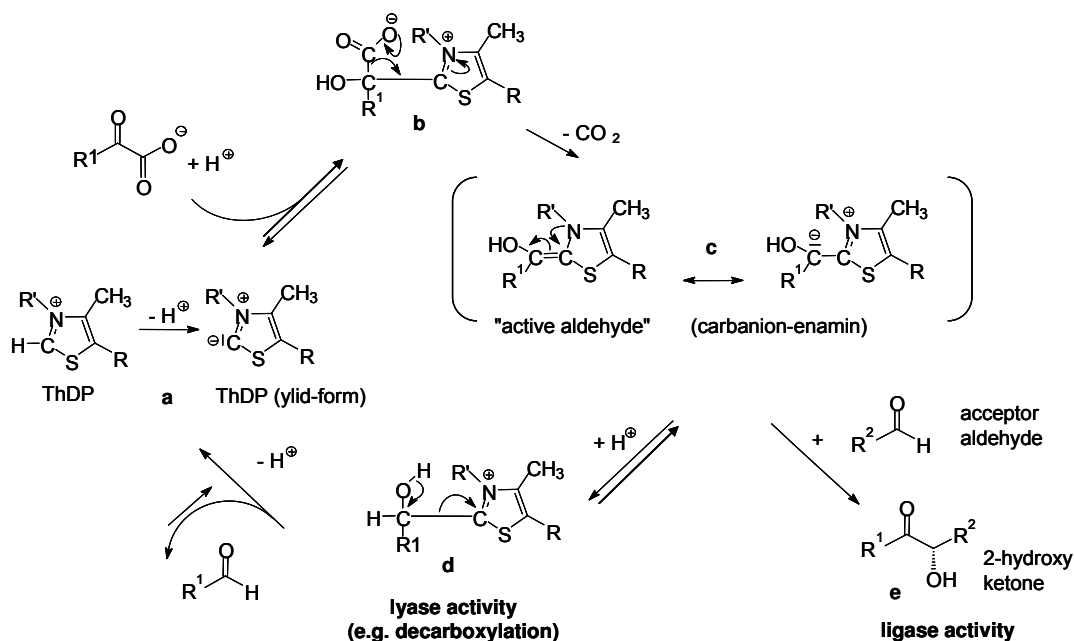
Fig. 1-8: Reaction mechanism of the lyase and ligase activity of 2-keto acid decarboxylases (according to Siegert *et al.*, 2005).

Conclusively this reaction mechanism explains the two possible products which can occur: if the acceptor is a proton, a simple decarboxylase reaction takes place and an aldehyde is released. This is the physiological activity of ThDP-dependent decarboxylases in the metabolism. If the acceptor is an aldehyde, a carboligation reaction can occur forming a 2-hydroxy ketone. An explanation for the physiological role of this promiscuous activity has not been found so far.

As both reactions take place at the same active site, the steric and electronic properties of the active centre influence both reactions similarily. Therefore, the investigation of the substrate range of decarboxylation is of interest in order to deduce information about the acyldonor (first substrate) spectrum for the carboligase activity of ThDP-dependent enzymes. If a respective 2-keto acid is a substrate for the decarboxylation the binding of the corresponding aldehyde to the C2-atom of ThDP located in the active centre is most likely, meaning that this aldehyde may be also a possible donor aldehyde in enzyme catalysed carboligation reactions.

1.4.5   Carboligase activity

As described above, carboligation can take place starting either from aldehydes or from the corresponding 2-keto acids. The first substrate bound to ThDP is called the *donor* substrate, whereas the second one is the *acceptor* substrate. If both substrates are the same, only one product can be gained, which can occur in the (*R*)- or (*S*)-configuration or a mixture of both (Fig. 1-2). The situation is more complex in case of mixed carboligations of two different substrates.

Chemoselectivity

The simplest mixed carboligation of an aromatic and an aliphatic aldehyde, which is catalysed by several ThDP-dependent enzymes and already well investigated (Dünkelmann *et al.*, 2002; Lingen *et al.*, 2004) is shown in Fig. 1-9. Starting from benzaldehyde and acetaldehyde, four different products can occur, each of them stereoselective in the (*R*)- or (*S*)-configuration or as a mixture of both. Acetoin is obtained by the self-ligation of acetaldehyde while benzoin results from the coupling of two benzaldehyde molecules. If a mixed product appears it can be either phenylacetylcarbinol (PAC), obtained from acetaldehyde as the donor and benzaldehyde as the acceptor, or 2-hydroxypropiophenone (2-HPP), resulting from benzaldehyde as the donor and acetaldehyde as the acceptor (Fig. 1-10).
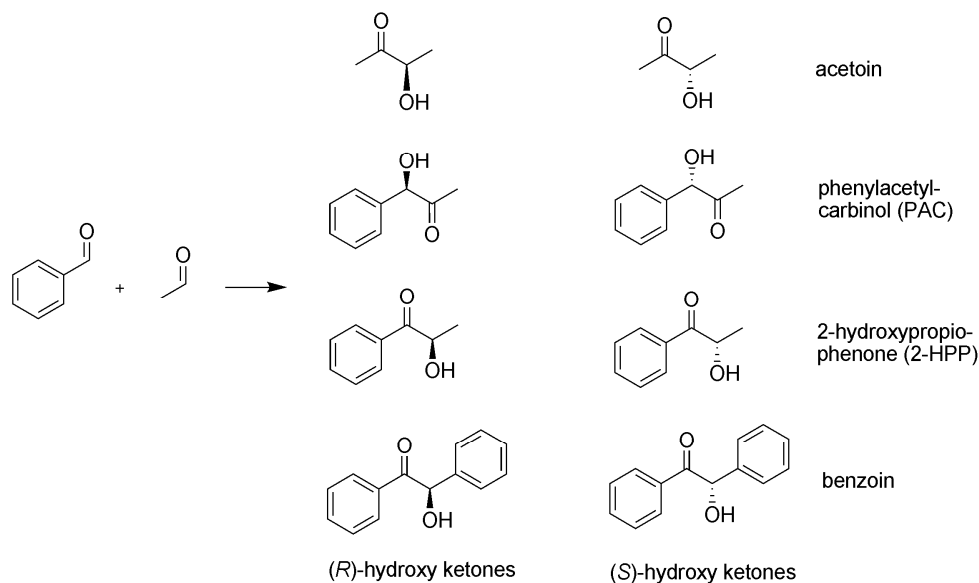
Fig. 1-9: Carboligation of benzaldehyde and acetaldehyde. Four different products can occur, each in the (*R*)- or (*S*)-conformation.
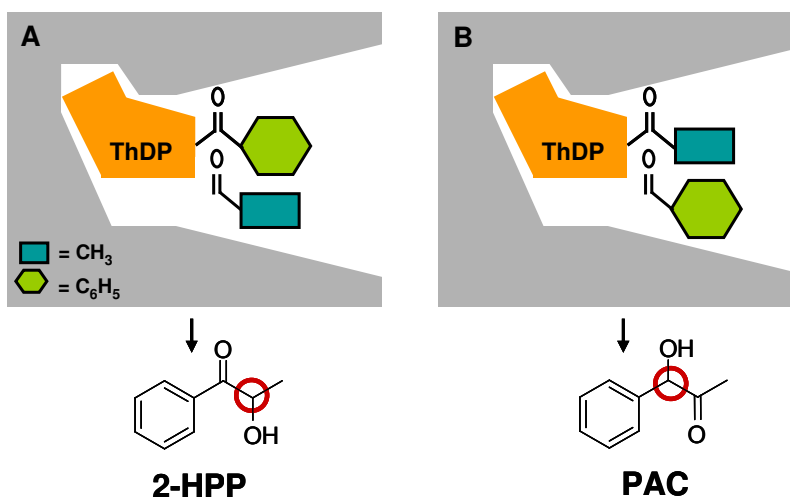
Fig. 1-10: The chemoselectivity of ThDP-dependent enzymes depends on the binding order of the substrates in the active site. If the benzaldehyde is bound to the C2-atom of ThDP it functions as the donor aldehyde and the aliphatic acetaldehyde as the acceptor aldehyde yielding 2-HPP (**A**). If acetaldehyde is the donor and benzaldehyde the acceptor PAC is formed (**B**).

Whether 2-HPP or PAC is obtained by the respective enzyme depends on the chemo-selectivity of the biocatalyst. Due to the variation of the shapes and amino acid residues forming the active sites of ThDP-dependent 2-keto acid decarboxylases the chemo-selectivity of these biocatalysts can differ significantly. As some examples in chapter 1.5 show, the chemoselectivity of the enzymes is often rather high. Factors influencing the chemoselectivity of ThDP-dependent enzymes will be analysed in the scope of this thesis.

For simplification within this thesis mixed products of differently substituted aromatic and aliphatic aldehydes are often named *PAC-analogues*, when the hydroxy group is situated next to the aromatic ring, or *HPP-analogues*, having the keto group positioned next to the ring (Fig. 1-10).

Stereoselectivity

In recent years the ability of ThDP-dependent enzymes to catalyse the stereoselective carboligation of aldehydes yielding diverse chiral 2-hydroxy ketones has been intensively studied with various wild type enzymes and variants thereof (Demir *et al.*, 2003; Demir *et al.*, 2002; Domínguez de María *et al.*, 2006; Dünkelmann *et al.*, 2004; Pohl *et al.*, 2002; Siegert *et al.*, 2005; Stillger *et al.*, 2006). From these studies it became apparent that the stereocontrol is only strict if the carboligation reaction encompasses at least one aromatic aldehyde, whereas with two aliphatic aldehydes only moderate enantiomeric excesses could be obtained (Domínguez de María *et al.*, 2007; Siegert *et al.*, 2005). It can be emphasised that in almost all mixed carboligations employing either solely aromatic or aromatic and aliphatic aldehydes exclusively the (*R*)-products are formed. Currently there is only one exception from this rule: BFD from *Pseudomonas putida* (*Pp*BFD) catalyses the formation of (*S*)-2-HPP (*ee* 92%) (Fig. 1-11) from benzaldehyde and acetaldehyde, but yields (*R*)-benzoin (*ee* 99%) in self-ligation reactions with benzaldehyde (Iding *et al.*, 2000). The structural basis for the differences in stereo control have recently been investigated by comparing the active sites of *Pp*BFD with benzaldehyde lyase from *Pseudomonas fluorescens* (*Pf*BAL), which is strictly (*R*)-specific (Knoll *et al.*, 2006). Modelling studies revealed a so-called *S*-pocket in *Pp*BFD, which could bind small acceptor aldehydes such as acetaldehyde, whereas no such pocket exists in *Pf*BAL (Fig. 1-11).
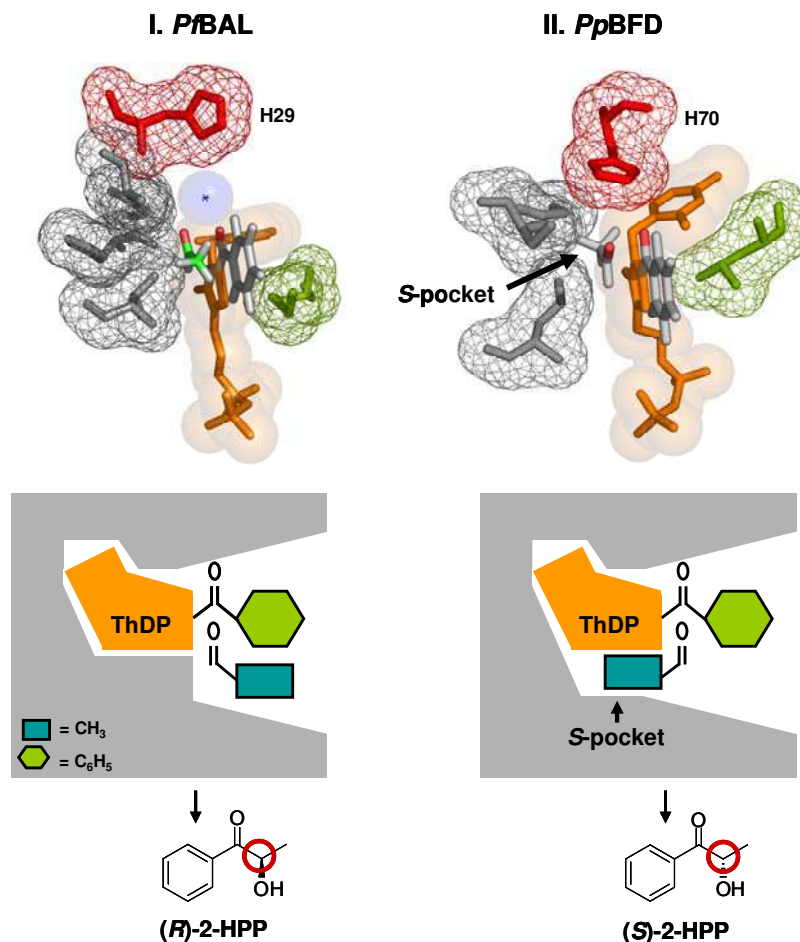
Fig. 1-11: Molecular models showing the assumption for the orientation of the acceptor aldehyde in *Pf*BAL (**I**) and *Pp*BFD (**II**). In both cases benzaldehyde is assumed to adapt a coplanar conformation relative to the thiazolium ring, due to mesomeric effects. The orientation of the oxygen of the donor aldehydes (benzaldehyde, grey) as well as the acceptor aldehydes (acetaldehyde, *Pf*BAL: green, *Pp*BFD: grey) is directed upwards towards H29 in *Pf*BAL (mediated by a water molecule) and H70 in *Pp*BFD, acting as proton relays. As there is no other possibility, in *Pf*BAL the acceptor aldehyde must be always arranged parallel to the donor aldehyde, as demonstrated for acetaldehyde (bright green) leading to the formation of (*R*)-HPP. In the case of *Pp*BFD the *S*-pocket allows the antiparallel orientation of the acceptor acetaldehyde relative to the donor benzaldehyde, resulting in the formation of (*S*)-HPP (according to Knoll *et al.*, 2006).

According to the modelling studies, the *S*-pocket in *Pp*BFD is just large enough to adopt the side chain of acetaldehyde, while larger aldehydes may not be bound to this pocket. Consequently, larger aldehydes are arranged parallel to the ThDP-bound donor aldehyde, resulting in the formation of (*R*)-2-hydroxy ketones. Thus, (*R*)-selectivity seems to be an intrinsic property of all 2-keto acid decarboxylases and *Pf*BAL. The confirmation of this hypothesis is part of this thesis.

1.5   Special enzymes

This chapter reviews the literature about the enzymes used in this work summarising the state of the art before this thesis was started and important publications which have appeared during the course of this thesis. The summary focuses on important aspects for

this work. Detailed information concerning the different enzymes is found in the publications, the general discussion and the cited references.

1.5.1 Benzoylformate decarboxylase

Benzoylformate decarboxylase (BFD, E.C. 4.1.1.7) activity in the bacterium *Pseudomonas putida* (ATCC 12633) was first described by Hegemann *et al.* in 1966. It is the third enzyme of the mandelic acid catabolism, where mandelic acid is converted to benzoic acid in order to finally be metabolised in the β-ketoadipate pathway (Stanier & Ornston, 1973) and the citric acid cycle. This enables the organism to grow on (*R*)-mandelic acid as a sole carbon source. The physiological function of BFD is the non-oxidative conversion of benzoylformate to benzaldehyde and CO_2. In 1990 Tsou *et al.* cloned the coding gene (*mdlC*) (Tsou *et al.*, 1990).

Just recently Henning (Henning *et al.*, 2006) isolated three further proteins with BFD activity, two of them originating from a chromosomal library of *Pseudomonas putida* ATCC 12633 and the third from an environmental-DNA library of a metagenome screening. Their physiological activity is still unknown. Several putative genes encoding potential BFDs can be found by sequence comparison with the BLAST tool (BLAST, 2007). For four of them, BFD from *Bradyrhizobium japonicum* (Wendorff, 2006), *Pseudomonas aeruginosa* (Barrowman *et al.*, 1986), *Pseudomonas stutzeri* (Saehuan *et al.*, 2007) and *Acinetobacter calcoaceticus* (Barrowman & Fewson, 1985) the enzymatic activity could be verified. However, the enzyme from *Pseudomonas putida* (*Pp*BFD) is by far the best characterised BFD. The substrate range for its physiological decarboxylase activity was thoroughly studied (Iding *et al.*, 2000; Siegert *et al.*, 2005). Kinetic parameters for the decarboxylation of benzoylformate showed a hyperbolic Michaelis-Menten plot (V_{max} of 320 U/mg, K_M value of 0.7 mM). Most interestingly the substrate range is highly specific for benzoylformate, while larger aromatic and small aliphatic 2-keto acids are only poorly accepted (Iding *et al.*, 2000).

In 1992 Wilcocks *et al.* first described the ability of *Pp*BFD to catalyse C-C-bond formations (Wilcocks & Ward, 1992; Wilcocks *et al.*, 1992). By incubating whole *Pseudomonas putida* cells with acetaldehyde and benzaldehyde (*S*)-2-HPP formation (*ee* of 91-92%) was gained. Detailed studies of the substrate range for the carboligation reaction using purified wild type enzyme followed (Dünnwald *et al.*, 2000; Iding *et al.*, 2000). *Pp*BFD is able to catalyse the ligation of a broad range of different aromatic, heteroaromatic, cyclic aliphatic as well as olefinic aldehydes yielding exclusively the HPP-analogue (Fig. 1-10). Hence the selectivity for aromatic aldehydes as donor aldehydes is very high (Dünkelmann *et al.*, 2002), which is also reflected in the relative activities of *Pp*BFD concerning the formation of (*S*)-2-HPP (7 U/mg; *ee* 92%), (*R*)-benzoin (0.25

U/mg; *ee* 99%) and acetoin (0.01 U/mg, *ee* 13%) (Iding *et al.*, 2000; Siegert *et al.*, 2005) (Fig. 1-9). Depending on the substitution pattern of the aromatic ring, diverse HPP-analogues are accessible in high yields and with good to high optical purity. Selectivity, activity and stability of *Pp*BFD have been optimised using reaction engineering. Best results have been obtained by adjusting very low benzaldehyde concentrations in a continuous reactor (Iding *et al.*, 2000). Varying techniques of directed evolution yielded variants with enhanced ligase activity for the formation of (*S*)-HPP, improved stability towards organic solvents and an enlarged donor substrate range (Lingen *et al.*, 2002; Lingen *et al.*, 2003). Since the crystal structure of *Pp*BFD could be solved with (*R*)-mandelate as an inhibitor (Polovnikova *et al.*, 2003) and without (Hasson *et al.*, 1995; Hasson *et al.*, 1998), rational protein design and mechanism based inhibitor studies (Bera *et al.*, 2007) are much more efficient nowadays.

1.5.2 Pyruvate decarboxylases

Pyruvate decarboxylases (PDCs, E.C. 4.1.1.1) catalyse the non-oxidative decarboxylation of pyruvate to acetaldehyde. They are the key enzymes in fermentative ethanol pathways (Fig. 1-12) and commonly found in

1. plants (especially in the seeds); e.g. *Oryza sativa* (rice) (Hossain *et al.*, 1996; Rivoal *et al.*, 1990)*, Zea mays* (maize) (Forlani, 1999) and *Pisum sativum* (pea) (Mücke *et al.*, 1995),
2. yeasts; e.g. *brewers yeast* (Holzer *et al.*, 1956), *Saccharomyces cerevisiae* (*Sc*PDC) (Schmitt & Zimmermann, 1982), *Saccharomyces kluyveri* (Moller *et al.*, 2004), *Kluyveromyces marxianus* (Holloway & Subden, 1993), *Zygosaccharomyces bisporus* (Neuser *et al.*, 2000),
3. fungi; e.g. *Neurospora crassa* (Alvarez *et al.*, 1993), *Aspergillus nidulans* (Lockington *et al.*, 1997), *Rhizopus oryzae* (Skory, 2003) and
4. bacteria; gram-negative bacteria: *Zymomonas mobilis* (*Zm*PDC) (Dawes *et al.*, 1966), *Zymobacter palmae* (*Zp*PDC) (Okamoto *et al.*, 1993) and *Acetobacter pasteurianus* (*Ap*PDC) (Raj *et al.*, 2001); gram-positive bacterium: *Sarcina ventriculi* (*Sv*PDC) (Lowe & Zeikus, 1992).

By contrast no PDCs have been found in mammals.
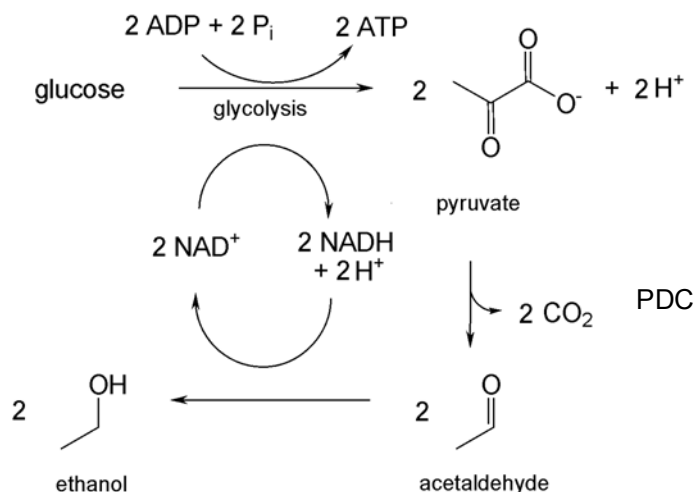
Fig. 1-12: Alcoholic fermentation (according to Campbell, 1997).

PDC from Saccharomyces cerevisiae (*Sc*PDC)

*Sc*PDC has been intensively studied concerning the mechanism of the decarboxylase reaction. Especially the substrate activation observed in brewers yeast and *Sc*PDC have been thoroughly studied (Jordan, 2003; König, 1998). The sigmoidal v/[S]-plot of yeast PDCs was already described 30 years ago (Hübner *et al.*, 1978; König, 1998). Similar results have been obtained with plant PDCs (Dietrich & König, 1997) and with the prokaryotic PDC from *Sarcina ventriculi* (*Sv*PDC) (Talarico *et al.*, 2001). Although the binding site for the activator is still a matter of discussion, it is evident that its binding leads to structural changes in the tetrameric arrangement resulting in a more compact active structure (chapter 1.4.3) (Arjunan *et al.*, 1996; Dyda *et al.*, 1993; König, 1998; Lu *et al.*, 1997).

The carboligase potential of PDC from yeast, in particular brewers yeast, is already known for more than 100 years (Neuberg & Karczag, 1911) and used for the production of (*R*)-PAC by whole cell biotransformation, as described above (Fig. 1-3). The investigation and optimisation of the PAC-synthesis is still a matter of actual research (Rosche *et al.*, 2005; Rosche *et al.*, 2003). Both, brewers yeast as well as *Sc*PDC show high enantioselectivity towards the (*R*)-product and high chemoselectivity. In contrast to *Pp*BFD, all PDCs prefer small aliphatic donor aldehydes and aliphatic or aromatic acceptor aldehydes. Thus, acetoin and PAC (Fig. 1-9) are typical ligation products of PDCs (Pohl *et al.*, 2002).

PDC from Zymomonas mobilis (*Zm*PDC)

Beside yeast PDCs the enzyme from *Zymomonas mobilis* has been intensively investigated. *Zymomonas mobilis* is a model microorganism for the Entner & Doudoroff pathway in which glucose is converted anaerobically to ethanol and CO_2 (Entner & Doudoroff, 1952). Dawes investigated this pathway in detail and detected the pyruvate decarboxylase taking part in this metabolism (Dawes *et al.*, 1966). After having found that *Zymomonas mobilis* is able to ferment glucose 6-7 times faster and with 5% higher yield compared to

Saccharomyces cerevisiae (Lee *et al.*, 1979), *Zm*PDC was isolated and partly characterised (Bringer-Meyer *et al.*, 1986). A detailed analysis of purified *Zm*PDC concerning kinetics, pH-stability, salt effects and thermostability followed (Pohl *et al.*, 1995) showing a Michaelis-Menten like kinetic for the decarboxylation of pyruvate (V_{max} of 120-150 U/mg, K_M of 1.1-1.3 mM). In contrast to *Sc*PDC no substrate activation was observed. First mutagenesis studies were performed based on a structural model derived from the 3D-structure of *Sc*PDC (Arjunan *et al.*, 1996) to investigate the reaction mechanism and to identify *hot spots* in the active site (Bruhn *et al.*, 1995; Candy & Duggleby, 1994; Candy & Duggleby, 1998; Candy *et al.*, 1996; Mesch, 1997; Pohl, 1997; Pohl *et al.*, 1998; Tittmann *et al.*, 1998). The crystal structure of *Zm*PDC was finally published in 1998 (Dobritzsch *et al.*, 1998), demonstrating a compact tetrameric structure, which resembles the quaternary structure of *Sc*PDC upon substrate activation (König, 1998).

The carboligase activity of *Zm*PDC was first described by Bringer-Meyer *et al.* (1988), investigating the formation of PAC and acetoin with partially purified *Zm*PDC (Bringer-Meyer & Sahm, 1988). A more detailed analysis demonstrated that although *Zm*PDC is strictly (*R*)-specific concerning the formation of PAC and PAC-analogues, it catalyses the formation of predominantly (*S*)-acetoin (*ee* 50-60%) (Bornemann *et al.*, 1993; Bornemann *et al.*, 1996; Bruhn *et al.*, 1995). The latter is in contrast to *Sc*PDC forming (*R*)-acetoin in slight excess (*ee* 20%) (Bornemann *et al.*, 1993).

Site-directed mutagenesis was used to improve carboligase activity towards PAC and to identify residues which are important for the enantioselectivity of the carboligation (Bruhn *et al.*, 1995; Pohl *et al.*, 1998; Siegert *et al.*, 2005). From these studies isoleucine 472 was determined to be a hot spot influencing enantio- and chemoselectivity (Pohl *et al.*, 1998; Siegert *et al.*, 2005). Tryptophane 392 was identified as a key residue limiting the access to the active centre for big aromatic substrates (Bruhn *et al.*, 1995; Mesch, 1997; Pohl, 1997). The highly potent variant *Zm*PDCW392M was tested in a continuous enzyme-membrane reactor giving space-time-yields of 81 g*L^{-1}*d^{-1}(Goetz *et al.*, 2001; Iwan *et al.*, 2001).

PDC from Acetobacter pasteurianus (*Ap*PDC)

Acetobacter pasteurianus (ATCC 12874) is an obligate oxidative bacterium. In physiological studies *Ap*PDC was found to be a central enzyme for the oxidative metabolism, which is in contrast to the fermentative ethanol pathways in plants, yeasts and other bacterial PDCs described above (Fig. 1-12) (Raj *et al.*, 2001). *Acetobacter pasteurianus*, a common spoilage organism of fermented juices, derives energy and carbon from the oxidation of D- and L-lactic acid. During growth *Ap*PDC cleaves the central intermediate pyruvate to acetaldehyde and CO_2.

Recombinant, purified *Ap*PDC was already described with respect to its pH- and temperature optimum for the decarboxylase reaction (Raj *et al.*, 2002). With pyruvate a kinetic according to Michaelis-Menten was detected similar to the one of *Zm*PDC (V_{max} of 97 U/mg at pH 5.0, K_M of 0.39 mM). There were no data available concerning carboligase activity of *Ap*PDC.

PDC from Zymobacter palmae (**Zp**PDC)

The ethanologenic bacterium *Zymobacter palmae* (ATCC 51623) was originally isolated from palm sap (Okamoto *et al.*, 1993) producing ethanol from hexose sugars and saccharides like most other PDCs (Horn *et al.*, 2000). The gene was cloned and *Zp*PDC was characterised with respect to its future application as a biocatalyst for fuel ethanol production from renewable biomass (Raj *et al.*, 2002). As has been observed for *Zm*PDC and *Ap*PDC, *Zp*PDC shows a Michaelis-Menten like kinetic for pyruvate (V_{max} of 140 U/mg at pH 7.0, K_M of 0.71 mM). Again no data concerning carboligase activity were available.

1.5.3 Branched-chain keto acid decarboxylases

Branched chain keto acid decarboxylases (E.C. 4.1.1.72) have been described from *Bacillus subtilis* (Oku & Kaneda, 1988), *Lactococcus lactis* IFPL730 (Amárita *et al.*, 2001) and from *Lactococcus lactis* subspec. *cremoris* (*Ll*KdcA) (Smit *et al.*, 2005a). Smit suggested the *Ll*KdcA to be the key enzyme for the metabolism from leucine to 3-methylbutanal (Fig. 1-13) (Smit *et al.*, 2005a).
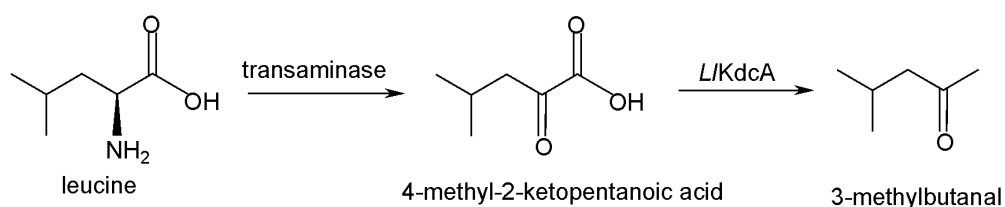
Fig. 1-13: Leucine metabolism in *Lactococcus lactis* (according to Smit *et al.*, 2005a).

Degradation of proteins to amino acids followed by transamination and decarboxylation is crucial for flavour formation during cheese ripening (McSweeney & Sousa, 2000). The *kdcA*-gene was isolated from *L. lactis* B1157, one of the few KdcA producing *Lactococcus* strains (Smit *et al.*, 2005a; Smit *et al.*, 2005b). The recombinant, purified enzyme was characterised concerning its substrate range for decarboxylase activity revealing a very broad spectrum with 3-methyl-2-oxobutanoic acid being converted with highest specific activity. Since valine is the corresponding amino acid to this 2-keto acid, the conversion of valine seems to be even more relevant than the initially screened leucine metabolism (Smit *et al.*, 2005a). Further substrates for *Ll*KdcA are other branched chain as well as aromatic and linear aliphatic keto acids, making the enzyme an interesting target for the

investigation of a putative carboligase activity. Just recently, a homology model was developed in order to understand the structural basis of the *Ll*KdcA substrate recognition and plan site directed mutagenesis (Yep *et al.*, 2006a).

1.5.4   Phenylpyruvate decarboxylases / Indole-3-pyruvate decarboxylases

Indole-3-pyruvate decarboxylases (In3PDC, E.C. 4.1.1.74) are involved in the biosynthesis of auxins, growth-promoting phytohormones in plants. Several biosynthetic pathways for the production of indole-3-acetic acid are in discussion (Costacurta *et al.*, 1994). The best known is the indole-3-pyruvate pathway, in which tryptophan seems to be first transaminated to indole-3-pyruvate, which is subsequently decarboxylated to indole-3-acetaldehyde by In3PDC and finally oxidised to the auxin indole-3-acetic acid (Costacurta & Vanderleyden, 1995; Koga, 1995; Patten & Glick, 1996). In 1992 Koga and coworkers described the cloning and characterisation of In3PDC from *Enterobacter cloacae* (Koga *et al.*, 1992). Highest decarboxylase activity was detected with indole-3-pyruvate while the activity with phenylpyruvate was negligible. Stopped-flow kinetics data showed no indication for substrate activation (Schütz *et al.*, 2003). The structure of this enzyme was later solved, showing the typical homo-tetrameric fold of ThDP-dependent enzymes (pdb-code: 1ovm, Schütz *et al.*, 2003). A similar gene was found in *Azospirillum brasilense* (*Ab*In3PDC) (Costacurta *et al.*, 1994), a gram-negative nitrogen-fixing rhizobacterium living in association with roots of grasses and cereals due to its growth-promoting effect on root hairs (Okon & Labandera-Gonzalez, 1994). It was recently demonstrated that *Ab*In3PDC is also involved in the biosynthesis from phenylalanine to phenylacetic acid, an auxin with antimicrobial activity (Somers *et al.*, 2005), demonstrating that the promiscuity of this enzyme may have an impact in two different pathways.

Phenylpyruvate decarboxylases (PhPDCs, E.C. 4.1.1.43) are involved in the Ehrlich pathway using amino acids as sole amino sources. The catabolism of phenylalanine comprises similar steps as the indole-3-pyruvate-pathway: an initial transamination to phenylpyruvic acid is followed by a PhPDC-catalysed decarboxylation to phenylacet-aldehyde. But the final step includes a reduction to phenylethanol, latter possessing a rose-like aroma interesting for the beverage and food industry (Vuralhan *et al.*, 2003). PhPDCs have been described in bacteria species of *Acromobacter*, *Acinetobacter* and *Thauera* (Ward & Singh, 2000) and in the yeast *Saccharomyces cerevisiae* (Vuralhan *et al.*, 2003). Their expression is inducible by growing the organism on phenylalanine, tryptophane or mandelate. Decarboxylase activity could be determined with phenylpyruvate, indole-3-pyruvate and further 2-keto acids with more than six carbon atoms in a straight chain (Ward & Singh, 2000). Asymmetric carboligase activity of PhPDCs has already been investigated by Guo *et al.* using phenylacetaldehyde (derived by *in situ* decarboxylation of phenylpyruvate) and acetaldehyde as substrates yielding the PAC-analogue 3-hydroxy-1-

phenyl-2-butanone by whole-cell catalysis of *Acromobacter eurydice*, *Pseudomonas aromatica* and *Pseudomonas putida* (Guo *et al.*, 1999). With purified *Acromobacter eurydice* PhPDC the ligation of phenylacetaldehyde (phenylpyruvate) with various aldehydes yielded also (*R*)-PAC derivatives with *ees* of 87-98% (Guo *et al.*, 2001).

1.5.5   Benzaldehyde lyase

Up to now the only biochemical characterised benzaldehyde lyase (BAL, EC 4.1.2.38) was found in *Pseudomonas fluorescens* Biovar I (*Pf*BAL) (González *et al.*, 1986; González & Vicuña, 1989). The gram-negative bacterium was isolated from wood splits of a cellulose factory showing the ability to grow on lignin-derived compounds like anisoin (4,4'-dimethoxybenzoin) and benzoin as the sole carbon and energy sources. Due to the lyase activity of *Pf*BAL the organism can cleave these 2-hydroxy ketones yielding aldehydes which are probably metabolised in the β-ketoadipate pathway (Stanier & Ornston, 1973). Sequence comparison using the BLAST-tool (BLAST, 2007) revealed two putative *bal*-genes in the purple non-sulfur bacterium *Rhodopseudomonas palustris* (ATCC BAA-98).

Detailed analysis of the stability and lyase activity of recombinant, purified *Pf*BAL was published by Janzen *et al.* revealing maximal specific activity for the cleavage of (*rac*)-benzoin of 74 U/mg (K_M of 0.05 mM) (Janzen, 2002; Janzen *et al.*, 2006). Due to its high (*R*)-selectivity *Pf*BAL can be used for kinetic resolution yielding (*S*)-benzoin which is otherwise difficult to access (Demir *et al.*, 2001). Furthermore the enzyme is able to cleave (*R*)-2-HPP (V_{max} of 3.6 mM; K_M of 0.3) (Janzen *et al.*, 2006). In contrast to the other ThDP-dependent enzymes described above, *Pf*BAL shows very little decarboxylase activity (0.02 U/mg for benzoylformate).

Demir *et al.* could testify that the lyase reaction is reversible, enabling *Pf*BAL also to form C-C bonds. Especially the self-ligation of benzaldehyde towards benzoin is catalysed with high reactivity (336 U/mg) and very high enantioselectivity (*ee* > 99% (*R*)) making this enzyme very interesting for industrial processes. For the benzoin formation the substrate range of aldehydes substituted in *ortho*-, *meta*- and *para*-position is very broad (Demir *et al.*, 2002). Mixed carboligations of several aromatic and aliphatic aldehydes result in the formation of HPP-analogues beside the benzoin derivative (Fig. 1-9) (Demir *et al.*, 2001; Demir *et al.*, 2002). Also phenylacetaldehyde is accepted as a donor aldehyde (Sanchez-Gonzalez & Rosazza, 2003). Beside acetaldehyde mono-, dimethoxyacetaldehyde and formaldehyde have been successfully applied as aliphatic acceptor aldehydes (Demir *et al.*, 2004; Demir *et al.*, 2003; Demir *et al.*, 2002; Dünkelmann *et al.*, 2004).

The transformation of aromatic aldehydes requires the use of cosolvents to improve the solubility of the aromatic substrates and products in certain cases. Dimethylsulfoxide

(DMSO) was shown to be extremely useful, not only for the solubility of the aromatic compounds but also for the stability of *Pf*BAL during biotransformations (Demir *et al.*, 2001). Recently, the carboligation of different 2-hydroxy ketones was studied in continuously operated membrane reactors using aqueous buffer systems with 30% (v/v) DMSO (Hildebrand *et al.*, 2007; Kurlemann & Liese, 2004; Stillger *et al.*, 2006). Further studies concerned the application of immobilised *Pf*BAL in cryogel beads which can be used in organic solvents. This two-phase system was successfully applied in batch as well as continuous reactions for the synthesis of (*R*)-3,3'-furoin (Ansorge-Schumacher *et al.*, 2006; Hischer *et al.*, 2005).

The crystal structure of this highly versatile biocatalyst was solved in 2005 (Mosbacher *et al.*, 2005). It opened the way to rationalise effects observed by mutations in the active site which have been produced based on structural models (Janzen, 2002; Janzen *et al.*, 2006; Kneen *et al.*, 2005) and a structure guided explanation for the differences observed between *Pf*BAL and *Pp*BFD (Knoll *et al.*, 2006).

1.6   Aim of the thesis

The general aim of this thesis is the design of a toolbox of ThDP-dependent enzymes with carboligase activity as diverse as possible in order to gain a highly substituted 2-hydroxy ketone platform. Main criterions for the selection of enzymes for the toolbox are high or complementary chemo- and enantioselectivities and varying substrate ranges.

This goal should be achieved by the following steps:

1. Development of a **high-throughput assay** for a quick detection of the potential of enzymes with carboligase activity. This test should be fast, easy to handle and reliable. Preferably it should be applicable for product detection after biotransformations with both, purified enzymes and whole cells. Most importantly, it should allow the sensitive detection of as many 2-hydroxy ketones as possible to be used as a fingerprinting tool for the rapid identification of the substrate range of new enzymes and enzyme variants.

2. Access to new enzymes with carboligase activity in order to **enlarge the enzyme toolbox**. Among the various possibilities to obtain new enzymes the focus in this work is laid on:
   - database mining,
   - screening of already known enzymes with decarboxylase activity for their C-C coupling potential,
   - alteration of the substrate range and selectivity of known enzymes by rational protein design.

3. **Enlargement of the hydroxy ketone platform**. Beside the implementation of new enzymes yielding different 2-hydroxy ketones the enzymes in the toolbox can also be screened for the acceptance of new substrates. First indications concerning activity and chemoselectivity should be deduced from the high-throughput assay. Additionally the preliminary results require verification by instrumental analysis using appropriate standards. Beside mutagenesis techniques variations of the reaction conditions can be used to vary chemo- and enantioselectivities of the biocatalysts yielding new products. The focus in this work is laid on so far hardly accessible types of 2-hydroxy ketones:

- branched-chain aliphatic hydroxy ketones and longer chained non-branched aliphatic aldehydes,
- substituted (*R*)-phenylacetylcarbinol-analogues,
- 2-hydroxy ketones from instable aldehydes e.g. indole-3-acetaldehyde and phenylacetaldehyde,
- and most importantly: access to (*S*)-2-hydroxy ketones.

As some of the techniques used to meet these challenges rely on a rational structure guided approach, structure-function relationships determining substrate range, chemoselectivity and enantioselectivity should be examined. Such studies are only possible if sufficient structural information is available and require the structural analysis of interesting enzymes and variants which is an additional target of this thesis.

If structural and biochemical investigations can be combined to generate a new mechanistic understanding this will enable a much faster, scientifically profound and more effective prediction and design of new catalysts with desired properties.

2 PUBLICATIONS

Publication I

A colorimetric assay for the detection of 2-hydroxy ketones.

Gocke, D., Wendorff, M. & Pohl, M.

submitted to Analytical Biochemistry

A colorimetric assay for the detection of 2-hydroxy ketones

Dörte Gocke, Marion Wendorff, Martina Pohl*

Institute of Molecular Enzyme Technology, Heinrich-Heine University Düsseldorf, Research Centre Jülich, D-52426 Jülich, Germany
* Corresponding author: ma.pohl@fz-juelich.de, Tel.: 0049-2461-613704,
Fax: 0049-2461-612490

Short title: An assay for 2-hydroxy ketones

Abstract
A simple and reliable tetrazolium chloride based colorimetric assay is described for the detection of a broad range of 2-hydroxy ketones, which is useful for high throughput analysis as well as for fingerprinting studies of the substrate range of enzymes. The assay has proved to be extremely sensitive for the detection of aromatic and aliphatic 2-hydroxy ketones with detection limits of 0.5-2 mM, whereas the sensitivity drops significantly with branched-chain aliphatic compounds. It is readily applicable for purified lyophilized enzymes as well as clarified crude cell extracts. Only few disturbing substances have been identified so far. Among these are glycerol, 2-methoxyacetaldehyde, and indole-3-pyruvate, which cause a more or less strong coloration of the assay.
The reliability of the assay is demonstrated in combination with data obtained from instrumental analysis.

Keywords: acyloin condensation, benzaldehyde lyase, benzoylformate decarboxylase, carboligation, pyruvate decarboxylase, thiamin diphosphate

Abbreviations: *Ap*PDC, *Acetobacter pasteurianus* pyruvate decarboxylase; BAL, *Pseudomonas fluorescens* benzaldehyde lyase; BFD, *Pseudomonas putida* benzoylformate decarboxylase; 2-HPP, 2-hydroxypropiophenone; *Ll*KdcA, *Lactococcus lactis* branched-chain keto acid decarboxylase; PAC, phenylacetylcarbinol; *Sc*PDC, *Saccharomyces cerevisiae* pyruvate decarboxylase; *Sc*PhPDC, *Saccharomyces cerevisiae* phenylpyruvate decarboxylase; TTC, 2,3,4-triphenyltetrazolium chloride; *Zp*PDC, *Zymobacter palmae* pyruvate decarboxylase; *Zm*PDC, *Zymomonas mobilis* pyruvate decarboxylase.

Introductory statements
Chiral 2-hydroxy ketones are important building blocks for chemical syntheses. They can be prepared with high chemo- and enantioselectivity using thiamin diphosphate (ThDP)-dependent enzymes [1, 2]. Meanwhile a broad range of various 2-hydroxy ketones is accessible by an enzyme toolbox consisting of 11 ThDP-dependent enzymes, among them are pyruvate decarboxylases from different organisms (*Zymomonas mobilis, Zm*PDC [3-5]; *Saccharomyces cerevisiae, Sc*PDC [1, 5]; *Acetobacter pasteurianus, Ap*PDC [6]), a branched-chain 2-ketoacid decarboxylase from *Lactococcus lactis* (*Ll*KdcA) [7, 8], a benzoylformate decarboxylase from *Pseudomonas putida* (BFD) [3, 9-11], a benzaldehyde

lyase from *Pseudomonas fluorescens* (BAL) [12-15], and more than 50 variants of these enzymes.

For rapid detection of carboligase activity, e.g. in the case of high throughput screening of enzyme libraries or for investigation of the substrate range we have developed an easy and reliable colorimetric assay based on the formation of a coloured formazane dye (Fig. 1).

This assay has originally been established to detect the formation of phenylacetylcarbinol (PAC) catalyzed by pyruvate decarboxylase [16] and was also shown to be useful for the detection of isomeric 2-hydroxypropiophenone (2-HPP) accessible by e.g. BFD catalysis [17] and for the detection of α,α'-dihydroxy ketones [18].

Here we describe the extension of its application to a broad range of aliphatic and aromatic acyloins including a thorough study of limitations and disturbing effects. Further, the reliability of the assay is proved by instrumental analysis.

Materials and Methods

All chemicals were purchased from Sigma-Aldrich or Fluka (St. Louis, MO, USA) and were of analytical grade. 2-Hydroxy ketones used for calibration and to determination of detection limits were either bought (Tab. 1, **1**, **5** (Fluka)) or synthesised. Branched-chain 2-hydroxy ketones (**3**, **4**) [12] were synthesised by non-enzymatic acyloins condensation as described by Fleming *et al.* [19]. **2**, **5**, **6**, **10-12** were synthesised enzymatically with benzaldehyde lyase from *Pseudomonas fluorescens* [9, 12, 14, 20] and **7** and **9** by enzymatic carboligation with pyruvate decarboxylase from *Acetobacter pasteurianus* (*Ap*PDC) [6]. After enzymatic synthesis products were purified chromatographically.

Assay conditions for crude extracts

To investigate the formation of 2-hydroxy ketones in crude cell extracts (e.g. during high throughput screening of enzyme libraries) recombinant *E. coli* clones containing a ThDP-dependent enzyme are inoculated in deep-well plates (polypropylen masterblock, 2 mL, Greiner bio-one International AG, Frickenhausen, D) containing LB-Medium (100 µL per well), including appropriate antibiotics. The wells are covered with an AirPore tape sheet (Qiagen, Hilden, D) and inoculated on an orbital shaker over night at 37°C, 600 rpm. After about 15 h the pre-culture in the wells is diluted with 600 µL LB-medium including antibiotics. After having shaken the plates for further 2.5 h at 30°C, 600 rpm, enzyme production is induced by addition of 10 µL 70 mM IPTG (16.67 mg/mL in 70% (v/v) ethanol, final concentration 1 mg/mL).

Subsequently, the plates are shaken over night at 37°C, 600 rpm. Prior to cell harvest by centrifugation (10°C, 4000 rpm, 20-30 min) replica plates are generated using a colony copier (96 pins). After centrifugation 500 µL of the supernatant of each well are removed carefully without touching the cell pellet. Then 250 µL potassium phosphate buffer (50 mM, pH 7, 2.5 mM MgSO$_4$, 0,1 mM ThDP) including lysozyme (1.4 mg/mL) are added. For a fast resuspension of the sedimented cells virtuous vortexing of the plates is advisable. The microtiter plates are shaken at 600 rpm, 30°C for at least 30 min. Cell debris are sedimented by centrifugation (10°C, 4000 rpm, 20-30 min) and 50-100 µL of the clear crude extract are transferred to a transparent 96 well microtiter plate with a flat bottom (Rotilabo-Microtest plates, Roth, D).

The activity assay is started with a biotransformation step employing the addition of equal amounts of crude extract and substrate solution (50-100 µL, depending on the enzyme activity), containing 36 mM substrate (aldehydes or 2-keto acids) in potassium phosphate buffer (50 mM, pH 7, 2.5 mM MgSO$_4$, 0.1 mM ThDP) and 40% (v/v) DMSO to ensure solubility of substrates and products. The final concentrations in the assay are 18 mM substrate, 0.35 mg/mL enzyme, and 20% (v/v) DMSO. The pH of the buffer may be adjusted

according to the requirements of the respective enzyme. A negative control (equal amounts of substrate solution and buffer without added enzyme or crude extract) and, if possible, a positive control should be added. It is advisable to separate the positive control by empty wells from the investigated samples to avoid false positive results. To minimize evaporation, the plates are covered with a self-adhesive plastic foil (SealPlate™, Roth, D) and an additional microtiter plate lid and are incubated over night (18 h) at 30°C slightly shaken on an orbital shaker (100 rpm). Due to the strong smell and the toxicity of certain aldehydes, all steps should be carried out under a hood.

For detection of the formed 2-hydroxy ketones 40 µL of an alkaline TTC-solution, containing 1 volume 2,3,5,-tetrazoliumchloride (0.4% (w/v) in 75% (v/v) ethanol) and 3 volumes 1 M NaOH, is added carefully without touching the solution in the well with the tips. This avoids cross contamination and allows the use of one set of tips on a multichannel pipette. Depending on the 2-hydroxy ketone and its concentration the appearance of a red coloration is immediately visible. The detection can be performed photometrically, following the increase of the red coloration over 2 minutes at 510 nm or as an endpoint determination. The continuous observation of the colour development is especially advantageous, if different strongly positive clones are compared concerning their relative activity, as differences are no longer detectable after 2 min.

However, for fingerprinting of the substrate range or a more general yes/no response a detection of the final colour intensity after 30 sec and 2 min is sufficient. For documentation microtiter plates are scanned after 2 min.

Assay conditions for purified enzymes
The enzyme solution (50 µL, 0.3-0.6 mg/mL) is incubated with the substrate solution (50 µL, 36 mM of each substrate in 50 mM potassium phosphate, pH 7, 2.5 mM $MgSO_4$, 0.1 mM ThDP, 40% (v/v) DMSO) as described above. Detection of the 2-hydroxy ketones formed is performed as described for crude extracts. Glycerol disturbs the detection.

Instrumental analysis was performed using HPLC, GCMS, NMR and CD as described elsewhere [12, 20, 21].

Results and Discussion
1. Detection limits of 2-hydroxy ketones
As demonstrated in Tab. 1 the assay allows the detection of a broad range of diverse 2-hydroxy ketones in a concentration range which is commonly achieved during screenings employing crude cell extracts. Although there is a clearly lower detection limit for aromatic and mixed aliphatic/aromatic 2-hydroxy ketones of about 0.5 mM, aliphatic acyloins such as acyloin (**1**), propioin (**2**) and 5-hydroxy-2,7-dimethyl-octan-4-one (**4**) can be also detected. However, the assay is not sensitive for the detection of 4-hydroxy-2,5-dimethyl-hexan-3-one (**3**), the ligation product of two isobutyraldehyde molecules (Tab. 1).

2. Influence of the enzyme preparation
The TTC-assay is very efficient with purified enzymes with no background activity observed. However, the enzyme preparation should not contain glycerol as this causes a strong red coloration. The assay is further useful for the detection of enzyme activity in crude cell extracts, but cell debris should be removed before starting the assay as false positive results might be caused depending on the aldehyde applied.

3. Substrates for enzymatic 2-hydroxy ketone syntheses

2-Hydroxy ketones can be produced enzymatically starting from aldehydes or 2-keto acids which are decarboxylated by 2-keto acid decarboxylases yielding aldehydes. While acetaldehyde, propanal, isovaleraldehyde, isobutyraldehyde, dimethoxyacetaldehyde, benzaldehyde, 3,5-dichlorobenzaldehyde, 3,5-dimethoxybenzaldehyde, glyoxylate and pyruvate cause no background colouration of the assay, care has to be taken with 3-methoxy acetaldehyde (>1 mM), indole-3-pyruvate (>1 mM), and phenylpyruvate (>10 mM). Substrates hardly soluble in aqueous solution, e.g. phenylacetaldehyde, 3,5-dichloro-acetaldehyde, and indole-3-pyruvate should be dissolved in 100% DMSO followed by addition of the buffer or a buffered solution of the second substrate in the case of a two substrate reaction.

4. Adjustment of the assay procedure

In contrast to a previous version of this assay described by Breuer *et al.* [16] – addition of the tetrazolium salt first followed by addition of NaOH - we shortened the procedure by simultaneous addition of a mixture of both solutions. Besides saving one pipetting step this makes the assay more reliable. Moreover, 2-hydroxy ketones derived from dimethoxy-acetaldehyde (**12**, Tab.1) are destroyed by incubation in alkaline solution in only a few minutes. Thus, for these products the reliability of the assay is significantly increased by a simultaneous addition of the tetrazolium salt and NaOH.

5. Reliability studies

The assay is calibrated by an appropriate dilution of the expected 2-hydroxy ketones. Further, new substrates and products should be checked with respect to their detection limits and possible background reactions. Taking this into account the assay is highly reliable and the results obtained show a good correlation with instrumental analysis as demonstrated for the following examples.

5.1 Carboligation of benzaldehyde and/or acetaldehyde

As an example, the substrate range and relative activities were estimated for the carboligation of acetaldehyde and benzaldehyde for wild type benzoylformate decarboxylase (BFDwt), two BFD variants and benzaldehyde lyase (BALwt). As demonstrated in Fig. 2 A, there is a very week acetoin (**1**) forming activity observed with BALwt, whereas acetoin formation is only hardly detectable with the BFD variants tested. These results agree well with data obtained by GC analysis as demonstrated in Tab. 2. Concerning the formation of benzoin from benzaldehyde BALwt shows a strong benzoin forming activity as is visible from the intensively red coloured assay. With respect to the formation of benzoin there is a clear difference visible for the BFD variants. While BFDwt and the variant BFDL461V show a similar benzoin forming activity, an only very week activity was observed with the variant BFDL461A [21]. Again these results correlate well with quantitative results obtained by HPLC (Tab. 2). When an equimolar mixture of both aldehydes is applied, all tested enzymes show the formation of comparable amounts of 2-hydroxy ketones, as can be deduced from the similar colour intensity observed in Fig. 2 A. However, it is not possible to decide whether predominantly one product or a mixture of different 2-hydroxy ketones is formed. In the case of BFDwt and BFDL461V a mixture of benzoin (**5**) and the mixed products 2-hydroxy propiophenone (2-HPP, **6**) as well as phenylacetylcarbinol (PAC, **7**) is possible, whereas the variant BFDL461A has an almost negligible benzoin forming activity and the observed product is thus most probably a mixed product. Detailed analysis of the product mixture by NMR showed that the formation of benzoin in the presence of acetaldehyde is mainly suppressed. Only in the case of BFDwt small amounts of benzoin (4%) could be detected in the reaction mixture (Tab. 2). By contrast, the benzoin forming activity of BALwt could not be

suppressed by addition of acetaldehyde. In this case the mixed carboligation results in a mixture of benzoin (**5**) and 2-HPP (**6**) [13].

5.2 Carboligation of benzaldehyde and propanal
In the next example the carboligation of benzaldehyde and propanal was investigated with the same set of enzymes (Fig. 2 B). BALwt is known to catalyze the synthesis of propioin [22], benzoin [14, 23], and the cross ligation product 2-hydroxy-1-phenyl-butan-1-one (**8**) (unpublished) (Tab. 1, 2). The same reactions were also tested with the BFDwt and variants, which have been designed to catalyze the formation of (*S*)-2-hydroxy-1-phenyl-butan-1-one (**8**) [6, 21]. With the variant BFDL461A the formation of the mixed product **8** is immediately visible, as no benzoin may occur as a by-product (Fig. 2 B, Tab. 2). Analysis of the product range by NMR demonstrates a mixture of almost all possible products in the case of BFDwt and BFDL461V, while BFDL461A catalyzes mainly the formation of the mixed product **8** beside small amount of benzoin (Tab. 2) [21].

6. Fingerprinting the substrate range
Due to the broad range of 2-hydroxy ketones which can be detected (Tab. 1) the TTC-assay is very useful to rapidly screen enzymes for the formation of new products. As demonstrated in Fig. 2 C, nine different 2-ketoacid decarboxylases were tested with respect to the carboligation of 3,5-dichlorobenzaldehyde with dimethoxyacetaldehyde: among these are seven pyruvate decarboxylases *Ap*PDC [6], *Sc*PDC, *Zp*PDC [6], *Zm*PDC and variants of the latter [3-5, 24], *Ll*KdcA [7, 8], and a phenylpyruvate decarboxylase from *Saccharomyces cerevisiae* (*Sc*PhPDC) [25].
Signals were obtained only with BALwt, demonstrating that beside the self-ligation products of both aldehydes also a mixed 2-hydroxy ketone is formed. Detailed analysis using HPLC and circular dichroism demonstrated that (*R*)-1-(3,5-dichoro-phenyl)-2-hydroxy-3,3-dimethoxy-propane-1-one (**13**, Tab. 1) is almost exclusively formed (Michael Müller, personal communication).

Conclusions
The TTC-assay allows the fast screening of enzyme libraries for diverse 2-hydroxy ketones. It is useful to obtain rapidly information about the substrate range and allows the comparison of relative activities. Due to its high sensitivity and the simplicity of the procedure the assay can be applied in high throughput screenings.

References
[1] M. Pohl, G. A. Sprenger, and M. Müller, A new perspective on thiamine catalysis, Curr. Opin. Biotechnol. 15 (2004) 335-342.
[2] M. Pohl, B. Lingen, and M. Müller, Thiamin-diphosphate dependent enzymes: new aspects of asymmetric C-C bond formation, Chem. Eur. J. 8 (2002) 5288-5295.

[3] P. Siegert, M. J. McLeish, M. Baumann, H. Iding, M. M. Kneen, G. L. Kenyon, and M. Pohl, Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*, Protein Eng. Des. Sel. 18 (2005) 345-57.

[4] M. Pohl, P. Siegert, K. Mesch, H. Bruhn, and J. Grötzinger, Active site mutants of pyruvate decarboxylase from *Zymomonas mobilis* - a site-directed mutagenesis study of L112, I472, I476, E473, and N482, Eur. J. Biochem. 257 (1998) 538-546.

[5] M. Pohl, Protein design on pyruvate decarboxylase (PDC) by site-directed mutagenesis. Application to mechanistical investigations, and tailoring PDC for the use in organic synthesis, Adv. Biochem. Eng. Biotechnol. 58 (1997) 15-43.

[6] D. Gocke, C. Berthold, T. Graf, H. Brosi, I. Frindi-Wosch, M. Knoll, T. Stillger, L. Walter, M. Müller, J. Pleiss, G. Schneider, and M. Pohl, Pyruvate decarboxylase from *Acetobacter pasteurianus*: Biochemical and structural characterisation, (2007) *submitted*.

[7] D. Gocke, C. L. Nguyen, M. Pohl, T. Stillger, L. Walter, and M. Müller, Branched-chain ketoacid decarboxylase from *Lactococcus lactis* (KdcA), a valuable thiamin diphosphate-dependent enzyme for asymmetric C-C bond formation, Adv. Synth. Catal. 349 (2007) 1425-1435.

[8] A. Yep, G. L. Kenyon, and M. J. McLeish, Determinants of substrate specificity in KdcA, a thiamin diphosphate-dependent decarboxylase, Bioorg. Chem. 34 (2006) 325-336.

[9] P. Dünkelmann, D. Kolter-Jung, A. Nitsche, A. S. Demir, P. Siegert, B. Lingen, M. Baumann, M. Pohl, and M. Müller, Development of a donor-acceptor concept for enzymatic cross-coupling reactions of aldehydes: the first asymmetric cross-benzoin condensation, J. Am. Chem. Soc. 124 (2002) 12084-12085.

[10] H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir, and M. Pohl, Benzoylformate decarboxylase from *Pseudomonas putida* as stable catalyst for the synthesis of chiral 2-hydroxy ketones, Chem. Eur. J. 6 (2000) 1483-1495.

[11] T. Dünnwald, A. S. Demir, P. Siegert, M. Pohl, and M. Müller, Enantioselective synthesis of (*S*)-2-hydroxypropanone derivatives by benzoylformate decarboxylase catalyzed C-C bond formation, Eur. J. Org. Chem. (2000) 2161-2170.

[12] P. Domínguez de María, M. Pohl, D. Gocke, H. Gröger, H. Trauthwein, T. Stillger, and M. Müller, Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes, Eur. J. Org. Chem. (2007) 2940-2944.

[13] T. Stillger, M. Pohl, C. Wandrey, and A. Liese, Reaction engineering of benzaldehyde lyase catalyzing enantioselective C-C bond formation, Org. Proc. Res. Develop. 10 (2006) 1172-1177.

[14] E. Janzen, M. Müller, D. Kolter-Jung, M. M. Kneen, M. J. McLeish, and M. Pohl, Characterization of benzaldehyde lyase from *Pseudomonas fluorescens* - a versatile enzyme for asymmetric C-C-bond formation, Bioorg. Chem. 34 (2006) 345-361.

[15] F. Hildebrand, S. Kühl, M. Pohl, D. Vasic-Racki, M. Müller, C. Wandrey, and S. Lütz, The production of (*R*)-2-hydroxy-1-phenyl-propan-1-one derivatives by benzaldehyde lyase from *Pseudomonas fluorescens* in a continuously operated membrane reactor, Biotechnol. Bioeng. 96 (2006) 835-843.

[16] M. Breuer, M. Pohl, B. Hauer, and B. Lingen, High-throughput assay of (*R*)-phenylacetylcarbinol synthesized by pyruvate decarboxylase, Anal. Bioanal. Chem. 374 (2002) 1069-1073.

[17] B. Lingen, J. Grötzinger, D. Kolter, M. R. Kula, and M. Pohl, Improving the carboligase activity of benzoylformate decarboxylase from *Pseudomonas putida* by a combination of directed evolution and site-directed mutagenesis, Protein Eng. 15 (2002) 585-593.

[18] M. E. Smith, U. Kaulmann, J. M. Ward, and H. C. Hailes, A colorimetric assay for screening transketolase activity, Bioorg. Med. Chem. 14 (2006) 7062-7065.

[19] I. Fleming, R. S. Roberts, and S. C. Smith, J. Chem. Soc. Perkin Trans. 1 (1998) 1215-1228.

[20] P. Domínguez de María, T. Stillger, M. Pohl, S. Wallert, K. Drauz, H. Gröger, H. Trauthwein, and A. Liese, Preparative enantioselective synthesis of benzoins and (*R*)-2-hydroxy-1-phenylpropanone using benzaldehyde lyase, J. Mol. Catal. B: Enzym. 38 (2006) 43-47.

[21] D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C. L. Berthold, G. Schneider, J. Pleiss, M. Müller, and M. Pohl, Rational Protein Design of ThDP-Dependent Enzymes: Engineering Stereoselectivity, ChemBioChem (2007) *accepted*.

[22] R. Mikolajek, A. C. Spiess, M. Pohl, S. Lamare, and J. Buchs, An Activity, Stability and Selectivity Comparison of Propioin Synthesis by Thiamine Diphosphate-Dependent Enzymes in a Solid/Gas Bioreactor, Chembiochem 8 (2007) 1063-1070.

[23] A. S. Demir, M. Pohl, E. Janzen, and M. Müller, 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 (2001) 633-635.

[24] G. Goetz, P. Iwan, B. Hauer, M. Breuer, and M. Pohl, Continuous production of (*R*)-phenylacetylcarbinol in an enzyme-membrane reactor using a potent mutant of pyruvate decarboxylase from *Zymomonas mobilis*, Biotechnol. Bioeng. 74 (2001) 317-325.

[25] Z. Vuralhan, M. A. Morais, S. L. Tai, M. D. Piper, and J. T. Pronk, Identification and characterization of phenylpyruvate decarboxylase genes in *Saccharomyces cerevisiae*, Appl. Environ. Microbiol. 69 (2003) 4534-4541.

Figure legends

Fig. 1: The reduction of 2,3,4-triphenyltetrazolium chloride (TTC) by 2-hydroxy ketones results in the formation of a red formazane.

Fig. 2: TTC assay monitoring the carboligation of acetaldehyde and benzaldehyde (**A**), propanal and benzaldehyde (**B**) and dimethoxyacetaldehyde and 3,5-dichlorobenzaldehyde (**C**) catalyzed by various wild type enzymes and variants. Plate **A** and **B** are performed with BFDwt, two BFD-variants and BALwt as a positive control. Plate **C** shows results for four wild type pyruvate decarboxylases (PDCs), *Ll*KdcA, *Sc*PhPDC, BALwt and four variants of the *Zm*PDC. As a negative control in the first column no catalyst was inserted. Each well contains 50 µL enzyme (0.7 mg/mL (**A** + **B**), 0.3 mg/mL (**C**)) and 50 µL substrate solution (36 mM of each substrate, 40% (v/v) DMSO in potassium phosphate buffer (50 mM, pH 7, 2.5 mM $MgSO_4$, 0.1 mM ThDP). The biotransformation was performed for 18 h at 30 °C.

Figures

Fig. 1

2,3,5-triphenyltetrazolium chloride   red formazane

Fig. 2

A

acetaldehyde

benzaldehyde

acetaldehyde + benzaldehyde

B

propanal

benzaldehyde

propanal+ benzaldehyde

C

dimethoxy-acetaldehyde

3,5-dichloro-benzaldehyde

dimethoxyacetaldehyde + 3,5-dichlorobenzaldehyde
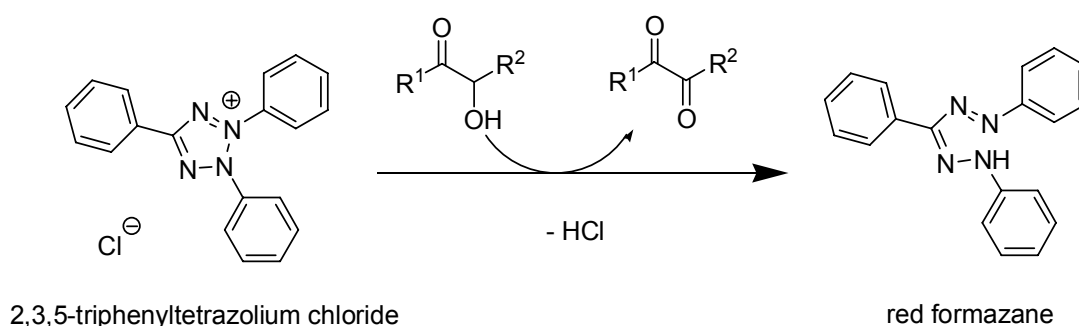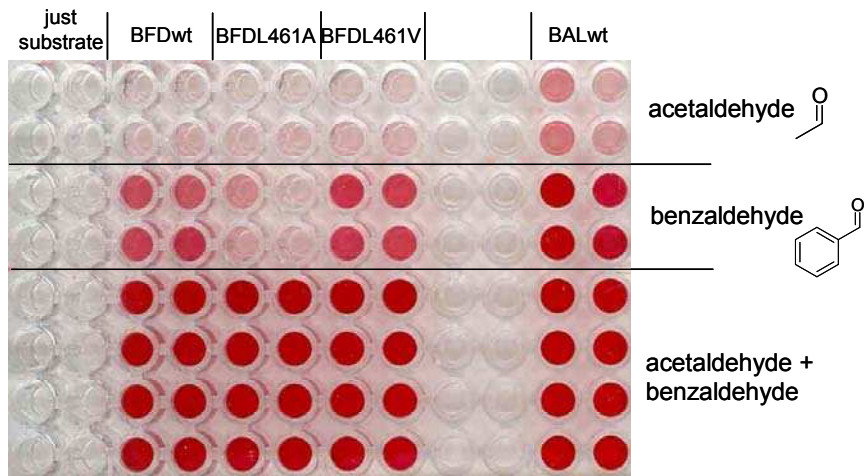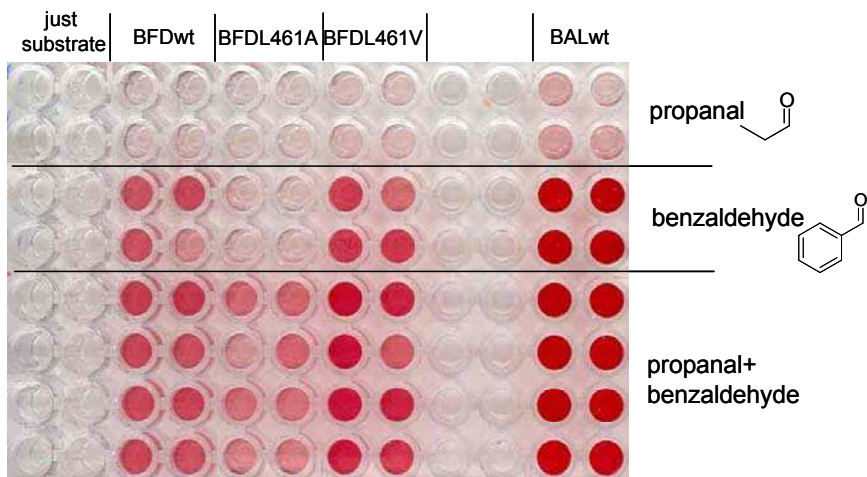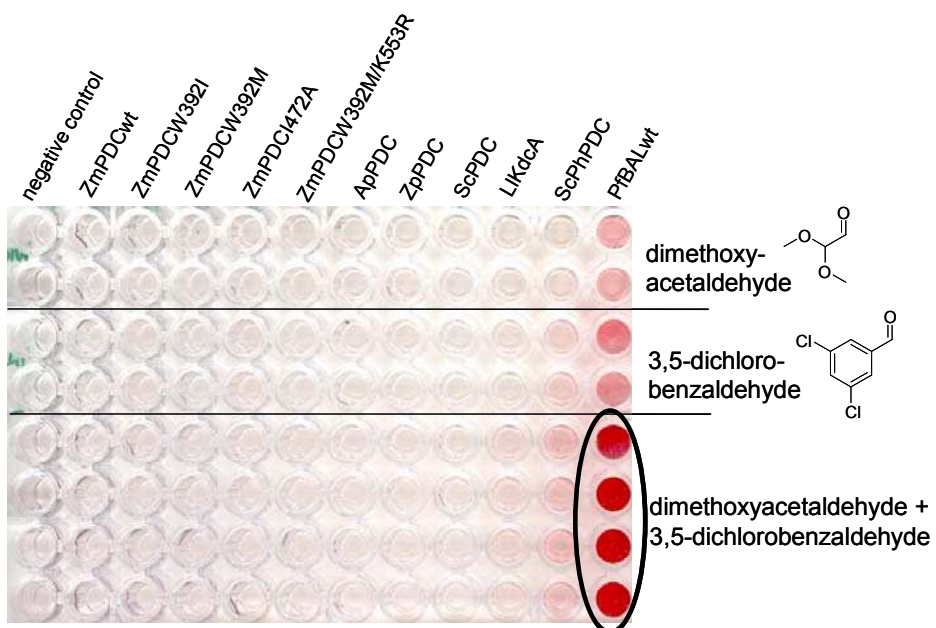
Table 1: Detection limits of various 2-hydroxy ketones.
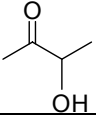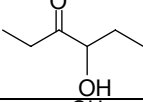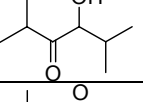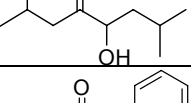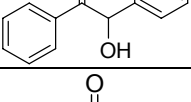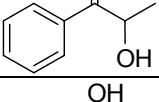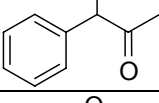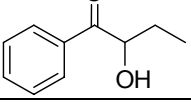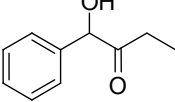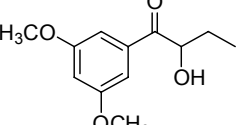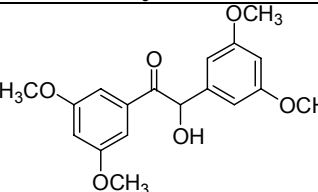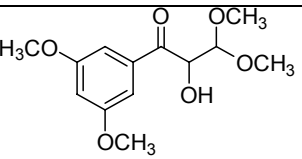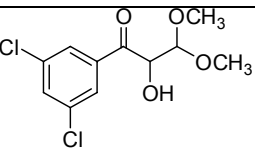
No	2-hydroxy ketone	name	lower detection limit [mM]
1		3-hydroxy-butan-2-one (acetoin)	0.5
2		4-hydroxy-hexan-3-one (propioin)	2
3		4-hydroxy-(2,5-dimethyl)-hexan-3-one	40
4		5-hydroxy-(2,7-dimethyl)-octan-4-one	5
5		2-hydroxy-1,2-diphenyl-ethanone (benzoin)	0.5
6		2-hydroxy-1-phenyl-propan-1-one, 2-hydroxy propiophenone (2-HPP)	0.5
7		1-hydroxy-1-phenyl-propan-2-one, phenylacetylcarbinol (PAC)	0.5
8		2-hydroxy-1-phenyl-butan-1-one	0.5
9		1-hydroxy-1-phenyl-butan-2-one	0.5
10		1-(3,5-dimethoxy-phenyl)-2-hydroxy-butan-1-one	0.5
11		1,2-bis-(3,5-dimethoxy-phenyl)-2-hydroxy-ethanone	0.5-1
12		1-(3,5-dimethoxy-phenyl)-2-hydroxy-3,3-dimethoxy-propan-1-one	1
13		1-(3,5-dichloro-phenyl)-2-hydroxy-3,3-dimethoxy-propan-1-one	n.d.

Table 2: Carboligation data obtained in analytical and preparative scale experiments with different BFD variants and BALwt. Products were identified by GC, HPLC, and NMR.

substrates	enzymes	products			
		acetoin (**1**)	mixed product (**6**)	mixed product (**7**)	benzoin (**5**)
acetaldehyde	BFDwt[a]	9.5%	-	-	-
	BFDL461A[a]	6%	-	-	-
	BFDL461V[a]	15%	-	-	-
	BALwt[d]	>90%	-	-	-
benzaldehyde	BFDwt[a]	-	-	-	23%
	BFDL461A[a]	-	-	-	4%
	BFDL461V[a]	-	-	-	27%
	BALwt[d]	-	-	-	90%
acetaldehyde + benzaldehyde	BFDwt[b]	-	81%	-	4%
	BFDL461A[b]	-	79%	-	-
	BFDL461V[b]	-	85%	-	-
	BALwt[c]	-	75%	-	-
		propioin (**2**)	mixed product (**8**)	mixed product (**9**)	benzoin (**5**)
propanal + benzaldehyde	BFDwt[b]	8%	6%	12%	8%
	BFDL461A[b]	6%	21.5%	-	1.5%
	BFDL461V[b]	4.9%	14.5%	8%	17.7%
	BALwt[c]	-	89%	11%	-

[a] Conversion after 46 h [21]. Reaction conditions: 40 mM acetaldehyde or 20 mM benzaldehyde in potassium phosphate buffer (50 mM, pH 7.5, containing 2.5 mM $MgSO_4$, 0.1 mM ThDP) 20% (v/v) DMSO, 0.3 mg/mL purified enzyme, 30 °C, 100 rpm. Quantification was done by GC or HPLC [21].

[b] Relative product distribution after 72 h conversion determined in mol% by NMR [21]. Reaction conditions: 18 mM of each substrate in potassium phosphate buffer (50 mM, pH 7.0, containing 2.5 mM $MgSO_4$, 0.1 mM ThDP) 20% (v/v) DMSO, 0.3 mg/mL purified enzyme, 30 °C, 100 rpm.

[c] Like [b] but at pH 8.0 and 3-fold excess of aliphatic aldehyde.

[d] Conversions obtained from preparative carboligation studies after product isolation. In these studies 50 mM of acetaldehyde [12], respectively 5 times 100 mM benzaldehyde [20] and 500 U BALwt were employed. Quantification was done by GC. Note that in these two reactions the amount of employed BALwt is significantly higher than in the TTC-assay, yielding into higher conversions than obtained in the microtiter plates.

Publication II

Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes.

Domínguez de María, P., Pohl, M., Gocke, D., Gröger, H.,
Trauthwein, H., Walter, L. & Müller, M.

Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes

Pablo Domínguez de María,[a] Martina Pohl,[b] Dörte Gocke,[b] Harald Gröger,[a] Harald Trauthwein,[a]* Thomas Stillger,[c] Lydia Walter,[c] and Michael Müller[c]*

[a] Degussa AG, Service Center Biocatalysis, Rodenbacher Chaussee 4, 63457 Hanau-Wolfgang, Germany
Corresponding Author: Dr. Harald Trauthwein. Tel: ++49-6241-402-5220; Fax: ++49-6151-1884-5220. E-mail: Harald.trauthwein@degussa.com
[b] Institut für Molekulare Enzymtechnologie, Forschungszentrum Jülich, 52426 Jülich, Germany
[c] Institut für Pharmazeutische Wissenschaften, Albert-Ludwigs-Universität, 79104 Freiburg, Germany

Dedicated to Professor Maria-Regina Kula on the occasion of her 70th birthday

Keywords: Enzyme catalysis; carboligation; benzaldehyde lyase; benzoylformate decarboxylase; aldehydes; 2-hydroxy ketones.

Abstract.
Benzaldehyde lyase (BAL) and benzoylformate decarboxylase (BFD) catalyze the asymmetric ligation of aliphatic aldehydes to enantiomerically enriched 2-hydroxy ketones. Carboligation of linear aldehydes with both enzymes results in high conversions and enantioselectivities of up to 80% ee. In case of branched aliphatic aldehydes BAL enables the carboligation of 3-methyl-butanal with high conversion and enantioselectivity of 89% ee.

Introduction.
Thiamindiphosphate (ThDP) dependent enzymes catalyze different type of reactions in biosynthesis.[1] It is well established that the asymmetric carboligation of two aromatic aldehydes (benzoin condensation) or the cross-carboligation between aromatic and aliphatic aldehydes can be performed with high regio- and stereocontrol using ThDP-dependent enzymes.[2] However, the asymmetric benzoin-type condensation of aliphatic aldehydes has been described rarely using either enzymatic or non-enzymatic methods. In this regard, Schmauder and Gröger described the whole-cell (*Saccharomyces cerevisiae*) catalyzed formation of acetoin and mixed aliphatic acyloins.[3] Chen and Jordan demonstrated that pyruvate decarboxylase (PDC) is the active catalyst in the formation of acetoin.[4] Later on, Bringer-Meyer and Sahm[5] and Crout et al.[6] described acyloin formation using PDC from *Zymomonas mobilis*. Berger et al. reported similar transformations using different yeast strains or isolated PDC.[7] In most of these transformations acetaldehyde, used directly or via decarboxylation of pyruvate, reacts predominantly as the donor substrate.

Non-enzymatic thiazolium- and triazolium-catalyzed asymmetric acyloin formation has been described, e.g., by Knight and Leeper[8] and by Enders and Breuer.[9] In both cases, however, enantiomeric excess of the products was below 33%. Other non-enzymatic methods for the enantioselective synthesis of aliphatic acyloins have been developed, e.g. using the Sharpless asymmetric dihydroxylation[10] or selective oxidation of vicinal diols.[11]
Moreover, the lipase-catalyzed racemic resolution of 2-hydroxy ketones[12] has been applied to aliphatic acyloins,[13a] as well as other enzymatic approaches like diketone reductions.[13b-d] The use of aliphatic aldehydes as substrates via a benzoin-type condensation enables access to enantiomerically pure acyloins in 100% theoretical yield.
Since the coupling with aromatic aldehydes has emerged in the last years as a powerful tool for the synthesis of 2-hydroxy ketones, we envisaged a similar transformation with aliphatic aldehydes. As enzymes we chose the ThDP-dependent enzymes benzaldehyde lyase (BAL) from *Pseudomonas fluorescens* biovar I (EC 4.1.2.38),[14] and benzoylformate decarboxylase (BFD) from *Pseudomonas putida* (EC 4.1.17)[15] which have proven to be very efficient catalysts for the production of benzoins

and 2-hydroxy propiophenones.[16,17] In the case of BFD the carboligation of two molecules of aliphatic aldehydes to yield acetoin (13% ee) or 2-cyclohexyl-2-hydroxypropanone (61% ee) has been reported by some of us.[18] The BAL-catalyzed formation of 4-hydroxy-3-octanone has been published recently.[19]

Herein we describe the enzymatic carboligation of linear and branched aliphatic aldehydes to yield aliphatic 2-hydroxy ketones in high yields and moderate-to-high enantiomeric excesses.

Results & Discussion.

The biocatalytic carboligation of aliphatic aldehydes was performed according to the reaction conditions for the synthesis of acetoin by the coupling of acetaldehyde (see Scheme 1).[15]
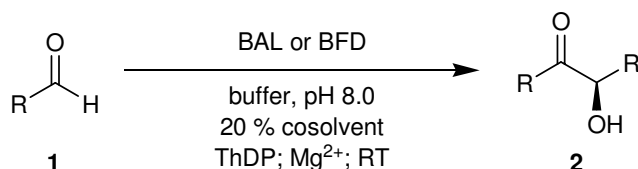
Scheme 1. Enzymatic carboligation of linear aliphatic aldehydes.

The reaction was performed in aqueous buffer with isolated enzyme (BAL or BFD respectively) in the presence of the cofactor ThDP and Mg^{2+}. Since cosolvents are assumed to influence the activity and selectivity, we included the investigation of DMSO and 2-propanol as cosolvents in our studies. The results for a range of linear and branched aldehydes with both enzymes are summarized in Table 1.

The first experiment we conducted was the carboligation of acetaldehyde (**1a**) (entry 1, 2). Both, BAL and BFD gave (*R*)-acetoin (**2a**) with high conversions of more than 90%. In both cases the enantioselectivities were low ($\leq$ 40% ee) which is in agreement with the literature.[15] Enantiomeric excess was determined by chiral GC, showing that BAL and BFD produce the same favoured enantiomer as PDC from *Saccharomyces cerevisiae*, leading to the assignment of the absolute configuration of acetoin (**2a**).[6] The tested cosolvents do not influence significantly the enantioselectivity of BAL and BFD in this reaction.

By taking propanal (**1b**) the desired product 4-hydroxy-hexan-3-one (**2b**) is formed with both enzymes BAL and BFD in conversions of more than 90% (entries 3-6). BAL gave the (*S*)-acyloin **2b** as the major enantiomer, the absolute configuration of which was asigned by the positive Cotton effect in the circular dichroism spectrum (CD) at 278 nm. For several aliphatic hydroxy ketones it has been shown that the positive and negative CD bands correspond to the (*S*)- and (*R*)-enantiomers, respectively.[20] Whereas the (*S*)-enantiomer is formed preferentially with BAL ($\leq$ 60% ee) BFD favours the opposite (*R*)-enantiomer with an ee of 63% (shown by chiral GC). The presence of cosolvents results in reduced enantioselectivity when using BAL, showing that this transformation is influenced considerably by external parameters.

n-Butanal (**1c**) is the most suitable substrate in this series of linear aliphatic aldehydes concerning enantioselectivity (entries 7-10). Without cosolvent the enantioselectivity is 50% ee in the BAL-catalyzed reaction forming (*R*)-5-hydroxy-octan-4-one (**2c**). Addition of DMSO enhances slightly the enantioselectivity to 55% ee. Remarkably, by adding 2-propanol the enantioselectivity is raised to 80% ee, indicating that reaction engineering can be used successfully for enhancing enantioselectivity in such transformations. Both, BAL and BFD produce the same enantiomer of **2c** (chiral GC) as is the case for *n*-pentanal (**1d**) as substrate. The absolute (*R*)-configuration was assigned to **2c** and **2d** using circular dichroism (negative band at 278 nm).

In all BFD-catalyzed reactions there is no significant influence of the cosolvent (data not shown) on the selectivity of BFD. The positive influence of 2-propanol on BAL is also confirmed when using *n*-pentanal (**1d**) as a substrate (**entries 11-13**). The enantioselectivity is enhanced from 30% ee (no cosolvent or DMSO) to 60% ee at a conversion of more than 90% for (*R*)-6-hydroxy-decan-5-one (**2d**). BFD furnishes enantioselectivities of about 65% ee at a high conversion of > 90%.

Table 1. Carboligation of linear aldehydes catalyzed by BAL and BFD. Enantiomeric excesses were determined via separation of compounds **2** by GC using chiral phases.

Entry	Aldehyde (R =)	Product	Enzyme (20% cosolvent)	Conv. [%]	ee [%] (**2**)
1	(**1a**) acetaldehyde (CH$_3$)	**2a**	BAL	> 90	40 (*R*)
2		**2a**	BFD	> 90	34 (*R*)
3		**2b**	BAL	> 90	60 (*S*)
4	(**1b**) propanal (CH$_3$CH$_2$)	**2b**	BAL (DMSO)	n.d.	53 (*S*)
5		**2b**	BAL (2-propanol)	> 90	20 (*S*)
6		**2b**	BFD	> 90	63 (*R*)
7		**2c**	BAL	> 90	50 (*R*)
8	(**1c**) *n*-butanal (CH$_3$(CH$_2$)$_2$)	**2c**	BAL (DMSO)	> 90	55 (*R*)
9		**2c**	BAL (2-propanol)	> 90	80 (*R*)
10		**2c**	BFD	> 90	80 (*R*)
11		**2d**	BAL	> 90	30 (*R*)
12	(**1d**) *n*-pentanal (CH$_3$(CH$_2$)$_3$)	**2d**	BAL (2-propanol)	> 90	60 (*R*)
13		**2d**	BFD	> 90	65 (*R*)
14	(**1e**) *iso*-butyraldehyde ((CH$_3$)$_2$CH)	**2e**	BAL	< 1	-
15		**2e**	BFD	< 1	-
16	(**1f**) pivaldehyde ((CH$_3$)$_3$C)	**2f**	BAL	< 1	-
17		**2f**	BFD	< 1	-
18		**2g**	BAL	> 90	89 (*R*)
19	(**1g**) *iso*-valeraldehyde	**2g**	BAL (DMSO)	> 90	86 (*R*)
20	((CH$_3$)$_2$CHCH$_2$)	**2g**	BFD	60	85 (*R*)
21		**2g**	BFD (DMSO)	n.d.	86 (*R*)
22	(**1h**) *tert*-butylacetaldehyde	**2h**	BAL	< 1	-
23	((CH$_3$)$_3$CCH$_2$)	**2h**	BFD	< 1	-

n.d.: not determined.

Thus, we could show that linear aliphatic 2-hydroxy ketones are accessible by enzymatic carboligation. Next, we studied the applicability of BAL- and BFD-catalyzed carboligation with branched aliphatic aldehydes. The existence of additional methyl groups close to the aldehyde group is regarded as a higher steric hindrance for the enzymes, which should reduce the reaction rate, but might result in a better enantioselectivity. We varied the degree of steric hindrance by using *iso*-propyl or *tert*-butyl groups. These groups were adjacent to the carbonyl group or were separated from the carbonyl group by a methylene spacer.

No 2-hydroxy ketone could be observed in the BFD-catalyzed carboligation with the branched aliphatic aldehydes *iso*-butyraldehyde (**1e**), pivaldehyde (**1f**), and *tert*-butyl acetaldehyde (**1h**). Wildtype BFD seems to be very sensitive to the steric demand of the aldehyde, as this is known from aromatic aldehydes as well.[21] Only in the case of *iso*-valeraldehyde (**1g**) (**entry 20, 21**) product formation was observed resulting in the same enantiomer as by using BAL. Thus, taking *iso*-valeraldehyde (**1g**) as a substrate, the formation of (*R*)-5-hydroxy-2,7-dimethyloctan-4-one (**2g**) was observed with an enantioselectivity of up to 89% ee. The absolute configuration of **2g** was established as (*R*) by CD (negative band at 278 nm) of the crude product. Additionally, the ee and absolute configuration of **2g** were confirmed by Mosher ester methodology (NMR and separation of diastereomers via GCMS). The high enantioselectivity might be caused by the steric demanding *iso*-propyl group. The other tested branched aliphatic substrates with *iso*-butyraldehyde (**1e**), pivaldehyde (**1f**), and *tert*-butyl acetaldehyde (**1h**) did not react in the presence of BAL. Here the steric demand is even higher compared to **1g** as no methylene spacer or even a *tert*-butyl group is present. Following the course of reactions of BAL and BFD in detail demonstrates the differences in activities of both enzymes. Employing 63 mM of **1g** the reaction catalyzed by BAL was completed after 5.5 h (Fig. 1), whereas the reaction catalyzed by BFD is significantly slower. The calculated space-time yields are 0.26 g L^{-1} d^{-1} (BFD) and 13.15 g L^{-1} d^{-1} (BAL) using 0.3 mg mL^{-1} enzyme.
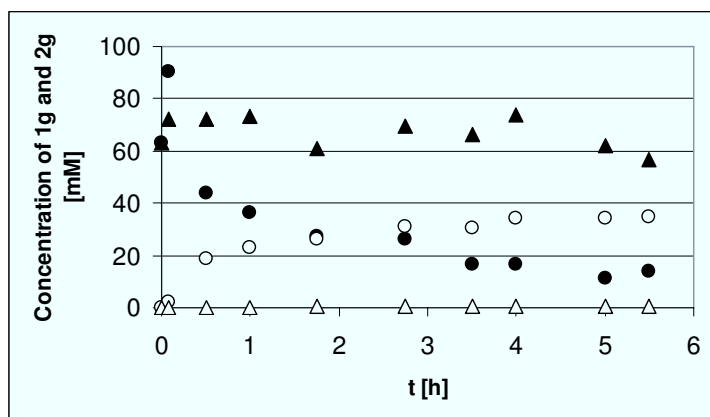
Figure. 1: Following the time course of 5-hydroxy-2,7-dimethyloctan-4-one (**2g**, open symbols) from *iso*-valeraldehyde (**1g**, closed symbols) catalyzed by purified BFD (triangels) and BAL (circles), respectively.

In order to confirm the racemic resolution of the enzymatically produced 2-hydroxy ketones (chiral GC) we synthesized compounds **2e** and **2g** via non-enzymatic acyloin condensation according to Fleming et al.[22] Analytical data (NMR, GCMS) of the enzymatic products and racemic material were identical in all respect. In case of **2b** the NMR data (enzymatic products) are not in accordance with literature values.[23] The correct structure of **2b** was confirmed by COSY and HSQC experiments.

Conclusions.

In summary, regarding the recently introduced donor-acceptor concept by Müller et al.[24] it has been shown, that aliphatic aldehydes are both suitable donors and acceptors as far as the steric hindrance is not too high. These examples also show that the substrate spectrum of these C-C-ligating ThDP-dependent enzymes is broader than described in the literature so far. Thus, further interesting substrates are expected to be applied in this attractive carboligating reaction. Through combination of reaction engineering and random and/or site-directed mutagenesis enantioselectivity might even be increased further.[25]

Experimental section.

Chemicals and biocatalysts: All reagents were commercially available from Sigma-Aldrich and were used without further purification. Benzaldehyde lyase from *Pseudomonas fluorescens* Biovar I and benzoylformate decarboxylase from *Pseudomonas putida* were produced in *E. coli* cells.[26] The preparation of BAL and BFD was done by sonication (ca. 3-5 min) of 2 g of the corresponding *E. coli* cells, which were dissolved in 20 mL of phosphate buffer 50 mmol L^{-1}, with 2.5 mmol L^{-1} $MgSO_4$ and 0.3 mmol L^{-1} ThDP. The disrupted cells were centrifuged at 4000 rpm during 20 minutes at 4 °C. The pellet was removed and the supernatant was applied for catalysis. The biocatalytic characterization was carried out as described in literature.[19,27] The volumetric activity of this solution was determined to 100 U mL^{-1}. One unit (U) of BAL activity is defined as the amount of enzyme which catalyzes the cleavage of 1 μmol benzoin (1.5 mmol L^{-1}) into benzaldehyde in potassium phosphate buffer (20 mmol L^{-1}, pH 7.0), containing $MgSO_4$ (2.5 mmol L^{-1}), ThDP (0.15 mmol L^{-1}) and PEG 400 (15 vol%) at 30°C. 1 Unit of BFD is defined as the amount of enzyme which catalyzes the cleavage of 1 μmol benzoylformate (30 mmol L^{-1}) into benzaldehyde and CO_2 in potassium phosphate buffer, 50 mmol L^{-1}, pH 6.5, containing $MgSO_4$ (2.5 mmol L^{-1}), ThDP (0.1 mmol L^{-1}) at 30°C.

Typical carboligation procedure: Variable amounts of BAL or BFD (500 U) were added to a 50 mL mixture of potassium phosphate buffer (50 mmol L^{-1}, with 2.5 mmol L^{-1} $MgSO_4$ and 0.3 mmol L^{-1} ThDP, pH 8.0) in the absence and presence of 20 vol% cosolvent (DMSO, 2-propanol) and the substrate (50 mmol L^{-1}) The reaction mixture was stirred for 24 h at room temperature. For the work-up, the reaction media was extracted with dichloromethane (4 x 50 mL) and the organic phase was washed several times with water. After drying the organic layer with $MgSO_4$ the solvent was removed under reduced pressure. The products were achieved practically pure without need of further purification.

In the case of 5-hydroxy-2,7-dimethyloctan-4-one (**2g**) the reaction course was followed in order to determine space-time yields for BAL and BFD. The reaction was performed using 63 mmol L^{-1} *iso*-valeraldehyde (**1g**) in 50 mmol L^{-1} potassium phosphate buffer with 2.5 mmol L^{-1} MgSO$_4$ and 0.1 mmol L^{-1} ThDP, pH 8.0, 30°C, 20% DMSO, and 0.3 mg purified BAL and BFD, respectively, per mL reaction volume. Samples taken every 0.5-1 h were analyzed by GC.

Analytical performance: The conversion and the enantiomeric excess were followed by chiral GC, employing a Chirasil-DEX CB (Varian), 25 m x 0.32 mm, or a FS Lipodex D 50 m x 0.25 mm, with a FID detector. NMR spectra were recorded on a Bruker DPX-400. Chemical shifts are reported in ppm relative to CHCl$_3$ (^{1}H NMR: δ = 7.27 ppm), CDCl$_3$ (^{13}C NMR: δ = 77.0 ppm) as internal standards. GCMS spectra were determined on a HP 6890 series GC-system fitted with a HP 5973 mass selective detector (Hewlett Packard; column HP-5MS, 30 m·250 µm; T_{GC}(injector) = 250 °C, T_{MS}(ion source) = 200 °C, time program (oven): $T_{0\ min}$ = 60 °C, $T_{3\ min}$ = 60 °C, $T_{14\ min}$ = 280 °C (heating rate 20 °C·min^{-1}), $T_{19\ min}$ = 280 °C). CD-spectra were recorded on a JASCO J-810 Spectropolarimeter using acetonitrile as solvent.

3-Hydroxy-2-butanone (**2a**): Colorless liquid. Conversion: > 90%. GC (Lipodex D, $T_{injector}$ = 190 °C, T_{column} 70 °C) R_t (*R*) = 13.4 min, R_t (*S*) = 18.3 min (ee 40% (*R*) BAL; 34% (*R*) BFD). ^{1}H-NMR: δ (ppm) = 1.4 (d, *J* = 7.3 Hz, 3H), 2.2 (s, 3H), 3.5 (s, 1H, OH), 4.25 (m, 1H).

4-Hydroxy-3-hexanone (**2b**): Colorless liquid. Conversion: > 90%. 54% isolated yield (BAL, without cosolvent). GC (Lipodex D, $T_{injector}$ = 190 °C, T_{column} 60 °C) R_t (*R*) = 30.6 min, R_t (*S*) = 31.8 min (ee 20-60% (*S*) BAL; 60% (*R*) BFD). ^{1}H-NMR (400 MHz, CD$_2$Cl$_2$): δ (ppm) = 0.95 (t, *J* = 7.3 Hz, 3H), 1.12 (t, *J* = 7.3 Hz, 3H), 1.61 (m, *J* = 14.4, 7.3, 6.9 Hz, 1H), 1.91 (dqd, *J* = 14.4, 7.3, 4.0 Hz, 1H), 2.47 (dq, *J* = 17.9, 7.3 Hz, 1H), 2.55 (dq, *J* = 17.9, 7.3 Hz, 1H),3.4 (br, 1H), 4.17 (dd, *J* = 6.9, 4.0 Hz, 1H). ^{13}C-NMR (100 MHz, CD$_2$Cl$_2$): δ (ppm) = 7.2, 8.6 (CH$_3$), 26.8, 30.9 (CH$_2$), 76.9 (CHOH), 212.8 (C=O). GCMS R_t = 4.13 min. MS (EI, 70 eV): *m/z* = 116 (M$^+$,3%), 87 (2), 73 (2), 59 (100).

5-Hydroxy-4-octanone (**2c**): Yellow liquid. Conversion: > 90%. 71% isolated yield (BAL, DMSO as cosolvent). GC (Lipodex D, $T_{injector}$ = 190 °C, T_{column} 70 °C) R_t (*R*) = 46.2 min, R_t (*S*) = 49.5 min (ee 50-80% (*R*) BAL; 80% (*R*) BFD). ^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 0.93 (t, *J* = 7.4 Hz, 3H), 0.94 (t, *J* = 7.2 Hz, 3H), 1.34 (m, 3H), 1.65 ('sextett', *J* = 7.3 Hz, 2H), 1.72-1.82 (m,1H), 2.36-2.51 (m, 2H), 4.16 (t, *J* = 3.5 Hz, 1H). ^{13}C-NMR (100 MHz, CDCl$_3$): δ (ppm) = 13.7, 13.8 (CH$_3$), 17.0, 18.1, 35.7, 39.7 (CH$_2$), 76.2 (CHOH), 212.2 (C=O). GCMS R_t = 6.27 min. MS (EI, 70 eV): *m/z* = 144 (M$^+$ 1%), 102 (12), 73 (66), 55 (100).

6-Hydroxy-5-decanone (**2d**): Colorless liquid. Conversion: > 90%. 91% isolated yield (BAL, DMSO as cosolvent). GC (Lipodex D, $T_{injector}$ = 190 °C, T_{column} 70 °C) R_t (*R*) = 99.5 min, R_t (*S*) = 99.7 min (ee 30-60% (*R*) BAL; 65% (*R*) BFD). ^{1}H-NMR (400 MHz, CDCl$_3$): δ (ppm) = 0.928 (t, *J* = 7.1 Hz, 3H), 0.933 (t, *J* = 7.3 Hz, 3H), 1.28-1.67 (m, 9H), 1.8-1.9 (m, 1H), 2.4-2.55 (m, 2H), 4.18 (dd, 1H, *J* = 7.4, 3.7 Hz, C*H*OH). ^{13}C-NMR (100 MHz, CDCl$_3$): δ (ppm) = 13.7, 13.8 (CH$_3$), 22.3, 22.5, 25.7, 26.9, 33.4, 37.5 (CH$_2$), 76.3 (CHOH), 212.5 (C=O). GCMS R_t = 7.99 min. MS (EI, 70 eV): *m/z* = 172 (M$^+$, 0.1%), 116 (5%), 87 (35%), 69 (100%), 57 (31%).

2,5-Dimethyl-4-hydroxy-3-hexanone (**2e**) (nonenzymatic synthesis): colorless liquid. ^{1}H-NMR (400 MHz, CDl$_3$): δ (ppm) = 0.70 (d, *J* = 6.7 Hz, 3 H), 1.11 (d, *J* = 6.8 Hz, 3 H), 1.12 (d, *J* = 6.8 Hz, 3 H), 1.13 (d, *J* = 6.8 Hz, 3H), 2.18 ('septett', *J* = 6.8 Hz, 1H), 2.80 (d'septett', *J* = 6.7, 2.5 Hz, 1H), 3.47 (br, 1H), 4.20 (br, 1H). ^{13}C-NMR (100 MHz, CDl$_3$): δ (ppm) = 14.9, 17.3, 19.7, 20.2 (CH$_3$), 31.1, 36.1 (CH), 79.1 (*C*HOH), 216.2 (C=O). GCMS R_t = 5.7 min. MS (EI, 70 eV): *m/z* = 144 (M$^+$, 3%), 102 (20), 87 (25), 73 (100), 55 (60).

2,7-Dimethyl-5-hydroxy-4-octanone (**2g**): Yellow liquid. Conversion: > 90%. 65% isolated yield (BAL, methyl-*tert*-butylether as cosolvent[28]). GC (Lipodex D, $T_{injector}$ = 190 °C, T_{column} 70 °C) R_t (*R*) = 82.0 min, R_t (*S*) = 86.1 min (ee 86-89% (*R*) BAL; 85% (*R*) BFD). ^{1}H-NMR (400 MHz, CDl$_3$): δ (ppm) = 0.94 (d, *J* = 6.7 Hz, 3H), 0.95 (d, *J* = 6.7 Hz, 3H), 0.97 (d, *J* = 6.6 Hz, 3H), 1.0 (d, *J* = 6.6 Hz, 3H), 1.39 (ddd, *J* = 13.7, 10.3, 4.2 Hz, 1H), 1.51 (ddd, *J* = 13.7, 9.6, 3.0 Hz, 1H), 1.96 (m, 1H), 2.23

(m, 1H), 2.36 (m, 2H), 3.51 (br, 1H), 4.15 (dd, J = 10.3, 3.0 Hz, 1H). ^{13}C-NMR (100 MHz, CDl$_3$): ☐ (ppm) = 21.2, 22.5, 22.6, 23.6 (CH$_3$), 24.6, 24.8 (CH), 42.6, 46.7 (CH$_2$), 75.3 (CHOH), 212.5 (C=O). GCMS R$_t$ = 7.51 min. MS (EI, 70 eV): m/z = 170 (M$^+$-2H, 10%), 116 (30), 85 (70), 69 (100), 57 (70).

Acknowledgments.

References.

[1] M. Pohl, G.A. Sprenger, M. Müller, *Current Opinion Biotechnol.* **2004**, *15*, 335-342.

[2] M. Pohl, B. Lingen, M. Müller, *Chem. Eur. J.* **2002**, *8*, 5288-5295.

[3] H.-P. Schmauder, D. Gröger, *Pharmazie* **1968**, *23*, 320-331.

[4] G.C. Chen, F. Jordan, *Biochemistry* **1984**, *23*, 3576-3582.

[5] S. Bringer-Meyer, H. Sahm, *Biocatalysis* **1988**, *1*, 321-331.

[6] D.H.G. Crout, H. Dalton, D.W. Hutchinson, M. Miyagoshi, *J. Chem. Soc., Perkin Trans. 1* **1991**, 1329-1334.

[7] a) F. Neuser, H. Zorn, R.G. Berger, *Z. Naturforsch.* **2000**, *55c*, 560-568; b) T. Kurniadi, R. Bel-Rhlid, L.-B. Fay, M.-A. Juillerat, R.G. Berger, *J. Agric. Food Chem.* **2003**, *51*, 3103-3107; c) F. Neuser, H. Zorn, R.G. Berger, *J. Agric. Food Chem.* **2000**, *48*, 6191-6195.

[8] R.L. Knight, F.J. Leeper, *Tetrahedron Lett.* **1997**, *38*, 3611-3614.

[9] a) D. Enders, K. Breuer, In *Comprehensive asymmetric catalysis*, E. N. Jacobsen, A. Pfaltz, H. Yamamoto, Eds.; Springer, 1999; pp 1093-1102; b) K. Breuer, *Doctoral thesis*, Technical University Aachen, **1997**.

[10] T. Hashiyama, K. Morikawa, K. B. Sharpless, *J. Org. Chem.* **1992**, *57*, 5067-5068.

[11] B. Plietker, *Org. Lett.* **2004**, *6*, 289-291.

[12] T.-H. Duh, Y.-F. Wang, M.-J. Wu, *Tetrahedron: Asymmetry* **1993**, *4*, 1793-1794.

[13] a) G. Scheid, W. Kuit, E. Ruijter, R.V.A. Orru, E. Henke, U. Bornscheuer, L.A. Wessjohann, *Eur. J. Org. Chem.* **2004**, 1063-1074; b) P. Besse, J. Bolte, A. Fauve, H. Veschambre, *Bioorg. Chem.*, **1993**, *21*, 342-353; c) R. Bel-Rhlid, A. Fauve, M.F. Renard, H. Veschambre, *Biocatalysis*, **1992**, *6*, 319-337; d) R. Bel-Rhlid, M.F. Renard, H. Veschambre, *Bull. Soc. Chim. Fr.* **1996**, *133*, 1011-1021.

[14] B. Gonzalez, R. Vicuña, *J. Bacteriol.* **1989**, *171*, 2401-2405.

[15] H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A.S. Demir, M. Pohl, *Chem. Eur. J.* **2000**, *6*, 1483-1495.

[16] A.S. Demir, M. Pohl, E. Janzen, M. Müller, *J. Chem. Soc., Perkin Trans. 1* **2000**, 633-635.

[17] A.S. Demir, O. Sesenoglu, E. Eren, B. Hosrik, M. Pohl, E. Janzen, D. Kolter, R. Feldmann, P. Dünkelmann, M. Müller, M. *Adv. Synth. Catal.* **2002**, *344*, 96-103.

[18] T. Dünnwald, A.S. Demir, P. Siegert, M. Pohl, M. Müller. *Eur. J. Org. Chem* **2000**, 2161-2170.

[19] E. Janzen, M. Müller, D. Kolter-Jung, M. M. Kneen, M. J. McLeish, M. Pohl, *Bioorg. Chem.* **2006**, *34*, 345-361.

[20] A. Baykal, S. Chakraborty, A. Dodoo, F. Jordan, *Bioorg. Chem.* **2006**, *34*, 380-393; and references cited therein.

[21] B. Lingen, D. Kolter, P. Dünkelmann, R. Feldmann, J. Grötzinger, M. Pohl, M. Müller, C*hembiochem* **2003**, *4*, 721-726.

[22] I. Fleming, R.S. Roberts, S.C. Smith, *J. Chem. Soc., Perkin Trans. 1* **1998**, 1215-1228.

[23] R.D. Chambers, J. Hutchinson, G. Sandford, A. Shah, J. Vaughan, *Tetrahedron* **1997**, *53*, 15833-15842. NMR data (CD$_2$Cl$_2$ as solvent) of **2b** align well with those published for 4-hydroxy-3-hexanone in CDCl$_3$ as solvent: K. Edegger, Doctoral Thesis, University of, Graz, Austria, **2005**.

[24] 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.

[25] M. T. Reetz, *Proc. Natl. Acad. Sci.* **2004**, *101*, 5716-5722.

[26] M. Pohl, M. Mueller, A. S. Demir, WO 2002002753 A2. **2002**.

[27] E. Janzen, Doctoral Thesis, Heinrich-Heine University of Duesseldorf, 2002.

[28] M. Villela Filho, T. Stillger, M. Müller, A. Liese, C. Wandrey, *Angew. Chem.* **2003**, *115*, 3101-3104; *Angew. Chem. Int. Ed.* **2003**, *42*, 2993-2996.

Publication III

Characterization of phenylpyruvate decarboxylase, involved in auxin production of *Azospirillum brasilense*.

Spaepen, S., Versées, W., Gocke, D., Pohl, M., Steyaert, J. & Vanderleyden, J.

Characterization of Phenylpyruvate Decarboxylase, Involved in Auxin Production of *Azospirillum brasilense*[▽]

Stijn Spaepen,[1] Wim Versées,[2,3] Dörte Gocke,[4] Martina Pohl,[4] Jan Steyaert,[2,3] and Jos Vanderleyden[1]*

*Centre of Microbial and Plant Genetics and INPAC, K.U. Leuven, Kasteelpark Arenberg 20, 3001 Heverlee, Belgium[1];
Department of Ultrastructure, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium[2]; Department of Molecular and
Cellular Interactions, VIB, Pleinlaan 2, 1050 Brussel, Belgium[3]; and Institute of Molecular Enzyme Technology,
Heinrich-Heine University Düsseldorf, Research Centre Jülich, 52426 Jülich, Germany[4]*

Azospirillum brasilense belongs to the plant growth-promoting rhizobacteria with direct growth promotion through the production of the phytohormone indole-3-acetic acid (IAA). A key gene in the production of IAA, annotated as indole-3-pyruvate decarboxylase (*ipdC*), has been isolated from *A. brasilense*, and its regulation was reported previously (A. Vande Broek, P. Gysegom, O. Ona, N. Hendrickx, E. Prinsen, J. Van Impe, and J. Vanderleyden, Mol. Plant-Microbe Interact. 18:311–323, 2005). An *ipdC*-knockout mutant was found to produce only 10% (wt/vol) of the wild-type IAA production level. In this study, the encoded enzyme is characterized via a biochemical and phylogenetic analysis. Therefore, the recombinant enzyme was expressed and purified via heterologous overexpression in *Escherichia coli* and subsequent affinity chromatography. The molecular mass of the holoenzyme was determined by size-exclusion chromatography, suggesting a tetrameric structure, which is typical for 2-keto acid decarboxylases. The enzyme shows the highest k_{cat} value for phenylpyruvate. Comparing values for the specificity constant k_{cat}/K_m, indole-3-pyruvate is converted 10-fold less efficiently, while no activity could be detected with benzoylformate. The enzyme shows pronounced substrate activation with indole-3-pyruvate and some other aromatic substrates, while for phenylpyruvate it appears to obey classical Michaelis-Menten kinetics. Based on these data, we propose a reclassification of the *ipdC* gene product of *A. brasilense* as a phenylpyruvate decarboxylase (EC 4.1.1.43).

Azospirillum brasilense is a gram-negative, nitrogen-fixing bacterium that lives in association with the roots of grasses and cereals (24). This bacterium belongs to the plant growth-promoting rhizobacteria with direct growth promotion. *Azospirillum* inoculation results in an increased number of plant root hairs and lateral roots, through which an improved uptake of water and nutrients can occur. This effect is primarily caused by the bacterial production of the phytohormone indole-3-acetic acid (IAA), the most abundant naturally occurring auxin (10). Tryptophan (Trp) is generally considered to be the precursor of IAA. The best-characterized pathways in bacteria for the conversion of Trp to IAA are the indole-3-acetamide pathway and the indole-3-pyruvate (IPyA) pathway (8, 25). In the IPyA pathway, Trp is first transaminated to IPyA, subsequently decarboxylated to indole-3-acetaldehyde, which can be oxidized to IAA. A gene (*ipdC*) encoding IPyA decarboxylase (IPDC) was isolated from *Enterobacter cloacae*, and a homologous gene has also been found in *A. brasilense* (7, 16). The characterization and structural analysis of IPDC from *E. cloacae* (IPDC$_{Ec}$) have been published by Schütz et al. (30, 31). IPDC$_{Ec}$ shows the highest catalytic efficiency with the natural substrate IPyA and low or no activity with pyruvate and phenylpyruvate. In contrast to some other thiamine diphosphate (ThDP)-dependent 2-keto acid decarboxylases, there were no

indications for substrate activation. In *A. brasilense* an *ipdC*-knockout mutant was found to produce less than 10% of the level of wild-type IAA production, indicating that the *ipdC* gene product is a key enzyme for IAA biosynthesis in this bacterium (7, 28). Expression analysis of *ipdC* revealed a strict regulation. The gene is expressed only in the late exponential phase and is induced by auxins (37, 38). Recently, Somers et al. demonstrated that the *ipdC* gene product is also involved in the biosynthesis of phenylacetic acid (PAA), an auxin with antimicrobial activity (34). An *A. brasilense ipdC*-knockout mutant shows significantly lower production of PAA (34). Here we report the detailed characterization of the *A. brasilense ipdC* gene product, which suggests its reclassification as a phenylpyruvate decarboxylase (PPDC).

MATERIALS AND METHODS

Strains and growth media. The strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were grown in Luria-Bertani (LB) medium at 37°C. *A. brasilense* strains were grown at 30°C in LB medium supplemented with 2.5 mM CaCl$_2$ and 2.5 mM MgCl (LB* medium) or in *Azospirillum* minimal medium with 0.5% (wt/vol) malate (39), phenylalanine or tryptophan as a carbon source, and 20 mM NH$_4$Cl or 0.5% (wt/vol) phenylalanine or tryptophan as a nitrogen source. The antibiotic kanamycin was used when appropriate at 25 μg ml^{-1}.

Growth experiments under different carbon and nitrogen sources. Bacterial strains were grown in shaken culture overnight in LB* medium supplemented with the appropriate antibiotics. Cells were pelleted and washed twice with an 0.85% (wt/vol) sodium chloride solution. The optical density at 600 nm was adjusted to 1, and the cultures were diluted 1,500-fold in minimal medium with different carbon and nitrogen sources. The growth was monitored every 30 min (600 nm) in a Bioscreen C apparatus for 72 h. The growth characteristics were identified by the growth rate constant calculated on the exponential part of the

* Corresponding author. Mailing address: Centre of Microbial and Plant Genetics, K.U. Leuven, Kasteelpark Arenberg 20, 3001 Heverlee, Belgium. Phone: 32 16321631. Fax: 32 16321963. E-mail: jozef.vanderleyden@biw.kuleuven.be.

TABLE 1. Strains and plasmids

Strain or plasmid	Relevant characteristic(s)	Reference or source
Strains		
Escherichia coli		
DH5α	*hsdR17 endA1 thi-1 gyrA96 relA1 recA1 supE44 ΔlacU169*(φ80*lacZΔM15*)	33
BL21-CodonPlus(DE3)-RP	*ompT hsdS*(r$_B^-$ m$_B^-$) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte [*argU proL* Cam^r]	Stratagene
Azospirillum brasilense		
Sp245	Wild-type strain	5
FAJ0009	Sp245 *ipdC*::Tn*5*	7
Plasmids		
pET28a	Km^r	Novagen
pCMPG9551	1.6-kb NdeI-XhoI fragment, carrying the open reading frame of the *ipdC* gene, in pET28a	This study

growth curve and defined by ln 2 g^{-1} (*g*, doubling time of an exponentially growing culture).

Construction of overexpression strain. The open reading frame of the *ipdC* gene of *A. brasilense* was amplified using PCR with the following primers: Azo-415 (5′-GATCCATATGAAGCTGGCCGAAGCCCTG-3′) and Azo-416 (5′-CTAGCTCGAGTTACTCCCGGGGCGCGGCGTG-3′), containing NdeI and XhoI restriction sites, respectively, at the 5′ end (underlined in nucleotide sequence). After digestion, the fragment was ligated in a pET28a vector (Novagen). *E. coli* DH5α was transformed with this construct (called pCMPG9551). After selection of a correct clone by a restriction digest, the expression host *E. coli* BL21-CodonPlus(DE3)-RP (Stratagene) was transformed with the purified plasmid. The sequence was verified on an ALF Sequencer (Amersham).

Expression and purification of recombinant enzyme. Two hundred fifty milliliters LB medium was inoculated with 12.5 ml of an overnight culture of BL21-CodonPlus(DE3)-RP(pCMPG9551) and grown at 37°C till it reached a turbidity (595 nm) of 0.5 to 0.7 (2 to 3 h). Protein expression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside. The cultures were grown at 30°C for a further 7 to 8 h. Cells were harvested by centrifugation at 4,000 × *g* for 20 min (4°C), and the pellets were stored at −20°C. For cell disruption, the pellets were resuspended in binding buffer (10 mM 2-morpholinoethanesulfonic acid [MES], 2.5 mM MgSO₄, 0.1 mM ThDP, 0.5 M KCl, 20 mM imidazole, pH 7.4) supplemented with 1 mg lysozyme per ml binding buffer. After incubation at 37°C for 30 min, the cells were sonicated (200 W; eight times for 10 s each). The cell lysate was cleared by centrifugation (10,000 × *g*; 25 min; 4°C), and the supernatant was applied to a HisTrap HP column (Amersham). After the column was washed with binding buffer, the recombinant protein was eluted using a linear gradient from 0 to 100% elution buffer (10 mM MES, 2.5 mM MgSO₄, 0.1 mM ThDP, 0.5 M KCl, 500 mM imidazole, pH 7.4). The protein eluted at an imidazole concentration of 250 mM. The elution fraction was dialyzed against the assay buffer (10 mM MES, 2.5 mM MgSO₄, 0.1 mM ThDP, pH 6.5), yielding a total dilution factor of 50,000. The protein-containing fractions, the molecular mass of the subunits, and the enzyme purity were determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by Coomassie blue staining, and the protein concentration was determined by direct optical measurement (ε_{280} = 36,840 M^{-1} cm^{-1}; determined using ProtParam tool; Swiss Institute of Bioinformatics [http://www.expasy.org/tools/protparam.html]) or by the Bradford method (6) using bovine serum albumin as a standard.

Size-exclusion chromatography. The molecular mass of the holoenzyme was estimated by size-exclusion chromatography on a Superdex G200 column, equilibrated with 10 mM MES, 2.5 mM MgSO₄, 0.5 mM ThDP, 150 mM KCl, pH 6.5 (flow rate of 1 ml min^{-1}). The molecular mass was estimated by comparing the elution volume (at least three repetitions) with those of known standard proteins (molecular mass in kDa is indicated in parentheses): RNase A (13.7), chymotrypsinogen A (25), ovalbumin (43), albumin (67), aldolase (158), catalase (232), ferritin (440), and thyroglobulin (669). Blue dextran 2000 was used to determine the void volume of the column.

Enzyme assay and kinetic analysis. The decarboxylation activity was measured using a coupled photometric assay with alcohol dehydrogenase as an auxiliary enzyme (43). A typical reaction mixture (pH 6.5) contained 10 mM MES, 2.5 mM MgSO₄, 0.1 mM ThDP, 0.35 mM NADH, and 0.125 U alcohol dehydrogenase from horse liver. To avoid interference with the substrates and products, the absorbance was measured at 366 nm ($\varepsilon_{366, NADH}$ = 2.87 mM^{-1} cm^{-1}) (30), using

a Cary 100 UV-Vis spectrophotometer (Varian Inc.). The reaction was followed for 15 min at 30°C. The reaction velocity was converted into the reaction rate (per monomer), and the data were plotted and fitted according to Michaelis-Menten kinetics or the simplified Hill equation using Origin software (version 7; OriginLab Corporation).

Phylogenetic analysis. The BLAST algorithm (1) with IPDC$_{Ec}$ and PPDC of *A. brasilense* (PPDC$_{Ab}$) as baits was used to identify (putative) bacterial PPDCs and IPDCs. Thirty-three amino acid sequences were selected for multiple sequence alignment with ClustalW (36). The alignment was further adapted with the Genedoc program (23) for publication. A phylogenetic tree based on the multiple sequence alignment was constructed using an UPGMA (unweighted-pair group method using average linkages) model with the MEGA software (18), and bootstrap values were used for reliability testing (1,000 replicates).

RESULTS

Growth experiments. The *A. brasilense* wild-type strain and *ipdC*-knockout mutant (FAJ0009) were grown in minimal medium with different nitrogen sources (Fig. 1). The two strains show the same growth characteristics in minimal medium with ammonium as the sole nitrogen source. The growth rate of the *ipdC*-knockout mutant (growth rate constant, 0.087 h^{-1}) in minimal medium with phenylalanine as the sole nitrogen source is, however, impaired compared to the wild-type strain (0.102 h^{-1}). Tryptophan is not a nitrogen source for the *A. brasilense* strains: the wild-type and *ipdC*-knockout mutant strains did not show any significant growth. *A. brasilense* was also unable to grow on minimal medium with the aromatic amino acid phenylalanine or tryptophan as the sole carbon source (data not shown).

Heterologous overexpression, purification, and size-exclusion chromatography of PPDC$_{Ab}$. Recombinant PPDC$_{Ab}$ was produced by cloning the open reading frame into an N-terminal His$_6$-tag-attaching overexpression vector. Heterologous overexpression in *E. coli* and subsequent Ni^{2+} chelate affinity chromatography resulted in isolation of recombinant PPDC$_{Ab}$. Expression level and purity were checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and the recombinant enzyme was purified in high yield and purity. A subunit molecular mass of about 60 kDa could be estimated from comparison with the standard proteins.

A native molecular mass of 240.6 kDa was determined by size-exclusion chromatography for the recombinant PPDC$_{Ab}$,
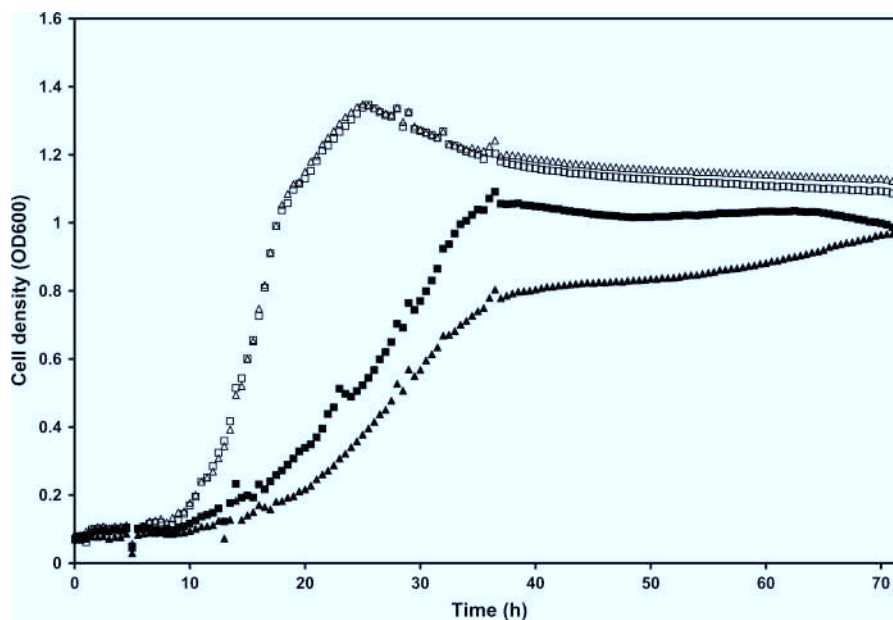
FIG. 1. Growth curves of *Azospirillum brasilense* and FAJ0009. Data represent means of five replicates of one experiment. Squares, *A. brasilense* Sp245; triangles, FAJ0009; open symbols, minimal medium for *Azospirillum* (MMAB) with ammonium as the sole nitrogen source; filled symbols, MMAB with phenylalanine as the sole nitrogen source.

which is consistent with the theoretical mass of a tetramer calculated from the amino acid sequence (240.0 kDa).

Substrate specificity of PPDC$_{Ab}$. The decarboxylase activity of the ThDP-dependent 2-keto acid decarboxylase PPDC$_{Ab}$ was followed using a coupled optical test. For substrates showing significant activity, the steady-state parameters were determined. Depending on the reaction rate versus substrate concentration curve, different models were used to analyze the data. The aromatic 2-keto acids phenylpyruvate, IPyA, 4-phenyl-2-oxobutanoic acid, and 5-phenyl-2-oxopentanoic acid are decarboxylated by PPDC$_{Ab}$ but with significantly different kinetic parameters (Table 2). No activity was observed with benzoylformate and indole-glyoxylic acid. The enzyme has a k_{cat}/K_m of 309 mM^{-1} min^{-1} for phenylpyruvate, showing apparently classical Michaelis-Menten kinetics (with a slight deviation from hyperbolicity: Hill coefficient [h] of 1.26 when fitted on the Hill equation). By contrast, the other aromatic substrates showed substrate activation, which is visible from the sigmoidal reaction rate-versus-substrate concentration curves and the deviation from linearity of Eadie-Hofstee plots

(Fig. 2). Therefore, the kinetic data were analyzed with the Hill model, yielding $k_{cat}/K_{0.5} = 32$ mM^{-1} min^{-1} for IPyA, 7 mM^{-1} min^{-1} for 4-phenyl-2-oxobutanoic acid, and 339 mM^{-1} min^{-1} for 5-phenyl-2-oxopentanoic acid. Hill coefficients for these aromatic substrates were in the range of 1.8 to 1.9. The substrate specificity of PPDC$_{Ab}$ was further tested for some commercially available aliphatic and branched-chain 2-keto acids. PPDC$_{Ab}$ exhibited no significant activity towards 3-methyl-2-oxopentanoic acid and 4-methyl-2-oxopentanoic acid (maximum substrate concentration, 4 mM). The aliphatic 2-keto acid 2-ketohexanoic acid was a substrate for PPDC$_{Ab}$ with $k_{cat}/K_{0.5} = 3$ mM^{-1} min^{-1}. 2-Keto acids with longer aliphatic chains (e.g., 2-keto-octanoic acid at 5 mM) were not converted.

Phylogenetic analysis. Sequences similar to PPDC$_{Ab}$ and IPDC$_{Ec}$ were extracted from the NCBI database using the amino acid sequence of these two well-characterized ThDP-dependent 2-keto acid decarboxylases as baits in a BLAST search (1). The multiple sequence alignment, generated by ClustalW, shows that most catalytically important residues are conserved among the 33 sequences, with representatives in

TABLE 2. Steady-state kinetic parameters of *A. brasilense* PPDC

Substrate[d]	h	K_m[a] (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})	Relative activity (%)[b]
Phenylpyruvate	1	1.08 ± 0.09	333.7 ± 7.8	309 ± 27	100.0
IPyA	1.85	0.13 ± 0.01	4.1 ± 0.1	32 ± 3	10.4
4-Phenyl-2-oxobutanoic acid	1.79	1.91 ± 0.11	12.5 ± 0.4	7 ± 0	2.3
5-Phenyl-2-oxopentanoic acid	1.93	0.30 ± 0.01	101.8 ± 0.8	339 ± 12	110.0
2-Ketohexanoic acid[c]	2.23	5.51 ± 0.23	17.2 ± 0.6	3 ± 0	1.0

[a] The parameter K_m is replaced by $K_{0.5}$ when using the Hill equation.
[b] Based on $k_{cat}/K_{0.5}$ in comparison with k_{cat}/K_m of phenylpyruvate.
[c] The enzyme activity of PPDC$_{Ab}$ with 2-ketohexanoic acid as substrate was subjected to substrate inhibition. The data were analyzed for lower substrate concentrations using the Hill model without substrate activation.
[d] The aromatic substrates, which have no activity towards PPDC$_{Ab}$, are benzoylformate and indole-glyoxylic acid. Most aliphatic 2-keto acids are also inactive, including the tested molecules 2-keto-octanoic acid, 3-methyl-2-ketopentanoic acid, and 4-methyl-2-ketopentanoic acid.
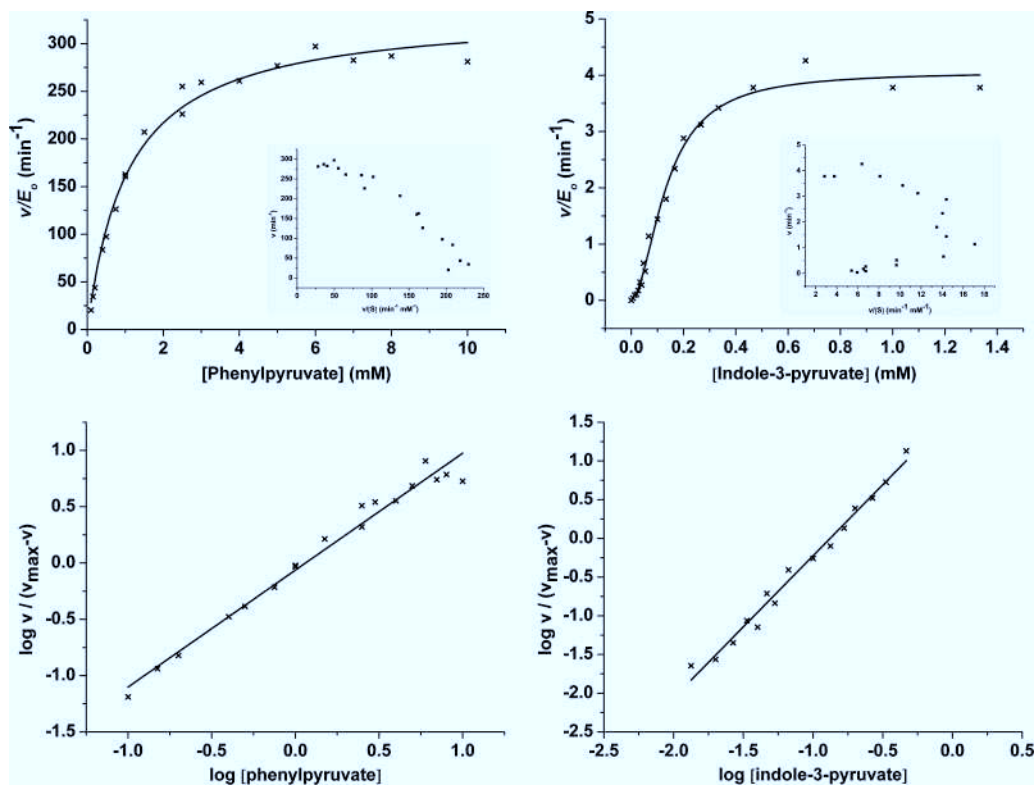
FIG. 2. Determination of the steady-state kinetic parameters of PPDC$_{Ab}$. (Top panels) Plots of initial reaction rates versus substrate concentrations. (Insets) Eadie-Hofstee plots. (Bottom panels) Hill plots of kinetic data. (Left panels) PPDC$_{Ab}$ activity with phenylpyruvate. (Right panels) PPDC$_{Ab}$ activity with IPyA. The reaction rate was determined as described in Materials and Methods. Data are shown in Table 2.

different subclasses of the proteobacteria, firmicutes, and cyanobacteria. In Fig. 3, the multiple sequence alignment of the region comprising the well-conserved ThDP-binding motif (GDGX$_{24}$NN [14]) is shown. However, the only nonconserved but catalytically important residue is glutamate 468 (numbering according to amino acid sequence of IPDC$_{Ec}$), which is exchanged by leucine in nine sequences (indicated with an arrow in Fig. 3). Based on the multiple sequence alignment, a phylogenetic tree was constructed (Fig. 4). The sequences fall into two clades: one contains IPDC$_{Ec}$, while the second contains PPDC$_{Ab}$. Eight sequences clustering with PPDC$_{Ab}$ all have Leu in the position corresponding to Glu468 (sequences indicated with bracket designated "PPDC"). However, the other sequences in this clade (middle bracket) have the conserved Glu in position 468.

DISCUSSION

The *ipdC* gene and gene product from *A. brasilense* have been characterized in different ways during the past several years. The amount of IAA produced by a mutant with a knockout mutation in the *ipdC* gene (FAJ0009) is less than 10% of that produced by the wild-type strain, indicating a central role of this gene in IAA biosynthesis (7, 28). Comparative in vitro growth tests with wheat seedlings inoculated with wild-type *A. brasilense* and FAJ0009 demonstrated a direct link between IAA production and root hair induction (10). The transcription of the *ipdC* gene is under tight control via a positive

feedback mechanism. The IAA production and *ipdC* gene induction are cell density regulated with a maximum in the stationary and late-exponential phases, respectively, and the end product IAA itself is responsible for the induction of the gene expression (37, 38). As this autoinduction is rather unusual for biosynthetic genes, the mechanism(s) for the *ipdC* gene regulation remains intriguing. A signaling role has been suggested for IAA both in bacterial and in bacterium-plant interactions (19, 29, 35). Recently, the production of PAA, a weak auxin with antimicrobial activity, was described by Somers et al. (34). The *ipdC* gene is also involved in the synthesis of PAA, as FAJ0009 produces significantly lower amounts of PAA. Furthermore, PAA can induce *ipdC* gene expression in a concentration range similar to that of IAA (34). Most results regarding PAA can be explained by the structural similarity between IAA and PAA (12). However, it is intriguing that, despite this similarity, IPDC$_{Ec}$ does not show any activity towards phenylpyruvate (30). Therefore, it was reasoned that a biochemical and phylogenetic analysis of the *ipdC* gene product of *A. brasilense* could reveal its distinction from IPDC$_{Ec}$.

Inactivation of the *ipdC* gene has no effect on growth in medium with ammonium as a nitrogen source (Fig. 1), indicating that its activity is not required in the central ammonium metabolism of *A. brasilense*. Nevertheless, when growth of the wild type and that of the mutant were compared in minimal medium with phenylalanine as the sole nitrogen source, FAJ0009 was slightly impaired in growth compared to the wild-type strain. However, a similar growth experiment could

```
A   :  ILVGDGAFQMTGWELGNCRRYGWDPIVIVFNNSSWEMLRTFQ
Rf  :  ILVGDGAFQMTGWELGNCRRYGWDPIVLFNNASWEMLRTFQ
Ab  :  TVVGDGAFQMTGWELGNCRRLGIDPIVILFNNASWEMLRTFQ
Al  :  TVVGDGAFQMTGWELGNCRRLGIDPIVILFNNASWEMLRTFQ
Dd  :  ILVGDGAFQMTGMELGNCTRLGIDPIVLVFNNSSWEMLHVFQ
Rr  :  ILVGDGAFQMTGWELGNCARYGWDPIVIVFNNASWEMLRTFQ
B   :  ILVGDGAFQMTGWELGNCRRYGLDPIVVLFNNCSWEMLRVFQ
Rp  :  VLVGDGAFQMTGWELGNCKRYGLDPIVVLFNNCSWEMLRVFQ
Av  :  ALVGDGAFHMTGMELLTAQRLRLNPIVIVINNGSFASLQAMG
Bm  :  SIIGDGAFQMTGMELSTIVRHGYDPVIIVLDNHCYGTERWLH
Gv  :  VLVGDGAFQMSGLELSTIARFGLDPIVVLLNNRGYGTERPMQ
Pca :  VLVGDGAFQMTGMELSTAARFGLSPIVVVLNNNGYGTERPMT
Mc  :  ILVGDGAFQMTGTELSTHARLGLAPIVVVLNNDNRGYSTERFIL
Hc  :  VVAGDGGFMMMCQSLSTLARNRVNAVIFVMSNGVYAIEQAFV
Ret :  VFAGDGGFQMTAQCLSTQTRFNLNPIIFVMDNGIYGVEQWLA
Sm  :  VISGDGGFQMTAQCLSTQTRFNLNPIIFVTDNGVYGVEQWLA
Lp  :  ACIGDGSFQMTAQELSTMIRYGYKPIIFLMNNACYTIEVQIH
Bc  :  LIIGDGSFQLTVQELSTILRKNLKPIIFLINNNGYTVERAIH
Ca  :  LLIGDGAFQMTAQELSTMLRLQIKPIIFLINNDGYTVERAIH
Np  :  LLVGDGSFQMTAQELSTILRYNLKPIIFLINNDGYTIERDIH
Bx  :  LFIGDGSFQLTAQELSSILRRDLKPVIFLLNNSGYTIERLIL
Reu :  LFIGDGSFQFTAQELSSILRRGLKPVIFLLNNGGYTIERMIL
Eca :  LLIGDGSLQLTVQELGSIIRDRLNMVIVVLNNEGYTVERAIH
Ym  :  LLVGDGAAQLTIQELGSMLRDGLNFTIFLLNNQGYTVERAIH
Ecl :  VITGDGAAQLTIQELGSMLRDKQHPILVLLNNEGYTVERAIH
Ppu :  VITGDGAAQLTIQELGSMLRDKQHPILVLLNNEGYTVERAIH
St  :  IIIGDGAAQLTIQEMGSMLRDGQAPVILLLNNDGYTVERAIH
Pag :  LLIGDGSAQLTIQELGSMLRDGQRLIIFLLNNDGYTVERAIH
Mt  :  LLIGDGAAQLTVQELGTFSREGLSPVIVVVNNDGYTVERAIH
Par :  LLIGDGSALLTIQEIAVMIQERINPVIVLINNDGYTVERAIH
Ll  :  LFIGDGSLQLTVQELGLSIREKLNPICFIINNDGYTVEREIH
Sa  :  LLIGDGSLQLTVQAISTMIRQHIKPVLFVINNDGYTVERLIH
Sv  :  LLSGDGSFQLTVQVSTMIRQKLNTVLFVVNNDGYTIERLIH
```
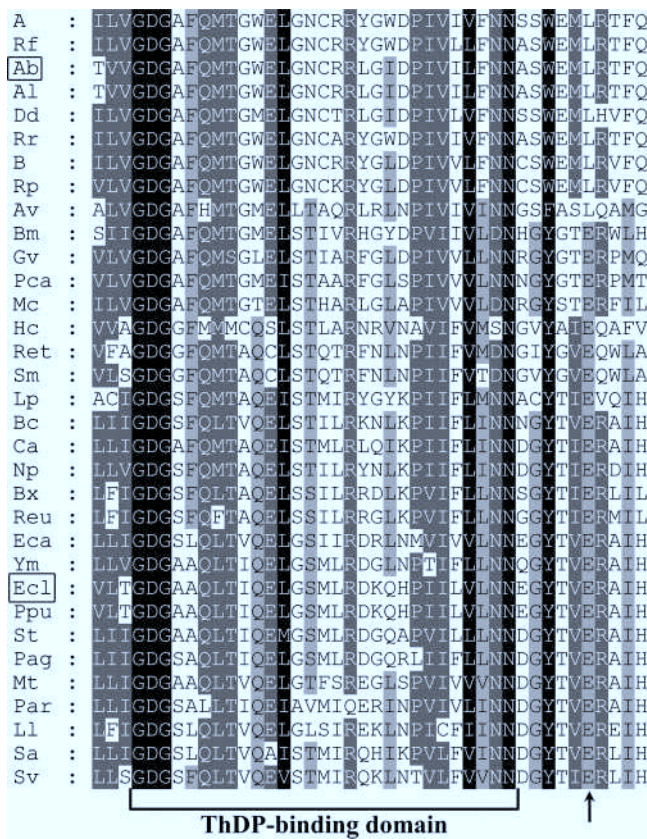
ThDP-binding domain

FIG. 3. Alignment of the amino acid sequences of the ThDP-binding domain. Details of the alignment of a set of (putative) PPDCs and IPDCs are shown. The ThDP-binding domain starts with GDG and ends with NN as defined in reference 14. The conserved glutamine residue, which is replaced by a leucine in (putative) PPDCs, is indicated with an arrow. Abbreviations: A, *Azoarcus* sp. strain EbN1; Ab, *Azospirillum brasilense* Sp245; Al, *Azospirillum lipoferum* FS; Av, *Anabaena variabilis* ATCC 29413; B, *Bradyrhizobium* sp. strain BTAi1; Bc, *Bacillus cereus* E33L; Bm, *Blastopirellula marina* DSM 3645; Bx, *Burkholderia xenovorans* LB400; Ca, *Clostridium acetobutylicum* ATCC 824; Dd, *Desulfovibrio desulfuricans* G20; Eca, *Erwinia carotovora* subsp. *atroseptica* SCRI1043; Ecl, *Enterobacter cloacae*; Gv, *Gloeobacter violaceus* PCC 7421; Hc, *Hahella chejuensis* KCTC 2396; Ll, *Lactococcus lactis* subsp. *lactis* IFPL730; Lp, *Legionella pneumophila* strain Lens; Mc, *Methylococcus capsulatus* strain Bath; Mt, *Mycobacterium tuberculosis* H37Rv; Np, *Nostoc punctiforme* PCC 73102; Pag, *Pantoea agglomerans*; Par, *Psychrobacter arcticus* 273-4; Pca, *Pelobacter carbinolicus* DSM 2380; Ppu, *Pseudomonas putida*; Ret, *Rhizobium etli* CFN 42; Reu, *Ralstonia eutropha* H16; Rf, *Rhodoferax ferrireducens* T118; Rp, *Rhodopseudomonas palustris* BisB5; Rr, *Rhodospirillum rubrum* ATCC 11170; Sa, *Staphylococcus aureus* RF122; Sm, *Sinorhizobium medicae* WSM419; St, *Salmonella enterica* serovar Typhimurium LT2; Sv, *Sarcina ventriculi*; Ym, *Yersinia mollaretii* ATCC 43969. Species names boxed in the figure contain IPDCs/PPDCs discussed in the text.

not be conducted with tryptophan as neither the wild type nor the mutant can grow with tryptophan as the sole ammonium source. Neither phenylalanine nor tryptophan can be used as a sole carbon/energy source by *A. brasilense*. This difference in growth could be partially linked to a difference in uptake of the amino acids. As *A. brasilense* cannot grow on minimal medium with phenylalanine as a carbon source, the generation of compounds necessary for energy formation is insufficient probably due to either a low uptake of phenylalanine, slow conversion by different enzymes (e.g., PPDC$_{Ab}$), or the absence of enzymes necessary for the degradation of intermediates to coenzyme A derivatives. The small but significant effect of the *ipdC* mutation on growth with phenylalanine as nitrogen source could point to a partial role of the *ipdC* gene product in the Ehrlich pathway which involves transamination followed by decarboxylation. In *Saccharomyces cerevisiae* the Ehrlich pathway has been studied extensively for the production of fusel oils. Transcriptome analysis of yeast cells grown in glucose-limited chemostats with phenylalanine as the sole nitrogen source in comparison with cells grown in ammonium revealed that the transcript level of *ARO10/YDR380w* was increased 30-fold (42). Aro10p was first identified as a PPDC in *S. cerevisiae* (9, 42); later a broader substrate specificity was proposed (41). The yeast enzyme Aro10p itself has not yet been biochemically characterized.

In a second part of this study, PPDC$_{Ab}$ was heterologously overexpressed in *E. coli* and subsequently purified via immobilized nickel chelate affinity chromatography. The recombinant protein appeared as a tetrameric form in size-exclusion chromatography, which is consistent with other ThDP-dependent 2-keto acid decarboxylases (2, 11, 13) and the recently determined crystal structure of PPDC$_{Ab}$, which shows the enzyme as a dimer of dimers (40). The kinetic analysis revealed quite surprising results. PPDC$_{Ab}$ shows near Michaelis-Menten kinetics toward phenylpyruvate ($h = 1.26$) with the highest specificity constant of all tested substrates (except for 5-phenyl-2-oxopentanoic acid). For all other tested substrates with significant activity, the enzyme is subject to clear substrate activation ($h = 1.8$ to 1.9). To determine the kinetic parameters of these substrates, the kinetic data were analyzed with the Hill model. The structurally similar substrate IPyA has a 100-fold-lower k_{cat} than does phenylpyruvate. Due to a low $K_{0.5}$, the $k_{cat}/K_{0.5}$ value with IPyA is 10-fold lower. To broaden our understanding of the substrate specificity, phenyl-substituted 2-keto acids with a longer alkyl side chain were tested. Surprisingly, 4-phenyl-2-oxobutanoic acid has a 100-fold-lower $k_{cat}/K_{0.5}$ value, while 5-phenyl-2-oxopentanoic acid (with a side chain elongated by one methylene group) has a $k_{cat}/K_{0.5}$ value comparable to the k_{cat}/K_m value for phenylpyruvate. The reason for the difference between these two compounds could be the position of the keto acid group with respect to the phenyl moiety. Crystal structures of the enzyme cocrystallized with substrate (analogues) should shed further light on this. Benzoylformate (phenylglyoxylic acid), which has a shorter side chain than does phenylpyruvate, and indole-glyoxylic acid are not converted. The substrate activity profile of PPDC$_{Ab}$ is clearly distinct from that of IPDC$_{Ec}$. IPDC$_{Ec}$ shows activity towards IPyA, benzoylformate, and 4-substituted derivatives with no indication for substrate activation behavior by any of these compounds. Phenylpyruvate is not a substrate for IPDC$_{Ec}$ (30). These differences in the substrate range of the *A. brasilense* protein and IPDC$_{Ec}$ indicate an evolutionary difference and would allow the functional classification of the *A. brasilense* decarboxylase as a PPDC. In general the catalytic activity of PPDC$_{Ab}$ is rather low compared to that of 2-keto acid decarboxylases but comparable to the catalytic activity of IPDC$_{Ec}$. As both enzymes are part of a pathway for the syn-
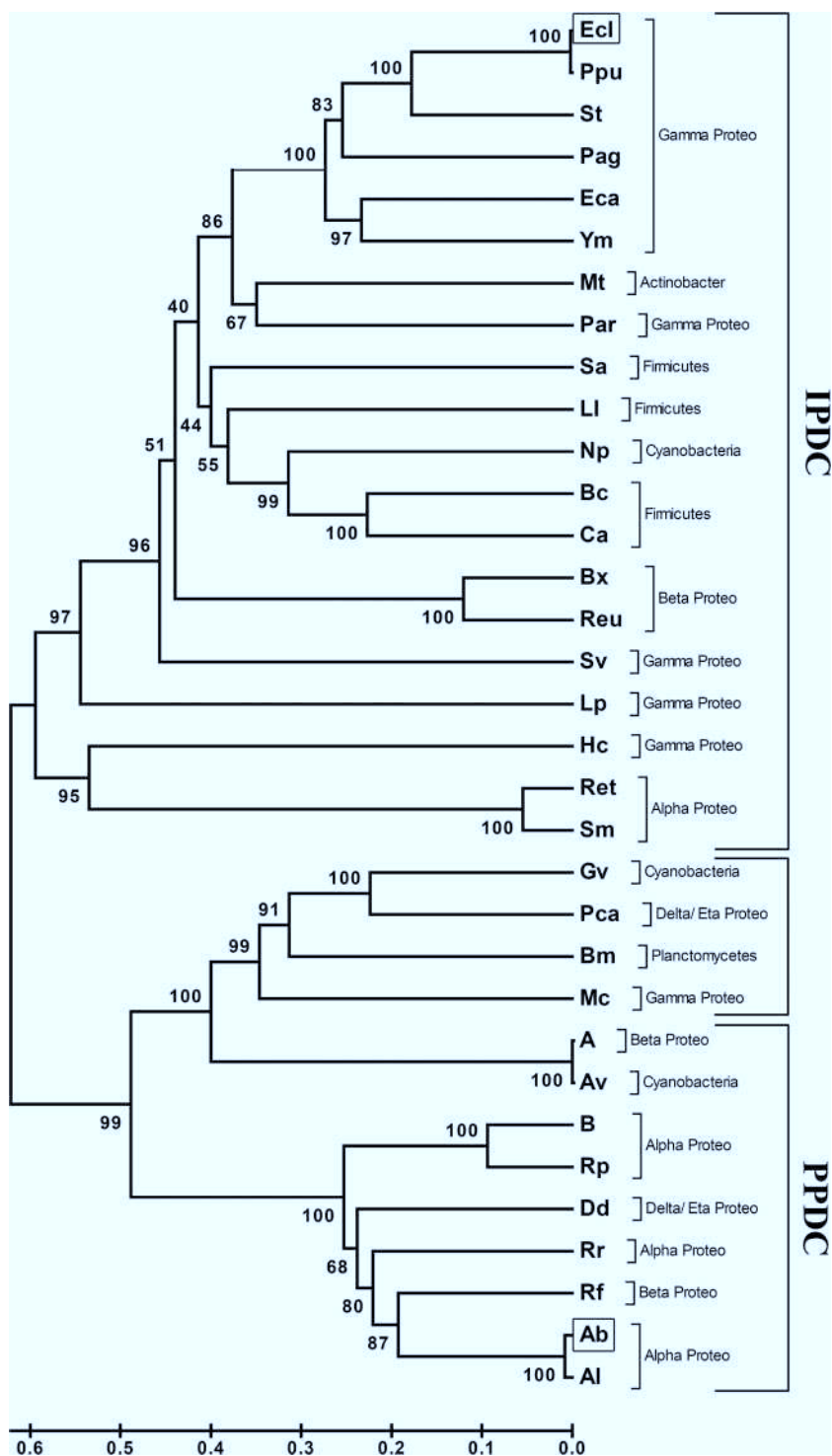
FIG. 4. Phylogenetic tree of (putative) PPDCs and IPDCs. ClustalW was used to align a set of (putative) PPDC and IPDC amino acid sequences. A phylogenetic tree was constructed based on the UPGMA method. The values shown at the branches are bootstrap values for 1,000 replicates. For species abbreviations, see the legend to Fig. 3. Boxed abbreviations indicate IPDCs/PPDCs discussed in the text.

thesis of the plant hormone IAA, to be considered as a secondary metabolite in microorganisms, their activity does not have to be as high as that of other enzymes, like pyruvate decarboxylases, which are involved in an important metabolic pathway (central intermediate metabolism).

Several ThDP-dependent 2-keto acid decarboxylases such as pyruvate decarboxylase from *Saccharomyces cerevisiae* (ScPDC) and from *Kluyveromyces lactis* are activated by their substrate (15, 17, 33). Most bacterial pyruvate decarboxylases, IPDC$_{Ec}$, and benzoylformate decarboxylase display hyperbolic

Michaelis-Menten kinetics without any indication of substrate activation (11, 27, 30). The combination of clear substrate activation for certain substrates and hyperbolic Michaelis-Menten kinetics (or only very weak substrate activation) for others in a single enzyme, as observed for the PPDC$_{Ab}$, is less common. The nature of this behavior is currently under investigation. The mechanism of substrate activation in ScPDC has in fact been the subject of discussion for quite a few years. In the most widely accepted mechanism, binding of the substrate to a cysteine residue (Cys221 in ScPDC) distant from the active site is proposed to be the first step in an information transfer cascade to the active site, leading ultimately to increased activity (3, 4). However, in a crystal structure cocrystallized with the substrate surrogate pyruvamide, the activator was found to bind to a site 10 Å away from the pivotal cysteine. Binding of the activator to ScPDC induces a conformational change in the enzyme: two (of the four) active sites transform from an open to a closed (=active) form, which might be the molecular explanation for the substrate activation behavior (21, 22). Since the pivotal cysteine in the proposed activation mechanism of ScPDC is not conserved in PPDC$_{Ab}$ (40), it will be interesting to unravel the activation mechanism used by the latter.

Numerous sequences annotated as putative IPDCs and PPDCs are present in the NCBI database, but only for two has the activity of the purified enzymes been determined (PPDC$_{Ab}$ [this study] and IPDC$_{Ec}$). To possibly distinguish between these two different kinds of enzyme specificities by in silico analysis, a multiple sequence alignment was performed and a phylogenetic tree was constructed. Strikingly, with the exception of one amino acid residue, all catalytically important residues are conserved in the aligned sequences, indicating the same catalytic mechanism suggested for other 2-keto acid decarboxylases. The multiple sequence alignment allows the division of the sequences into two groups based on the amino acid corresponding to position 468 in IPDC$_{Ec}$ (477 in ScPDC numbering): those with Glu and those with Leu. Glu468 is a conserved catalytic residue in pyruvate decarboxylase and IPDC$_{Ec}$, which has been proposed to be involved in both pre- and postdecarboxylation steps. More specifically a role in protonation of the enamine-carbanion intermediate has been proposed (20, 26, 32). Enzymes with amino acid substitutions in this position showed an impaired release of the aldehyde product. In the benzoylformate decarboxylase of *Pseudomonas putida*, the glutamate residue is also replaced by Leu. Here, protonation of the enamine-carbanion intermediate is proposed to be mediated by a catalytic histidine residue in the active site (27). In PPDC$_{Ab}$, it was proposed that the leucine, which is located underneath the carboxylate group of the lactyl-ThDP intermediate, increases the hydrophobic character of the active site and consequently stabilizes the zwitterionic intermediates (40). In view of this clear partitioning, site-directed mutagenesis studies at this and other amino acid residue positions (see below) are ongoing, to provide information on their role in catalysis and substrate activation.

The phylogenetic tree based on the complete sequences (Fig. 4) also consists of two clades, indicating that the sequences have evolved from at least two different ancestors. For IPDC$_{Ec}$ and PPDC$_{Ab}$, this evolutionary difference has been demonstrated via biochemical analysis: PPDC$_{Ab}$ converts both

phenylpyruvate and IPyA, while IPDC$_{Ec}$ shows no activity for phenylpyruvate. Based on the phylogenetic tree, it is tempting to speculate that all members indicated by the bracket "PPDC" can be classified as PPDCs as they cluster with PPDC$_{Ab}$ and have a conserved Leu at position 468. However, it should be kept in mind that the biochemical activity of the purified enzyme has been determined for only one representative. The middle bracket of the phylogenetic tree, grouping four sequences, is part of the same clade, but these sequences have a conserved glutamate residue at position 468, in contrast to the other members of the clade. This implies that other amino acid residues are probably involved in determining the substrate specificity, unless a yet-unknown catalytic activity, different from that with IPyA or phenylpyruvate as a substrate, is represented by these four sequences.

Taking previous data and data presented in this study together, we propose the reclassification of the *ipdC* gene product of *A. brasilense* as a PPDC (EC 4.1.1.43).

ACKNOWLEDGMENTS

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:** 3389–3402.
2. **Arjunan, P., T. Umland, F. Dyda, S. Swaminathan, W. Furey, M. Sax, B. Farrenkopf, Y. Gao, D. Zhang, and F. Jordan.** 1996. Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* at 2.3 angstrom resolution. J. Mol. Biol. **256:** 590–600.
3. **Baburina, I., Y. Gao, Z. Hu, F. Jordan, S. Hohmann, and W. Furey.** 1994. Substrate activation of brewers' yeast pyruvate decarboxylase is abolished by mutation of cysteine 221 to serine. Biochemistry **33:**5630–5635.
4. **Baburina, I., H. J. Li, B. Bennion, W. Furey, and F. Jordan.** 1998. Interdomain information transfer during substrate activation of yeast pyruvate decarboxylase: the interaction between cysteine 221 and histidine 92. Biochemistry **37:**1235–1244.
5. **Baldani, V. L. D., M. A. D. Alvarez, J. I. Baldani, and J. Döbereiner.** 1986. Establishment of inoculated *Azospirillum* spp in the rhizosphere and in roots of field-grown wheat and sorghum. Plant Soil **90:**35–46.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
7. **Costacurta, A., V. Keijers, and J. Vanderleyden.** 1994. Molecular cloning and sequence analysis of an *Azospirillum brasilense* indole-3-pyruvate decarboxylase gene. Mol. Gen. Genet. **243:**463–472.
8. **Costacurta, A., and J. Vanderleyden.** 1995. Synthesis of phytohormones by plant-associated bacteria. Crit. Rev. Microbiol. **21:**1–18.
9. **Dickinson, J. R., L. E. Salgado, and M. J. Hewlins.** 2003. The catabolism of amino acids to long chain and complex alcohols in *Saccharomyces cerevisiae*. J. Biol. Chem. **278:**8028–8034.
10. **Dobbelaere, S., A. Croonenborghs, A. Thys, A. Vande Broek, and J. Vanderleyden.** 1999. Phytostimulatory effect of *Azospirillum brasilense* wild type and mutant strains altered in IAA production on wheat. Plant Soil **212:**155–164.
11. **Dobritzsch, D., S. König, G. Schneider, and G. Lu.** 1998. High resolution crystal structure of pyruvate decarboxylase from *Zymomonas mobilis*. Implications for substrate activation in pyruvate decarboxylases. J. Biol. Chem. **273:**20196–20204.
12. **Ferro, N., A. Gallegos, P. Bultinck, H. J. Jacobsen, R. Carbo-Dorca, and T. Reinard.** 2006. Coulomb and overlap self-similarities: a comparative selectivity analysis of structure-function relationships for auxin-like molecules. J. Chem. Inf. Model. **46:**1751–1762.
13. **Hasson, M. S., A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A. Gerlt, G. L. Kenyon, G. A. Petsko, and D. Ringe.** 1998. The crystal structure of benzoylformate decarboxylase at 1.6 angstrom resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. Biochemistry **37:** 9918–9930.
14. **Hawkins, C. F., A. Borges, and R. N. Perham.** 1989. A common structural motif in thiamin pyrophosphate-binding enzymes. FEBS Lett. **255:**77–82.

15. **Hübner, G., R. Weidhase, and A. Schellenberger.** 1978. The mechanism of substrate activation of pyruvate decarboxylase: a first approach. Eur. J. Biochem. **92:**175–181.

16. **Koga, J., T. Adachi, and H. Hidaka.** 1991. Molecular cloning of the gene for indolepyruvate decarboxylase from *Enterobacter cloacae*. Mol. Gen. Genet. **226:**10–16.

17. **Krieger, F., M. Spinka, R. Golbik, G. Hübner, and S. König.** 2002. Pyruvate decarboxylase from *Kluyveromyces lactis*. An enzyme with an extraordinary substrate activation behaviour. Eur. J. Biochem. **269:**3256–3263.

18. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. **5:**150–163.

19. **Lambrecht, M., Y. Okon, A. Vande Broek, and J. Vanderleyden.** 2000. Indole-3-acetic acid: a reciprocal signalling molecule in bacteria-plant interactions. Trends Microbiol. **8:**298–300.

20. **Lobell, M., and D. H. G. Crout.** 1996. Pyruvate decarboxylase: a molecular modeling study of pyruvate decarboxylation and acyloin formation. J. Am. Chem. Soc. **118:**1867–1873.

21. **Lu, G. G., D. Dobritzsch, S. Baumann, G. Schneider, and S. König.** 2000. The structural basis of substrate activation in yeast pyruvate decarboxylase. A crystallographic and kinetic study. Eur. J. Biochem. **267:**861–868.

22. **Lu, G. G., D. Dobritzsch, S. König, and G. Schneider.** 1997. Novel tetramer assembly of pyruvate decarboxylase from brewer's yeast observed in a new crystal form. FEBS Lett. **403:**249–253.

23. **Nicholas, K. B., H. B. Nicholas, and D. W. Deerfield II.** 1997. GeneDoc: analysis and visualization of genetic variation. EMBNET.news **4:**1–4.

24. **Okon, Y., and C. A. Labandera-Gonzalez.** 1994. Agronomic applications of *Azospirillum*: an evaluation of 20 years worldwide field inoculation. Soil Biol. Biochem. **26:**1591–1601.

25. **Patten, C. L., and B. R. Glick.** 1996. Bacterial biosynthesis of indole-3-acetic acid. Can. J. Microbiol. **42:**207–220.

26. **Pohl, M., P. Siegert, K. Mesch, H. Bruhn, and J. Grotzinger.** 1998. Active site mutants of pyruvate decarboxylase from *Zymomonas mobilis*—a site-directed mutagenesis study of L112, I472, I476, E473 and N482. Eur. J. Biochem. **257:**538–546.

27. **Polovnikova, E. S., M. J. McLeish, E. A. Sergienko, J. T. Burgner, N. L. Anderson, A. K. Bera, F. Jordan, G. L. Kenyon, and M. S. Hasson.** 2003. Structural and kinetic analysis of catalysis by a thiamin diphosphate-dependent enzyme, benzoylformate decarboxylase. Biochemistry **42:**1820–1830.

28. **Prinsen, E., A. Costacurta, K. Michiels, J. Vanderleyden, and H. Van Onckelen.** 1993. *Azospirillum brasilense* indole-3-acetic acid biosynthesis: evidence for a non-tryptophan dependent pathway. Mol. Plant-Microbe Interact. **6:**609–615.

29. **Remans, R., S. Spaepen, and J. Vanderleyden.** 2006. Auxin signaling in plant defense. Science **313:**171.

30. **Schütz, A., R. Golbik, K. Tittmann, D. I. Svergun, M. H. Koch, G. Hübner, and S. König.** 2003. Studies on structure-function relationships of indolepyruvate decarboxylase from *Enterobacter cloacae*, a key enzyme of the indole acetic acid pathway. Eur. J. Biochem. **270:**2322–2331.

31. **Schütz, A., T. Sandalova, S. Ricagno, G. Hübner, S. König, and G. Schneider.** 2003. Crystal structure of thiamin diphosphate-dependent indolepyruvate decarboxylase from *Enterobacter cloacae*, an enzyme involved in the biosynthesis of the plant hormone indole-3-acetic acid. Eur. J. Biochem. **270:**2312–2321.

32. **Sergienko, E. A., and F. Jordan.** 2001. Catalytic acid-base groups in yeast pyruvate decarboxylase. 2. Insights into the specific roles of D28 and E477 from the rates and stereospecificity of formation of carboligase side products. Biochemistry **40:**7369–7381.

33. **Sergienko, E. A., and F. Jordan.** 2002. New model for activation of yeast pyruvate decarboxylase by substrate consistent with the alternating sites mechanism: demonstration of the existence of two active forms of the enzyme. Biochemistry **41:**3952–3967.

34. **Somers, E., D. Ptacek, P. Gysegom, M. Srinivasan, and J. Vanderleyden.** 2005. *Azospirillum brasilense* produces the auxin-like phenylacetic acid by using the key enzyme for indole-3-acetic acid biosynthesis. Appl. Environ. Microbiol. **71:**1803–1810.

35. **Spaepen, S., J. Vanderleyden, and R. Remans.** 2007. Indole-3-acetic acid in microbial and microorganism-plant signaling. FEMS Microbiol. Rev. **31:**425–448.

36. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. Clustal-W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.

37. **Vande Broek, A., P. Gysegom, O. Ona, N. Hendrickx, E. Prinsen, J. Van Impe, and J. Vanderleyden.** 2005. Transcriptional analysis of the *Azospirillum brasilense* indole-3-pyruvate decarboxylase gene and identification of a cis-acting sequence involved in auxin responsive expression. Mol. Plant-Microbe Interact. **18:**311–323.

38. **Vande Broek, A., M. Lambrecht, K. Eggermont, and J. Vanderleyden.** 1999. Auxins upregulate expression of the indole-3-pyruvate decarboxylase gene in *Azospirillum brasilense*. J. Bacteriol. **181:**1338–1342.

39. **Vanstockem, M., K. Michiels, J. Vanderleyden, and A. P. Van Gool.** 1987. Transposon mutagenesis of *Azospirillum brasilense* and *Azospirillum lipoferum*: physical analysis of Tn*5* and Tn*5*-Mob insertion mutants. Appl. Environ. Microbiol. **53:**410–415.

40. **Versées, W., S. Spaepen, J. Vanderleyden, and J. Steyaert.** 2007. The crystal structure of phenylpyruvate decarboxylase from *Azospirillum brasilense* at 1.5 angstrom resolution: implications for its catalytic and regulatory mechanism. FEBS J. **274:**2363–2375.

41. **Vuralhan, Z., M. A. Luttik, S. L. Tai, V. M. Boer, M. A. Morais, D. Schipper, M. J. Almering, P. Kotter, J. R. Dickinson, J. M. Daran, and J. T. Pronk.** 2005. Physiological characterization of the *ARO10*-dependent, broad-substrate-specificity 2-oxo acid decarboxylase activity of *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. **71:**3276–3284.

42. **Vuralhan, Z., M. A. Morais, S. L. Tai, M. D. Piper, and J. T. Pronk.** 2003. Identification and characterization of phenylpyruvate decarboxylase genes in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. **69:**4534–4541.

43. **Weiss, P. M., G. A. Garcia, G. L. Kenyon, W. W. Cleland, and P. F. Cook.** 1988. Kinetics and mechanism of benzoylformate decarboxylase using ^{13}C and solvent deuterium isotope effects on benzoylformate and benzoylformate analogues. Biochemistry **27:**2197–2205.

Publication IV

Pyruvate decarboxylase from *Acetobacter pasteurianus*: biochemical and structural characterisation.

Gocke, D., Berthold, C. L., Graf, T., Brosi, H., Frindi-Wosch, I.,
Knoll, M., Stillger, T., Walter, L., Müller, M., Pleiss, J.,
Schneider, G. & Pohl, M.

Pyruvate Decarboxylase from *Acetobacter pasteurianus*: BIOCHEMICAL AND STRUCTURAL CHARACTERISATION

Dörte Gocke[1], Catrine L. Berthold[2], Thorsten Graf[1], Helen Brosi[1], Ilona Frindi-Wosch[1], Michael Knoll[3], Thomas Stillger[4], Lydia Walter[4], Michael Müller[4], Jürgen Pleiss[3], Gunter Schneider[2], Martina Pohl[1]*

[1]*Institute of Molecular Enzyme Technology, Heinrich-Heine University Düsseldorf, Research Centre Jülich, D-52426 Jülich, Germany;* [2]*Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden;* [3]*Institute of Technical Biochemistry, University of Stuttgart, D-70569 Stuttgart, Germany;* [4]*Institute of Pharmaceutical Sciences, Albert-Ludwigs-University Freiburg, 79104 Freiburg, Germany*

*Corresponding author.
Tel.: +49-(0)2461-613704; Fax: +49-(0)2461-612490; E-mail: ma.pohl@fz-juelich.de

***Keywords*:** *Acetobacter pasteurianus;* pyruvate decarboxylase; enzymatic carboligation; thiamin diphosphate; X-ray structure; 2-hydroxyketones

Summary
Pyruvate decarboxylase from *Acetobacter pasteurianus* (*Ap*PDC), a thiamin diphosphate (ThDP)-dependent enzyme, has been investigated concerning its decarboxylase activity as well as the potential to catalyse the carboligation of aldehydes to chiral 2-hydroxyketones. Beside a detailed characterisation of the substrate range, the stability of *Ap*PDC was studied with regard to the influences of different pH-values and temperatures. Data were compared to those obtained with similar PDCs from *Zymomonas mobilis*, *Zymobacter palmae* and *Saccharomyces cerevisiae*. Although *Ap*PDC resembles to a large extent the well characterised enzyme from *Zymomonas mobilis* with respect to sequence, kinetic behaviour and stability, there are some remarkable differences concerning the decarboxylation of aromatic 2-ketoacids and the stereoselectivity of the carboligase reaction, which hint to a larger active site in *Ap*PDC. These results can be explained based on the X-ray structure of the enzyme, which was determined to a resolution of 2.75 Å.

1. Introduction
Pyruvate decarboxylases (PDC, E.C. 4.1.1.1) catalyse the non-oxidative decarboxylation of pyruvate to acetaldehyde requiring the cofactors thiamin diphosphate (ThDP) and divalent cations like Mg^{2+}. They are the key enzymes in homo-fermentative pathways cleaving the central intermediate pyruvate into acetaldehyde and CO_2. PDCs are commonly found in plants, yeasts and fungi, but absent in mammals [1]. Currently four PDCs from prokaryotes have been cloned and characterised, encompassing three PDCs from the gram-negative bacteria *Zymomonas mobilis* (*Zm*PDC) [2], *Zymobacter palmae* (*Zp*PDC) [3], and *Acetobacter pasteurianus* (*Ap*PDC) [4], as well as the PDC from the gram-positive bacterium *Sarcina ventriculi* (*Sv*PDC) [5].
A. pasteurianus, a common spoilage organism of fermented juices, is an obligate oxidative bacterium and in contrast to the fermentative ethanol pathways in plants, yeast and bacterial PDCs, *Ap*PDC is a central enzyme of the oxidative metabolism [4].
Pyruvate decarboxylases are key enzymes in the production of biofuels from bulk plant materials like sugars or cellulose. Currently bacterial strains with engineered ethanol pathways, e.g. *Escherichia coli* strains with an inserted *Zm*PDC, alcohol dehydrogenase (ADH) and acyltransferase are investigated for industrial applications [6]. Moreover, PDCs are able to catalyse an acyloin-condensation like carboligation of aldehydes. Since the 1930s this activity of PDC from yeast is used in an industrial fermentative process for the production of (*R*)-phenylacetylcarbinol (**20**), a pre-step in the synthesis of ephedrine [7, 8]. A similar carboligase activity yielding chiral 2-hydroxy ketones was detected and intensively studied for other PDCs and further ThDP-dependent enzymes like benzoylformate decarboxylase (BFD, E.C. 4.1.1.7), branched-chain 2-keto acid decarboxylase (KdcA, E.C. 4.1.1.72) and benzaldehyde lyase (BAL, E.C. 4.1.2.38) [9-18]. Different substrate preferences of these enzymes open the way to a broad range of 2-hydroxy

ketones. Our intention is to create a 2-hydroxy ketone platform encompassing a high diversity of enantiocomplementary molecules. To achieve this goal a detailed investigation of structure-function relationships of ThDP-dependent enzymes is necessary in order to elucidate the mechanisms regulating chemo- and enantioselectivity and to enable the design of biocatalysts with tailor-made carboligase activity. Using a structure guided approach we have recently opened the way to access (*S*)-hydroxy ketones by site-directed mutagenesis of the predominantly strictly *R*-selective ThDP-dependent enzymes [19, 20].

Here we report the characterisation of the pyruvate decarboxylase from *Acetobacter pasteurianus* (*Ap*PDC) with regard to the substrate range of the decarboxylase, the carboligase reaction and the pH- and temperature dependent stability and activity in comparison to other PDCs from *Zymomonas mobilis* (*Zm*PDC), *Zymobacter palmae* (*Zp*PDC) and *Saccharomyces cerevisiae* (*Sc*PDC). Furthermore the three-dimensional (3D) structure of *Ap*PDC was determined by X-ray crystallography demonstrating a significant larger active site compared to other PDCs of known structure, which explains the observed differences in the substrate range and the stereoselectivity of *Ap*PDC.

2. Materials and methods

2.1. Cloning of ApPDC

Genomic DNA of *Acetobacter pasteurianus* subspec. ascendens (DSM-No. 2347 (= ATCC 12874), DSMZ) was isolated with the DNeasy Tissue Kit (Qiagen) using the manufacturers protocol for gram-negative bacteria (DNeasy Tissue Kit Handbook 05/2002). The gene was amplified by PCR using the following primers:

*Ap*PDC_hisC-up: *Ap*PDC_hisC-down:
5'-ATAT**CATATG**ACCTATACTGTTGGCAT–3'   5'-ATAT**CTCGAG**GGCCAGAGTGGTCTTGC-3'
    ***Nde*I**<u>*start*</u>   ***Xho*I**

After a single digest of the amplified gene with *Xho*I the gene was first ligated into the vector pBluescriptII (Fermentas) in order to allow restriction by *Nde*I and *Xho*I. The 1674 bp long gene was digested with both restriction endonucleases and ligated into the vector pET22b (Novagen) carrying an ampicillin-resistance and the information for a C-terminal His$_6$-tag. Transformation of the expression host *E. coli* BL21(DE3) (Novagen) was performed by electroporation. The DNA sequence was verified by sequencing (Sequiserve) and compared to the published sequence [4] (see supplementary material). *Zp*PDC was cloned into pET28a(+) carrying a C-terminal His$_6$-tag.

2.2. Cultivation and expression

Cultivation, expression, purification and storage of *Zp*PDC and *Sc*PDC were performed as described for *Ap*PDC in the following. High cell density cultivation according to the method of Korz et al. [21] was performed in a 40 L Techfors reactor (Infors AG, CH) at 30°C, pH 7.0. Starting on minimal media the fed-batch process was run by a continuous feed of glucose. pH 7.0 was maintained by NH$_3$-addition. Dissolved oxygen saturation was regulated between 30-40% saturation by increasing the stirrer speed and the air flow rate and by a stepwise pressure increase during the high oxygen demanding expression. After a preliminary growth phase protein production was induced by the addition of 1.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an OD$_{600}$~60. Cells were harvested by a separator (Westfalia Separator AG) and stored at -20°C.

2.3. Purification and storage

Harvested *E. coli* cells were suspended by addition of 1:10 (w/v) buffer A (50 mM potassium phosphate buffer, pH 6.5, including 2.5 mM MgSO$_4$ and 0.1 mM ThDP) and 10 mM. After addition of 0.33 mg/mL lysozyme and incubation for 30 min on ice, cells were further disrupted by ultra-sonication (10 times 30 sec) followed by centrifugation.

Purification of *Ap*PDC for biochemical studies was performed by immobilised nickel chelate chromatography and additional size exclusion chromatography, using a purification protocol previously developed for *Zm*PDC [22] with the following alterations: The pH of all 50 mM potassium phosphate buffers was set to 6.5; all buffers contained 2.5 mM MgSO$_4$ and 0.1 mM ThDP; non-specifically bound proteins were eluted with 50 mM imidazole; *Ap*PDC was eluted

with 250 mM imidazole. The enzyme was either freeze dried or diluted with 50% (v/v) glycerol and stored at -20 °C.

For crystallisation of *Ap*PDC the purification protocol was kept but potassium buffer was exchanged by Mes/NaOH-buffer as following: Ni-NTA-chromatography: disintegration buffer (50 mM Mes/NaOH-buffer, pH 6.5, 2.5 mM $MgSO_4$, 0.1 mM ThDP), washing buffer (50 mM Mes/HCl-buffer, pH 6.5, 50 mM imidazole), elution buffer (50 mM Mes/HCl-buffer, pH 6.5, 250 mM imidazole); G25-chromatography: 20 mM Mes/NaOH-buffer, pH 7.0, 2.5 mM $MgSO_4$, 0.1 mM ThDP. The enzyme solution was concentrated up to 35 mg/mL using vivaspin 20 centrifuge columns (Sartorius, cut off 10 kDa). For storage the enzyme solution was shock-frozen in liquid nitrogen and kept at -20 °C.

2.4. Determination of the molecular mass

Size-exclusion chromatography was performed using a Superdex G200 16/60 prep grade column (Amersham) and buffer A (pH 6.5) including 150 mM KCl. The coefficients of available volume ($K_{av} = (V_e-V_o)/(V_t-V_o)$, with V_e = elution volume of the respective protein and V_t = total volume, elution volume of blue dextran) for *Ap*PDC and the standard proteins were determined twice with a standard deviation of 0.2% (see supplementary material).

Calibration was performed using ribonuclease A (13.7 kD, K_{av} = 0.63), chymotrypsinogen A (25 kD, K_{av} = 0.59), ovalbumin (43 kD; K_{av} = 0.46), BSA (67 kD, K_{av} = 0.37), aldolase (158 kD, K_{av} = 0.27), catalase (232 kD, K_{av} = 0.23), ferritin (440 kD, K_{av} = 0.12) and thyroglobulin (669 kD, K_{av} = 0.05). The lyophilisate (56 mg) was solved in 600 μL buffer A (pH 6.5) including 60 μL 1.5 M KCl; flow 1 mL/min. The K_{av}-value of *Ap*PDC was determined to 0.125. Data were plotted as $logM_r$ over K_{av} resulting in a linear correlation with R^2 = 0.993.

2.5. Decarboxylase activity

One unit of decarboxylase activity is defined as the amount of *Ap*PDC which catalyses the decarboxylation of 1 μmol pyruvate (**1**) per minute under standard conditions (pH 6.5, 30°C). Protein determination was performed according to Bradford [23] using BSA as the standard.

Two continuous decarboxylase assays were used. In the *coupled decarboxylase assay* yeast alcohol dehydrogenase (yeast ADH) reduces the aldehydes obtained after decarboxylation of 2-keto acids by *Ap*PDC. The simultaneous consumption of NADH was followed spectrophotometrically for 90 sec at 340 nm (molar extinction coefficient ε of NADH = 6.22 L mmol^{-1} cm^{-1}). Assay composition: 890 μL buffer A (pH 6.5), 50 μL sodium pyruvate or another 2-keto acid (350 mM) in buffer A (final concentration 17.5 mM), 50 μL NADH 4.6 mM in buffer A (final concentration 0.23 mM), 10 μL yeast ADH (Roche, EC 1.1.1.1, 1 g in 34 mL 3.2 M ammonium sulphate, pH 6.0, 300 U/mg at 25°C). The assay was started by addition of 50 μL *Ap*PDC. The concentration of *Ap*PDC was chosen in such a way that a linear decay of NADH was observed over 90 sec (ca. 0.2 mg/mL *Ap*PDC with pyruvate as the substrate). The concentration of yeast ADH was sufficient to make certain that the coupling reaction was not rate limiting. The 3.2 M ammonium sulphate storage buffer of yeast ADH did not influence the results.

Further, a *direct decarboxylase assay* following the direct decay of pyruvate was developed to measure *Ap*PDC activity under NADH degrading or ADH inactivating conditions. Assay composition: 950 μL sodium pyruvate (15 mM in buffer A, pH 6.5). The reaction was started by addition of 50 μL *Ap*PDC solution (ca 0.8 mg/mL). The decay of pyruvate was followed at 320 nm ($ε_{pyruvate}$ = 0.022 L mmol^{-1} cm^{-1}) in quartz cuvettes.

Kinetic constants were calculated by non-linear regression using the Michaelis-Menten equation in Origin 7G SR4 (OriginLab Coop., Northampton) except for *Sc*PDC, where the Hill equation including substrate inhibition was used (see supplementary material). *Sc*PDC decarboxylase activity was measured using the same assay as described for *Ap*PDC. The activity assay for *Zp*PDC was performed at pH 7.0, which is the optimum for initial rate activity of *Zp*PDC (Fig. 2). Assays with *Zm*PDC were performed in 50 mM Mes/KOH-buffer, pH 6.5, including 2 mM $MgSO_4$ and 0.1 mM ThDP [24].

Both decarboxylase assays show a standard deviation of about 10%. Therefore all decarboxylase activities in this paper are given as the average of minimum 3-6 measurements.

Substrate range of decarboxylase activity- Activity toward different 2-keto acids was measured with the coupled decarboxylase assay using horse liver ADH as coupling enzyme. It was verified

that the respective aldehydes, which are obtained by decarboxylation, are substrates for the horse liver ADH. *Zm*PDC activity was detected in Mes/KOH-buffer, pH 6.5 [24] while the presented data were all measured in potassium phosphate buffer, pH 6.5. All substrates were applied at a concentration of 30 mM, except for indole-3-pyruvate (**14**), being applied at 1 mM due to low solubility and strong absorbance.

Optima and stability investigations- For *Ap*PDC, *Sc*PDC and *Zp*PDC the pH- and temperature optima were measured using the direct decarboxylase assay. To investigate the stability towards pH, temperature and organic solvents, *Ap*PDC was incubated under the conditions given in the figure legends. Residual activity was assayed with the coupled decarboxylase assay. It was verified, that the pH in the cuvettes was stable during the time of detection.

2.6. Carboligation

Benzoin formation- To avoid evaporation of the aldehydes the reaction batch was divided into 200 µL tubes. Reactions conditions (50 µL for each tube): 0-89 mM benzaldehyde in 15 µL dimethylsulfoxide (DMSO) (final concentration: 0-38 mM benzaldehyde, 30% (v/v) DMSO). The reaction was started by addition of 35 µL *Ap*PDC (4.2 mg/mL) in buffer A (pH 6.5), 30 °C, 100 rpm. The reaction was stopped by addition of 50 µL acetonitrile followed by vortexing and centrifugation of the precipitate. Calibration curves with benzoin were handled like the reaction samples.

The conversion was determined by HPLC, employing a Dionex HPLC (Germering), equipped with a 250 x 4.6 Multohyp ODS-5 µ (CS-Chromatography) and a UV-detector (mobile phase 59.5% H_2O : 40% acetonitrile: 0.5% acetic acid, flow 1.1 mL/min, pressure 130 bar, 20 µL injection volume, detection at $\lambda = 250$ nm), R_t (benzoin) = 32.2 min.

Mixed ligations of aromatic and aliphatic aldehyde- The conversion was followed by GCMS, employing a HP 6890 series GC-system fitted with a HP 5973 mass selective detector (Hewlett Packard; column HP-5MS, 30 m x 250 µm; T_{GC} (injector) = 250°C, T_{MS} (ion source) = 200°C, time program (oven): $T_{0\,min} = 60°C$, $T_{3\,min} = 60°C$, $T_{14\,min} = 280°C$ (heating rate 20°C·min^{-1}), $T_{19\,min} = 280°C$). The enantiomeric excess (*ee*) was determined by chiral phase HPLC employing a HP 1100 HPLC system (Agilent) fitted with a diode-array detector. Yields were calculated in mol% after NMR detection. NMR spectra were recorded on a Bruker DPX-400. Chemical shifts are reported in ppm relative to CHCl₃ (^{1}H NMR: $\delta = 7.27$) and CDCl₃ (^{13}C NMR: $\delta = 77.0$) as internal standards. CD-spectra were recorded on a JASCO J-810 spectropolarimeter using acetonitrile as solvent.

(R)-1-Hydroxy-1-phenylpropan-2-one (**20**)-

Reactions conditions: benzaldehyde (150 mg, 30 mM) and acetaldehyde (110 mg, 50 mM) were dissolved in 40 mL buffer A (pH 7.0) and 10 mL DMSO. After addition of purified *Ap*PDC dissolved in the same buffer (63 U/mL decarboxylase activity) the reaction was stirred slowly at 30°C. After 24 h additional acetaldehyde (50 mM) and *Ap*PDC (42 U/mL) were added. After 72 h the reaction mixture was extracted three times with ethyl acetate (25 mL) and the organic layer was dried over Na₂SO₄. The solvent was evaporated and the crude product was dissolved in diethylether (5 mL). The solution was washed with brine and dried over Na₂SO₄ followed by evaporation of the solvent. Purification of the crude product by flash chromatography revealed the designated product. *Analytical performance:* flash chromatography: cyclohexane/ethyl acetate 10:1; yield 30%; HPLC: Chiralcel-OD-H, n-hexane/2-propanol 90:10, 0.75 mL min^{-1}, 25 °C, R_t [(S)-**20**] = 11.25 min, R_t [(*R*)-**20**] = 12.56 min; ^{1}H NMR (400 MHz, CDCl₃, 300 K, ppm), $\delta = 2.10$ (s, 3 H, CH₃), 4.40 (bs, 1 H, OH), 5.10 (s, 1 H, C*H*OH), 7.30-7.41 (m, 5 H, Ar-H); ^{13}C NMR (100 MHz, CDCl₃, 300 K, ppm), $\delta = 25.3$ (CH₃), 80.2 (CHOH), 127.4 (CH), 128.8 (CH), 129.1 (CH), 138.1 (Cq), 207.2 (C=O); GCMS $R_t = 7.83$ min; MS (70 eV, EI): *m/z* (%): 150 (3%) [M$^+$], 107 (100%), 79 (68%).

(R)-1-Hydroxy-1-phenylbutan-2-one (**21**)-

Reactions conditions: benzaldehyde (32 mg, 20 mM) and propanal (35 mg, 40 mM) were dissolved in 12 mL buffer A (pH 7.0) and 3 mL DMSO. After addition of *Ap*PDC (72 U/mL) the reaction was stirred slowly at 30°C for 44 h before additional propanal (35 mg, 40 mM) and *Ap*PDC (72 U/mL) were added. After 72 h product extraction was performed as described for **20**. After purification of the crude product by thin layer chromatography 84 mol% **21** and 16 mol% propioin were obtained. *Analytical performance:* thin layer chromatography: cyclohexane/ethyl acetate 5:1, yield = 45%; HPLC: Chiralcel OD-H, n-hexane/2-propanol 95:5, 0.5 mL min^{-1}, 40°C, R_t [(*R*)-**21**] =

19.69 min, R_t [(S)-**21**] = 16.83 min; ^{1}H NMR (400 MHz, CDCl$_3$, 300 K, ppm), δ = 1.01 (t, J = 7.4 Hz, 3 H, CH$_3$), 2.29-2.46 (m, 2 H, CH$_2$), 4.30-4.4 (bs, 1 H, OH), 5.10 (s, 1 H, C*H*OH), 7.30-7.41 (m, 5 H, Ar-H); ^{13}C NMR (100 MHz, CDCl$_3$, 300 K, ppm), δ = 7.6 (CH$_3$), 31.2 (CH$_2$), 79.5 (CHOH), 127.3 (CH), 128.6 (CH), 128.9 (CH), 138.3 (Cq), 210.1 (C=O); GCMS R_t = 8.51 min, MS (70 eV, EI): *m/z* (%): 164 (0.1%) [M[+]], 107 (100%), 79 (81%).

(R)-1-Cyclopropyl-2-hydroxy-2-phenylethanone (**22**)-
Reactions conditions: Reaction was performed as described for **21** but with cyclopropyl-carbaldehyde (40 mM). After purification of the crude product by thin layer chromatography 76 mol% **22** and 24 mol% 2-hydroxy-1,2-dicyclopropylethanone were obtained. *Analytical performance:* thin layer chromatography: cyclohexane/ethyl acetate 5:1, yield = 6%; HPLC: Chiralcel OD-H, n-hexane/2-propanol 95:5, 0.5 mL min[-1], 40°C: R_t [(S)-**22**] = 16.6 min, R_t [(R)-**22**] = 21.15 min; CD (acetonitrile): λ ($\Delta\varepsilon$) [nm] (mol. CD) = 276 (-7.19), 219 (+8.36), 209 (+8.30); ^{1}H NMR (400 MHz, CDCl$_3$, 300 K, ppm), δ = 0.78-0.86 (m, 1 H), 0.94-1.10 (m, 1 H), 1.16-1.20 (m, 2 H), 1.83-1.90 (m, 1 H, CH), 4.39 (bs, 1 H, OH), 5.27 (s, 1 H, C*H*OH), 7.30-7.45 (m, 5 H, Ar-H); ^{13}C NMR (100 MHz, CDCl$_3$, 300 K, ppm), δ = 12.1 (CH$_2$), 12.7 (CH$_2$), 17.6 (CH), 80.1 (CHOH), 127.7 (CH), 128.6 (CH), 128.9 (CH), 138.1 (Cq), 209.6 (C=O); GCMS R_t = 9.45 min, MS (70 eV, EI): *m/z* (%): 176 [M[+]], 107 (100%), 79 (71%).

2.7. Crystallisation

*Ap*PDC was crystallised using the hanging drop vapor diffusion method. Droplets were set up for crystallisation by mixing 2 µL of purified protein solution containing 3 mg/mL protein in 20 mM Mes/NaOH-buffer, pH 6.5, 2.5 mM MgSO$_4$, 0.1 mM ThDP with 2 µL of reservoir solution. Screening and optimisation revealed a reservoir solution of 16-21% (w/v) PEG 1500, 0.1 M Mes/malic acid/Tris-buffer, pH 7-7.2 and 5-10 mM dithiothreithol (DTT) to be optimal. Diffraction-quality crystals appeared after three days of equilibration.

2.8. Data collection and processing

For cryoprotection the crystals were quickly dipped into the well solution supplemented with 25% ethylene glycol before being frozen in a cryogenic nitrogen gas stream at 110 K. Data were collected to a resolution of 2.75 Å at beamline I911-2 at Max-lab, Lund, Sweden. Images were processed using MOSFLM [25], where unit-cell parameters were determined by the auto indexing option. The data set was scaled with the program SCALA implemented in the CCP4 program suite [26].
The crystal belongs to the space group C2 with eight monomers packed in the asymmetric unit. Data collection statistics are given in Tab. 1.

2.9. Structure solution and crystallographic refinement

The structure of *Ap*PDC was determined by molecular replacement using the program MOLREP [26, 27]. As the sequence similarity between *Ap*PDC and *Zm*PDC is 62% (see supplementary material), the structure of *Zm*PDC (pdb accession code: 1zpd) [28] was used as the search model to place the eight monomers in the asymmetric unit. Atomic positions and B-factors of the model were refined by the maximum likelihood method in REFMAC5 [26, 29] which was interspersed with rounds of manual model building in COOT [30]. Non-crystallographic symmetry restraints were applied using tight main chain and side chain restraints. Water assignment was performed in COOT. The quality of the final structure was examined using PROCHECK [26, 31]. Statistics of the refinement and final model are given in Tab. 2. The coordinates of *Ap*PDC, in addition to the structure factors, have been deposited at the Research Collaboratory for Structural Bioinformatics Protein Databank PDB with accession code **2vbi**.

2.10. Structural analysis and models

Structural comparison of *Ap*PDC with *Zm*PDC (Fig. 5) was done with COOT (32). Comparison of *Ap*PDC with other ThDP-dependent enzymes (Fig. 7+8) was performed using the PyMOL program for visualisation [32] and the SwissPDB Viewer for superimposition [33].

3. Results and Discussion

3.1. Cloning, overexpression, purification and storage

The coding gene (pdc, gene bank AAM21208) was cloned and expressed in *E. coli* as a C-terminal hexahistidine fusion protein. Comparison of the amino acid sequences of *Ap*PDC with the one published by Raj et al. [4] revealed only 98% identity. Beside 9 amino acids exchanges in different positions of the protein, the *Ap*PDC described in this work contains 558 amino acids and is thus one amino acid longer than the published sequence [4] (for sequence alignment see supplementary material). A further C-terminal elongation by at least eight amino acids is due to the hexahistidine tag. Although the strains used in this work (DSM-No. 2347, DSMZ) and by Raj et al. [4] (ATCC 12874, American Type Culture Collection) were described to be the same the differences might be explained by varying sources of the *Acetobacter pasteurianus* strain sold by the companies. Because of the type of mutations it is unlikely that the differences appeared due to mistakes in the polymerase chain reaction. The sequence was confirmed twice in different clones.

A 15 L fed-batch fermentation of the recombinant *E. coli* strain yielded 1.42 kg cells with a cell specific activity of 750 U/g_{cells} (decarboxylase activity) in the crude extract and an overall activity of 1.07×10^6 U. The enzyme was purified by immobilised metal chelate chromatography yielding 48 g lyophilisate comprising, beside freeze-dried salts from the buffer solution, 14 g pure *Ap*PDC (protein purity according to SDS-gel > 95%). The specific activity measured directly after desalting was 110 U/mg. For the detection of the best storage conditions the specific activity of *Ap*PDC was followed for 4 weeks. Lyophilised *Ap*PDC as well as frozen fractions, both stored at -20 °C, kept an activity of 100 U/mg. However, storage under sterile conditions at 4 °C in 10 mM potassium phosphate buffer, pH 6.5, containing 2.5 mM $MgSO_4$ and 0.1 mM ThDP resulted in a drastically reduced catalytic activity (46 U/mg after 4 weeks).

3.2. Determination of the native molecular weight

The native molecular weight of the recombinant *Ap*PDC was determined by size-exclusion chromatography. The main peak (85.4%) eluted at an apparent molecular weight of 226 ± 1 kDa. As the calculated size of the monomeric subunit is 59.9 kDa, the observed data correlate best with a tetrameric structure of *Ap*PDC in the native state under the applied conditions. This is in accordance to most other tetrameric 2-keto acid decarboxylases, such as *Zm*PDC [28], *Sc*PDC [34, 35] and BFD [36], except for KdcA [37] showing a dimeric quaternary structure. A second small peak (14.6%) was detected corresponding to a molecular weight of 476 kDa, correlating well with an octamer and showing a specific activity comparable to the tetramer. Similar findings have also been described by Raj et al. [38]. In plant PDCs oligomers up to 300-500 kDa are rather common [39]. Small amounts of octamers were also found in *Zp*PDC (data not shown) and *Zm*PDC [40].

3.3. Substrate range and kinetics of decarboxylase activity

The investigation of the substrate range of the decarboxylase reaction is of interest in order to deduce information about the acyl donor spectrum for carboligase activity of *Ap*PDC. If a respective 2-keto acid is a substrate for decarboxylation, the binding of the corresponding aldehyde to the C2-atom of ThDP located in the active centre is most likely, meaning that this aldehyde may be a possible donor aldehyde in enzyme catalysed carboligation reactions.

As demonstrated in Tab. 3, *Ap*PDC prefers aliphatic non-branched 2-keto acids. The longer the side chain, the less activity is observed. Nevertheless, activity is also evident with branched-chain (**7-9**) and aromatic substrates (**11-13**), with an exceptionally high activity for benzoylformate compared to other PDCs (**11**). Thus, the substrate range of *Ap*PDC is similar to *Zp*PDC and broader compared to *Zm*PDC [24] and *Sc*PDC (Tab. 3). It should be mentioned that the substrate range was investigated with a concentration of 30 mM for each substrate. It is not very likely but still possible that the maximum rate of each reaction could not be detected correctly due to high K_M values or substrate surplus effects.

Kinetic constants were determined for the physiological substrate pyruvate (**1**). *Ap*PDC shows a hyperbolic v/[S]-plot up to 75 mM with a maximal velocity (V_{max}) of 110 U/mg and a K_M of 2.7 mM (Tab. 4, Fig. 1A). These values agree well with the published data obtained in citrate/phosphate buffer ([38], Tab. 4). Whereas *Zm*PDC [41] and *Zp*PDC show similar hyperbolic v/[S]-plots, substrate activation in *Sc*PDC gives rise to a sigmoidal v/[S]-plots. Additionally the plot shows a decrease of activity at concentrations > 30 mM pyruvate (Fig. 1B). The latter might be

a result of substrate surplus inhibition or increased ionic strength. The sigmoidal v/[S]-plot of brewers yeast PDC [42] and *Sc*PDC [39] have already been described and similar results have been obtained with plant PDCs [43] as well as with the prokaryotic PDC from *Sarcina ventriculi* (*Sv*PDC) [44]. Just recently the mechanism for the allosteric activation of yeast PDC could be elucidated [45, 46].

3.4. Optimal reaction conditions

Optimal cofactor concentration- As in all ThDP-dependent enzymes the cofactors in *Ap*PDC are bound non-covalently to the active site. The main contributions to the binding arise from the coordinative interaction of the diphosphate moiety via Mg^{2+} to the protein as well as from hydrophobic and ionic interactions between protein side chains and the thiazol- and pyrimidine ring of ThDP [47]. For stability of the holoenzyme most ThDP-dependent enzymes require the addition of cofactors to the buffer. Under the applied *Ap*PDC concentrations the addition of 2.5 mM $MgSO_4$ and 0.1 mM ThDP is sufficient to keep the enzyme stable and active in buffer for several days.

pH-dependent activity and stability- For *Ap*PDC maximal initial rate activity for the decarboxylation of pyruvate (**1**) was observed in an exceptional broad pH-range of pH 3.5-6.5 (Fig. 2). Similar results were observed by Raj et al. in sodium citrate buffer at 25°C [38]. However, the stability of *Ap*PDC decreases significantly at pH 4 (half-life time: 2.3 h), while it is completely stable in the range of pH 5-7 within 120 min (data not shown). In comparison the pH-optima of *Sc*PDC (pH 5-7), *Zp*PDC (pH 7) (data not shown) and *Zm*PDC (pH 6-6.5) [48] are significantly less broad and overlap well with the stability optima.

Temperature-dependent activity and stability- Under initial rate conditions (90 sec) the temperature optimum of *Ap*PDC was observed at 65°C with a specific activity of 188 U/mg. A further temperature increase resulted in a fast decay of activity with a midpoint of thermal inactivation (Tm) at about 70°C (Fig. 3). This optimum is considerably higher than that of *Sc*PDC (43°C) and *Zp*PDC (55°C) (data not shown) and comparable to data obtained with *Zm*PDC (60°C) [49].

From these data the activation energy of 27.1 kJ/mol was calculated for the decarboxylase reaction from an $lnV_{max}/[1/T]$-plot in the linear range of 20-50°C (see supplementary material), which is lower compared to other pyruvate decarboxylases. For *Zp*PDC an activation energy of 41 kJ/mol and for *Zm*PDC 43 kJ/mol were determined.

In order to identify optimal conditions for the application of *Ap*PDC in enzymatic syntheses, the temperature stability of the enzyme was determined (Tab. 5). While the enzyme is sufficiently stable up to 30°C, it rapidly looses activity at higher temperatures. The same is valid for *Zp*PDC, whereas the stability optimum of *Sc*PDC is below 30°C (Tab. 5). In comparison *Zm*PDC is more thermostable with a half-life time of 24 h at 50°C [49].

Stability towards dimethylsulfoxide (DMSO)- The biotransformation of aromatic aldehydes is often hampered by their low solubility in aqueous systems. Since the addition of 20% (v/v) DMSO has been successfully applied for BAL, BFD and KdcA catalysed carboligase reactions [9, 11, 12, 14, 18, 50] addition of DMSO was tested also with *Ap*PDC. As demonstrated in Fig. 4, *Ap*PDC is not just stable in the presence of up to 30% (v/v) DMSO (half-life: ~430 h), the stability is even considerably enhanced compared to the buffer without DMSO addition.

3.5. Carboligase activity

The substrate range of the decarboxylase reaction suggests *Ap*PDC to prefer short non-branched aliphatic donor aldehydes in carboligase reactions. We investigated the carboligase activity in mixed carboligation reactions with benzaldehyde and various aliphatic aldehydes. The results in Tab. 6 show that the main product in such biotransformations is indeed always the phenylacetyl-carbinol derivative (**20-22**), which results from the reaction of an aliphatic donor aldehyde (**17-19**) and benzaldehyde (**16**) as the acceptor aldehyde. As the aliphatic aldehyde was applied in excess, its self-ligation product is obtained as a side product (~16 mol% 4-hydroxyhexan-3-one (entry **2**, Tab. 6) and 24 mol% 1,2-dicyclopropyl-2-hydroxyethanone (entry **3**, Tab. 6), respectively). However, neither benzoin nor mixed carboligation products derived from benzaldehyde as a donor aldehyde have been observed. Even with benzaldehyde as the only substrate just a very low benzoin forming activity was determined (2.4 10^{-3} U/mg), which is typical for pyruvate decarboxylases [19].

Concerning the carboligation *Ap*PDC is less enantioselective (*ee* 81-91%, Tab. 6) than *Zm*PDC, which is strictly *R*-selectivity (98%) with respect to the formation of **20** [17].

The reduced stereo control of *Ap*PDC suggests that its active site might be different to *Zm*PDC. In order to rationalise these differences, the enzyme was crystallised to determine the structure.

3.6 Crystal structure

Quality of the electron density map and the model- The crystal structure of *Ap*PDC, determined to 2.75 Å resolution, was solved by molecular replacement. The asymmetric unit contains eight monomers, arranged as two tetrameric *Ap*PDC units. The structure was refined to an R-factor of 21.9% and an R_{free} of 24.0%.

8 x 554 amino acids (comprising residue 2-555) could be placed in the electron density map, with each subunit binding one ThDP molecule and one Mg^{2+}-ion anchoring the cofactor. 91.3% of the amino acid residues in the final model are placed in the most favored regions of the Ramachandran plot. One residue, S74, has a strained conformation in the disallowed region of the plot, but is well defined in the electron density (Tab. 1).

*Overall and quaternary structure- Ap*PDC shows the general fold of ThDP-dependent enzymes consisting of three domains with α/β-topology [35, 39] (Fig. 5). Each domain is comprised of a central β-sheet surrounded by α-helices. The N-terminal PYR (pyrimidine)- and C-terminal PP (pyrophosphate)-domains are topologically equivalent (residues 1-188 and 349-555) while the R (regulatory)-domain has a different fold (residues 189-348).

Two monomers form the minimal functional unit, with the coenzyme bound in the cleft between the PYR-domain of one monomer and the R-domain of the second monomer. ThDP is bound in the conserved V-conformation, forced by a large hydrophobic residue next to the thiazolium ring (Ile-411), positioning the N4'-imino group of the pyrimidine ring in a reactive distance to the C2 of the thiazolium ring where the catalytic reaction takes place. The packing of tetrameric units in the crystal supports the size exclusion chromatography results indicating that *Ap*PDC is a tetramer in solution. The oligomeric form can be best described as a dimer of dimers. Superimposition of *Zm*PDC and *Ap*PDC (Fig. 5) reveals the backbones of all three domains to be very similar. There are only minor differences in the loops connecting the domains (encompassing the residues 181-190 and 345-356 in *Ap*PDC) and the C-terminus, which is elongated in *Zm*PDC by four extra residues inserted in a loop relative to *Ap*PDC. Like *Zm*PDC, *Ap*PDC shows a similar compact arrangement of the tetrameric structure.

ThDP-binding site- A superimposition of the crystal structure of *Ap*PDC with the known structures of *Zm*PDC [28], *Sc*PDC [34], KdcA [37], BFD [36] and BAL [51] combined with the biochemical data concerning substrate specificities of these enzymes and variants thereof provides insights into several aspects of the active site architecture in these enzymes. In Tab. 7 important residues lining the substrate binding site including the *S*-pocket [19], catalytically important residues as well as amino acids forming the substrate channel are given. Their relative positions are visualised in Fig. 6.

Substrate channel and entrance- In contrast to BFD, which has no C-terminal helix [20], the entrance to the substrate channel in *Ap*PDC is not directly visible on the surface of the 3D structure due to a C-terminal helix covering this region (area C-Fig. 6, I). A similar 4-turn helix is found in *Zm*PDC while *Sc*PDC has a helix with 2-3 turns and BAL a shorter one with 1.5 turns (Tab. 7). It is assumed that the movement of the C-terminal helix is important for the catalytic activity [28]. This is supported by results published by Chang et al. [52] who investigated variants of *Zm*PDC with a C-terminal truncation. Whereas the deletion of the seven very last C-terminal amino acids has no effect on the decarboxylase activity, the next few, R561 and S560, are critical not only for the decarboxylase activity but also for the cofactor binding.

A comparison of the substrate channels leading to the active centres shows a very wide and straight substrate channel for BFD (Fig. 6, II), whereas the channels in *Ap*PDC and *Zm*PDC are less wide and curved (Fig. 6, I). The dimensions of the channels in BAL, KdcA and *Sc*PDC range between those of BFD and *Ap*PDC/*Zm*PDC.

Donor binding site- The reaction mechanism of ThDP-dependent decarboxylation has been intensively studied for PDCs [35, 53, 54] and BFD [55, 56]. The substrate is attacked by the activated ylid of ThDP followed by the release of CO_2 and the formation of the covalent α-carbanion/enamine intermediate, the so-called "active aldehyde". In the decarboxylation reaction

an aldehyde is released after protonation of the α-carbon, while during carboligation it reacts with a second aldehyde as an acyl acceptor [57]. Whereas the decarboxylation requires the binding of only one substrate (a 2-keto acid), the arrangement of two substrates (aldehydes) in close proximity in the active site is necessary for carboligation. The binding site and orientation of the second (acceptor) aldehyde is defined by the space which is left after the first (donor) aldehyde is bound to the C2 atom of the ThDP in V-conformation, and the necessity to arrange the carbonyl groups of the reacting aldehydes close to a histidine residue operating as a proton relay system during the catalysis [16] (Fig. 6). As a consequence overlapping but clearly distinct binding regions can be assumed for the keto-acid or respectively the donor and the acceptor aldehyde (Fig. 6).

Differences in the substrate range of the ThDP-dependent lyases are rationalised by different sizes of the respective binding sites. The preference of PDCs for short aliphatic 2-keto acids as well as short aliphatic donor aldehydes is mainly a result of the amino acids in position B (Fig. 6), which is occupied by a bulky tryptophane in bacterial PDCs (Tab. 7). Compared to *Zm*PDC the binding site for 2-keto acids and donor aldehydes is slightly larger in *Ap*PDC. This might explain why *Ap*PDC is able to decarboxylate benzoylformate at least with low activity in contrast to *Zm*PDC, not showing any product formation (Tab. 3). Modeling studies demonstrate that the space in the substrate binding site, which is mainly restricted by W388, is just large enough for a ThDP-bound benzaldehyde (Fig. 7). Nevertheless, aliphatic aldehydes are preferred as donors in mixed carboligations with aliphatic and aromatic aldehydes.

In BFD as well as in BAL and KdcA this part of the active site is wide enough to bind large aromatic substrates. However, the size of the binding site is not the only criterion to explain substrate transformation. Optimal stabilisation of a molecule in a binding site is also important to allow the rapid formation of an enzyme-substrate complex. This may explain why BFD shows only very little activity towards aliphatic 2-keto acids and aliphatic donor aldehydes although there is enough space available in the substrate binding site.

The aspect of optimal substrate stabilisation in the binding site is extremely important in mixed carboligations of two different aldehydes. The product range accessible by such biotransformations can yield a maximum of four different products each as pair of enantiomers [57]. Fortunately, the various constrains in the different enzymes result in an almost selective formation of one predominant ligation product, if the two aldehydes to be ligated are sufficiently different. Thus, in the case of an aliphatic and an aromatic aldehyde, PDCs prefer the binding of the aliphatic aldehyde as the donor, whereas BFD prefers the aromatic aldehyde. There are two possibilities for the acceptor aldehyde to approach the ThDP-bound donor: in a parallel mode, where both side chains of the donor- and the acceptor aldehyde point into the substrate channel or in an antiparallel mode, with the acceptor side chain pointing into the so-called *S*-pocket [20]. This step is decisive for the stereoselectivity of the ligation step and the chemoselectivity of the reaction, since both aldehydes present in the reaction mixture can also react as acceptors. Again selectivity is decided by steric requirements and optimal stabilisation as was recently described for BFD and BAL [19,20]. Currently BFD is the only known enzyme, where such an *S*-pocket is accessible in the wild-type enzyme making (*S*)-2-hydroxy propiophenones available.

S-pocket- This structural feature of the substrate binding site enables BFD to catalyse *S*-selective carboligation (Fig. 6). Recently we described similar but due to sterically reasons not accessible pockets in the 3D structures of other ThDP-dependent decarboxylases, such as KdcA, *Sc*PDC and *Zm*PDC [19]. In *Ap*PDC this pocket is by far the largest pocket observed so far (Fig. 8). However, the *S*-pockets in PDCs are blocked by bulky isoleucine residues (Tab. 7), which is the main reason for the usually high *R*-selectivity of the PDCs [19]. For *Zm*PDC it was demonstrated that chemo- and stereoselectivity of the carboligase reaction can be influenced by point mutation of this position (*Zm*PDCI472A) [17]. Since *Ap*PDC (91% **20**) is less enantioselective compared to *Zm*PDC (98% **20**) [17] concerning the carboligation of acetaldehyde and benzaldehyde, it can be assumed that the large *S*-pocket in *Ap*PDC is not completely blocked by isoleucine 468, which would explain the relative large amount of (*S*)-products formed, lowering the enantiomeric excess of the predominant (*R*)-product (Tab. 6). This hypothesis is supported by modelling studies which show that the *S*-pocket in *Ap*PDC is large enough to accept the bulky benzaldehyde substrate, if flexible side chains that adopt to the bound substrate (especially in the entrance region to the large *S*-pocket) are assumed.

4. Conclusions

*Ap*PDC resembles very much the well characterised enzyme from *Zymomonas mobilis* with respect to sequence, kinetic behaviour and stability. However, there are some remarkable differences concerning the catalytic activity: the decarboxylation of aromatic 2-ketoacids, especially benzoylformate and a significantly lower stereoselectivity during carboligation. Both observations can be explained based on the 3D structure, showing a larger active site compared to *Zm*PDC. *Ap*PDC prefers aliphatic donor aldehydes, as these fit best into the binding site. But the donor binding site is even large enough to accept benzaldehyde, which is not possible with *Zm*PDC. There are also differences in the acceptor binding site. In *Zm*PDC the acceptor binding site (Fig. 6) is wide enough to bind aromatic aldehydes, such as benzaldehyde, in one preferred orientation: with the aromatic side chain being directed into the substrate channel. Thus *Zm*PDC catalyses the synthesis of enantiomerically pure (*R*)-phenylacetylcarbinol from acetaldehyde as the donor and benzaldehyde as the acceptor. In *Ap*PDC the situation is different. The *S*-pocket, as part of the acceptor binding site, is also accessible. As the *S*-pocket in *Ap*PDC is large enough to bind even aromatic side chains, the stereocontrol of carboligase reactions with benzaldehyde is less strict then for *Zm*PDC. The amount of (*S*)-product formed is directly correlated with the size of the donor aldehyde. Thus if acetaldehyde is replaced by propanal or cyclopropylcarbaldehyde the amount of *S*-product increases resulting in a decrease of the enantioselectivity of the (*R*)-product (Tab. 6). Here the larger donor aldehydes occupy more space in the substrate binding site, thereby making the antiparallel arrangement of benzaldehyde in the *S*-pocket more likely. Together with our previous studies on BFD and BAL [19, 20], these structure-function studies pave the way for the design of *S*-specific ThDP-dependent enzymes.

Acknowledgements

References

[1]   M. Pohl, Protein design on pyruvate decarboxylase (PDC) by site-directed mutagenesis, Adv. Biochem. Eng./Biotechnol. 58 (1997) 15-43.

[2]   E.A. Dawes, D.W. Ribbons, P.J. Large, The Route of Ethanol Formation in *Zymomonas mobilis*, Biochem. J. 98 (1966) 795-803.

[3]   T. Okamoto, K. Taguchi, H. Nakamura, H. Ikenaga, H. Kuraishi, K. Yamasato, *Zymobacter palmae* gen. nov., sp. nov., a new ethanol-fermentating peritrichious bacterium isolated from palm sap, Arch. Microbiol. 160 (1993) 333-337.

[4]   K.C. Raj, L.O. Ingram, J.A. Maupin-Furlow, Pyruvate decarboxylase: a key enzyme for the oxidative metabolism of lactid acid by *Acetobacter pasteurianus*, Arch. Microbiol. 176 (2001) 443-451.

[5]   S.E. Lowe, J.G. Zeikus, Purification and characterization of pyruvate decarboxylase from *Sarcina ventriculi*, J. Gen. Microbiol. 138 (1992) 803-807.

[6]   R. Kalscheuer, T. Stolting, A. Steinbüchel, Microdiesel: *Escherichia coli* engineered for fuel production, Microbiology 152 (2006) 2529-2536.

[7]   C. Neuberg, J. Hirsch, Über ein Kohlenstoffketten knüpfendes Ferment (Carboligase), Biochem. Zeitschr. 115 (1921) 282.

[8]   V.B. Shukla, P.R. Kulkarni, L-Phenylacetylcarbinol (L-PAC): biosynthesis and industrial applications, World J. Microbiol. Biotechnol. 16 (2000) 499-506.

[9]   A.S. Demir, O. Şeşenoglu, P. Dünkelmann, M. Müller, Benzaldehyde lyase-catalyzed enantioselective carboligation of aromatic aldehydes with mono- and dimethoxy acetaldehyde, Org. Lett. 5 (2003) 2047-2050.

[10]   A.S. Demir, O. Şeşenoglu, E. Eren, B. Hosrik, M. Pohl, E. Janzen, D. Kolter, R. Feldmann, P. Dünkelmann, M. Müller, Enantioselective synthesis of alpha-hydroxy ketones via benzaldehyde lyase-catalyzed C-C bond formation reaction, Adv. Synth. Catal. 344 (2002) 96-103.

[11]   P. Domínguez de María, M. Pohl, D. Gocke, H. Gröger, H. Trauthwein, L. Walter, M. Müller, Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes, Eur. J. Org. Chem. (2007) 2940-2944.

[12]   P. Domínguez de María, T. Stillger, M. Pohl, S. Wallert, K. Drauz, H. Gröger, H. Trauthwein, A. Liese, Preparative enantioselective synthesis of benzoins and (R)-2-hydroxy-1-phenylpropane using benzaldehyde lyase, J. Mol. Catal. B- Enzym. 38 (2006) 43-47.

[13]   P. Dünkelmann, M. Pohl, M. Müller, Enantiomerically pure 2-hydroxy carbonyl compounds through enzymatic C-C bond formation, Chim. Oggi/Chemistry Today supplement Chiral Catalysis 22 (2004) 24-28.

[14]   D. Gocke, C.L. Nguyen, M. Pohl, T. Stillger, L. Walter, M. Müller, 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 (2007) 1425-1435.

[15]   M. Pohl, B. Lingen, M. Müller, Thiamin-diphosphate-dependent enzymes: new aspects of asymmetric C-C bond formation, Chem. Eur. J. 8 (2002) 5288-5295.

[16]   M. Pohl, G.A. Sprenger, M. Müller, A new perspective on thiamine catalysis, Curr. Opin. Biotech. 15 (2004) 335-342.

[17]   P. Siegert, M.J. McLeish, M. Baumann, H. Iding, M.M. Kneen, G.L. Kenyon, M. Pohl, Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*, Protein Eng. Des. Sel. 18 (2005) 345-357.

[18]   T. Stillger, M. Pohl, C. Wandrey, A. Liese, Reaction engineering of benzaldehyde lyase catalyzing enantioselective C-C bond formation, Org. Process Res. Dev. 10 (2006) 1172-1177.

[19]   D. Gocke, L. Walter, E. Gauchenova, G. Kolter, M. Knoll, C.L. Berthold, G. Schneider, J. Pleiss, M. Müller, M. Pohl, Rational Protein Design of ThDP-dependent Enzymes: Engineering Stereoselectivity, ChemBioChem 9 (2008) 406-412.

[20]   M. Knoll, M. Müller, J. Pleiss, M. Pohl, Factors Mediating Activity, Selectivity, and Substrate Specificity for the Thiamin Diphosphate-Dependent Enzymes Benzaldehyde Lyase and Benzoylformate Decarboxylase, ChemBioChem 7 (2006) 1928-1934.

[21]   D.J. Korz, U. Rinas, K. Hellmuth, E.A. Sanders, W.-D. Deckwer, Simple fed-batch technique for high cell density cultivation of *Escherichia coli*, J. Biotechnol. 39 (1995) 59-65.

[22]   M. Pohl, P. Siegert, K. Mesch, H. Bruhn, J. Grötzinger, Active site mutants of pyruvate decarboxylase from *Zymomonas mobilis* - a site-directed mutagenesis study of L112, I472, I476, E473, and N482, Eur. J. Biochem. 257 (1998) 538-546.

[23]   M.M. Bradford, A rapid and sensitive method for the quantification of microgram quantities of protein using the principle of protein-binding dye, Anal. Biochem. 72 (1976) 248-254.

[24]   P. Siegert, Vergleichende Charakterisierung der Decarboxylase- und Carboligasereaktion der Benzoylformiatdecarboxylase aus *Pseudomonas putida* und der Pyruvatdecarboxylase aus *Zymomonas mobilis* mittels gerichteter Mutagenese, doctoral thesis, Heinrich-Heine University Duesseldorf (2000).

[25]   A.G.W. Leslie, Recent changes to the MOSFLM package for processing film and image plate data, Joint CCP4 + ESF-EAMCB Newsl. Protein Crystallogr. 26 (1992).

[26]   Collaborative Computational Project Number 4, The CCP4 suite: programs for protein crystallography, Acta Crystallogr. Sec. D: Biol. Crystallogr. 50 (1994) 760-763.

[27]   A. Vagin, A. Teplyakov, MOLREP: an automated program for molecular replacement, J. Appl. Crystallogr. 30 (1997) 1022-1025.

[28]   D. Dobritzsch, S. König, G. Schneider, G. Lu, High resolution crystal structure of pyruvate decarboxylase from *Zymomonas mobilis*. Implications for substrate activation in pyruvate decarboxylases, J. Biol. Chem. 273 (1998) 20196-20204.

[29]   G.N. Murshudov, A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. Sect. D: Biol. Crystallogr. 53 (1997) 240-255.

[30]   P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. Sec. D: Biol. Crystallogr. 60 (2004) 2126-2132.

[31]   R.A. Laskowsky, M.W. McArthur, D.S. Moss, J.M. Thornton, PROCHECK, J. Appl. Crystallogr. 26 (1993) 282-291.

[32]   W.L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific, San Carlo, CA, USA, 2002.

[33]   N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714-2723.

[34]   P. Arjunan, T. Umland, F. Dyda, S. Swaminathan, W. Furey, M. Sax, B. Farrenkopf, Y. Gao, D. Zhang, F. Jordan, Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* at 2.3 A resolution, J. Mol. Biol. 256 (1996) 590-600.

[35]   F. Jordan, M. Liu, E. Sergienko, Z. Zhang, A. Brunskill, P. Arjunan, W. Furey, Yeast pyruvate decarboxylase: New features of the structure and mechanism, in: Jordan, F. and Patel, R.N. (Eds.), Thiamine-Catalytic Mechanisms in Normal and Disease States, Marcel Dekker Inc., New York / Basel, 2004.

[36]   M.S. Hasson, A. Muscate, M.J. McLeish, L.S. Polovnikova, J.A. Gerlt, G.L. Kenyon, G.A. Petsko, D. Ringe, The crystal structure of benzoylformate decarboxylase at 1.6 Å resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes, Biochemistry 37 (1998) 9918-9930.

[37]   C.L. Berthold, D. Gocke, M.D. Wood, F.J. Leeper, M. Pohl, G. Schneider, Crystal structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemo- and enantioselective carboligation reaction, Acta Crystallogr. Sec. D: Biol. Crystallogr. 63 (2007) 1217-1224.

[38]   K.C. Raj, L.A. Talarico, L.O. Ingram, J.A. Maupin-Furlow, Cloning and characterization of the *Zymobacter palmae* pyruvate decarboxylase gene (pdc) and comparison to bacterial homologues, Appl. Environ. Microbiol. 68 (2002) 2869-2876.

[39]   S. König, Subunit structure, function and organisation of pyruvate decarboxylases from various organisms, Biochim. Biophys. Acta 1385 (1998) 271-286.

[40]   M. Pohl, J. Grötzinger, A. Wollmer, M.R. Kula, Reversible dissociation and unfolding of pyruvate decarboxylase from *Zymomonas mobilis*, Eur. J. Biochem. 224 (1994) 651-661.

[41]   S. Bringer-Meyer, K.L. Schimz, H. Sahm, Pyruvate decarboxylase from *Zymomonas mobilis*. Isolation and partial characterization, Arch. Microbiol. 146 (1986) 105-110.

[42]   G. Hübner, R. Weidhase, A. Schellenberger, The mechanism of substrate activation of pyruvate decarboxylase. A first approach, Eur. J. Biochem. 92 (1978) 175-181.

[43]   A. Dietrich, S. König, Substrate activation behaviour of pyruvate decarboxylase from *Pisum sativum* cv. Miko, FEBS Lett. 400 (1997) 42-44.

[44]   L.A. Talarico, L.O. Ingram, J.A. Maupin-Furlow, Production of the Gram-positive *Sarcina ventriculi* pyruvate decarboxylase in *Escherichia coli*, Microbiology 147 (2001) 2425-2435.

[45]   G. Lu, D. Dobritzsch, S. Baumann, G. Schneider, S. König, The structural basis of substrate activation in yeast pyruvate decarboxylase. A crystallographic and kinetic study, Eur. J. Biochem. 267 (2000) 861-868.

[46]   S. Kutter, M. Weiss, G. Wille, R. Golbik, M. Spinka, S. König, Covalently bound substrate at the regulatory site triggers allosteric enzyme activation. Nat. Proc. (2008) precedings.nature.com.

[47]   R.A.W. Frank, F.J. Leeper, B.F. Luisi, Structure, mechanism and catalytic duality of thiamine-dependent enzymes, Cell. Mol. Life S. 64 (2007) 829-905.

[48]   A.D. Neale, R.K. Scopes, R.E. Wettenhall, N.J. Hoogenraad, Pyruvate decarboxylase of *Zymomonas mobilis*: isolation, properties, and genetic expression in *Escherichia coli*, J. Bacteriol. 169 (1987) 1024-1028.

[49]   M. Pohl, K. Mesch, A. Rodenbrock, M.R. Kula, Stability investigations on the pyruvate decarboxylase from *Zymomonas mobilis*, Biotechnol. Appl. Biochem. 22 (1995) 95-105.

[50]   F. Hildebrand, S. Kühl, M. Pohl, D. Vasic-Racki, M. Müller, C. Wandrey, S. Lütz, The production of (*R*)-2-hydroxy-1-phenyl-propan-1-one derivatives by benzaldehyde lyase from *Pseudomonas fluorescens* in a continuously operated membrane reactor, Biotechnol. Bioeng. 96 (2007) 835-843.

[51]   T.G. Mosbacher, M. Müller, G.E. Schulz, Structure and mechanism of the ThDP-dependent benzaldehyde lyase from *Pseudomonas fluorescens*, FEBS J. 272 (2005) 6067-6076.

[52]   A.K. Chang, P.F. Nixon, R.G. Duggleby, Effects of deletions at the carboxyl terminus of *Zymomonas mobilis* pyruvate decarboxylase on the kinetic properties and substrate specificity, Biochemistry 39 (2000) 9430-9437.

[53]   D. Kern, G. Kern, H. Neef, K. Tittmann, M. Killenberg-Jabs, C. Wikner, G. Schneider, G. Hübner, How thiamine diphosphate is activated in enzymes, Science 275 (1997) 67-70.

[54]   A. Schellenberger, Sixty years of thiamin diphosphate biochemistry, Biochim. Biophys. Acta 1385 (1998) 177-186.

[55]   E.S. Polovnikova, M.J. McLeish, E.A. Sergienko, J.T. Burgner, N.L. Anderson, A.K. Bera, F. Jordan, G.L. Kenyon, M.S. Hasson, Structural and kinetic analysis of catalysis by a thiamin diphosphate-dependent enzyme, benzoylformate decarboxylase, Biochemistry 42 (2003) 1820-1830.

[56]   P.M. Weiss, G.A. Garcia, G.L. Kenyon, W.W. Cleland, P.F. Cook, Kinetics and mechanism of benzoylformate decarboxylase using ^{13}C and solvent deuterium isotope effects on benzoylformate and benzoylformate analogues, Biochemistry 27 (1988) 2197-2205.

[57]   H. Iding, P. Siegert, K. Mesch, M. Pohl, Application of alpha-keto acid decarboxylases in biotransformations, Biochim. Biophys. Acta 1385 (1998) 307-322.

Figure legends

Fig. 1. Kinetics of *Ap*PDC (A) and *Sc*PDC (B) for the decarboxylation of pyruvate. h = Hill coefficient (if > 1, cooperativity confirmed), K_s= inhibition constant (B). Data were measured in 50 mM potassium phosphate buffer, pH 6.5, 2.5 mM $MgSO_4$, 0.1 mM ThDP.

Fig. 2. Determination of the pH-optimum of *Ap*PDC for the decarboxylation of pyruvate. Data were obtained using the direct decarboxylase assay under initial rate conditions.

Fig. 3. Determination of the temperature optimum of *Ap*PDC for the decarboxylation of pyruvate. Data were obtained using the direct decarboxylase assay under initial rate conditions.

Fig. 4. Stability of *Ap*PDC in the presence and absence of DMSO. Incubation in the presence ♦ and absence ◊ of 30% (v/v) DMSO in 50 mM potassium phosphate puffer, pH 6.5 (pH was adjusted before DMSO addition), 2.5 mM $MgSO_4$, 0.1 mM ThDP at 30 °C. Residual activities were measured using the coupled decarboxylase assay.

Fig. 5. Ribbon representation of the *Ap*PDC subunit (green) superimposed with the *Zm*PDC subunit (blue) [28]. Superposition of the two structures results in an rms deviation of 0.81 Å over 549 Cα-atoms. The three domains are coloured in different shades. ThDP is shown as sticks and the Mg^{2+} ion is shown as a sphere coloured in magenta.

Fig. 6. Schematic presentation of the substrate channel and the active site of *Ap*PDC/*Zm*PDC (I) and BFD (II). ThDP is bound in the V-conformation in the active site. x = C2-atom of the thiazolium ring, A = residues lining the *S*-pocket, A1 = most prominent residues determining the size of the *S*-pocket, A2 = residues defining the entrance to the *S*-pocket, a = acceptor binding site; B = bulky residue in *Ap*PDC/*Zm*PDC situated opposite to the substrate binding site where the donor binds (I), b = binding site for 2-keto acid and donor aldehyde, x = C2-atom of the thiazolium ring; C = C-terminal α-helix covering the entrance of the substrate channel in *Ap*PDC/*Zm*PDC (I).

Fig. 7. Comparison of the substrate binding sites of *Ap*PDC and BFD. Superimposition of the crystal structures of the binding site of *Ap*PDC (green) and BFD (blue). In *Ap*PDC residue W388 (marked in orange) limits the pockets size.

Fig. 8. Superimposition of the active site of *Ap*PDC (green), *Zm*PDC (light blue) and *Sc*PDC (dark blue). The pocket is mainly defined by the side chain of a conserved glutamate residue (E469 in *Ap*PDC, E473 in *Zm*PDC and E477 in *Sc*PDC). Differences in the size of the *S*-pocket result from differences in the orientation of this glutamate residue and are not due to alternate backbone conformations.

Tables

Tab. 1. Data collection statistics of *Ap*PDC. Values in parentheses are given for the highest resolution interval.
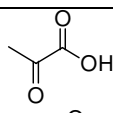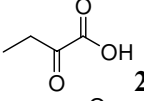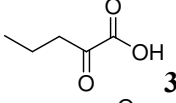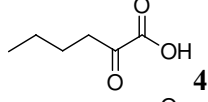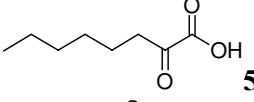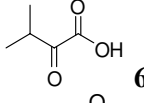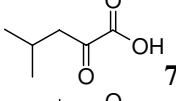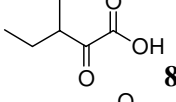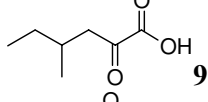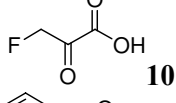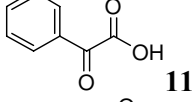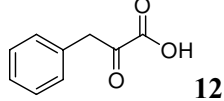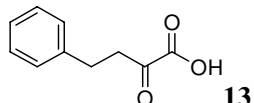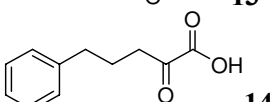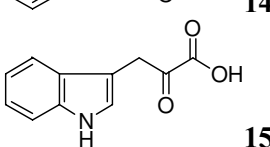
Space group	C2
Molecules in asymmetric unit	8
Unit cell dimensions	$a = 179.62$ Å, $b = 162.43$ Å, $c = 169.33$ Å
No. of observations	289630 (42113)
No. of unique reflections	119238 (17599)
Resolution	2.75 (2.9-2.75) Å
R-merge	0.083 (0.433)
Mn (I/σ(I))	10.8 (2.0)
Completeness	97.2 (98.5)%
Multiplicity	2.4 (2.4)%
Wilson B-factor	69.7 Å^2

Tab. 2. Structure refinement and final model statistics. TLS model = twin lattice symmetry model; rms deviation = root means square deviation

Refinement program	REFMAC5
TLS model	8 TLS groups
Reflections in working set	107178
Reflections in test set	5995
R-factor	21.9%
R-free	24.0%
Atoms modeled	33860
Number of amino acids	8 x 554
Number of ThDPs	8
Number of magnesium ions	8
Number of waters	92
Average B-factor protein*	42.4 Å^2
Average B-factor ligands	38.3 Å^2
Average B-factor magnesium ions	34.7 Å^2
Average B-factor waters	29.0 Å^2
RMS deviations from ideals	
Bonds	0.009 Å
Angles	1.1044 °
Ramachandran distribution	
Most favored	91.3%
Additionally allowed	8.5%
Generously allowed	0.0%
Disallowed	0.2%

* Contains residual B-factors only, does not include TLS contribution

Tab. 3. Substrate range of the decarboxylation reaction of *Ap*PDC compared to *Zm*PDC [24], *Sc*PDC and *Zp*PDC. Activity towards different 2-keto acids was measured using the coupled decarboxylase assay. All substrates were applied in a concentration of 30 mM, except **15** (1 mM).

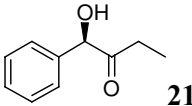| 2-keto acid	specific activity [U/mg]			
	*Ap*PDC	*Zm*PDC	*Sc*PDC	*Zp*PDC
1	89.3	120	43.4	147.0
2	60	79	16.9	85.1
3	12.9	13	18.8	20.7
4	4.2	0.2	5.3	6.2
5	1.1	-	0.0	0.6
6	-	0.0	6.9	10.4
7	1.1	0.3	0.3	2.8
8	2.2	0.0	0.0	2.8
9	0.6	0.0	0.0	0.4
10	0.3	-	0.0	0.6
11	1.1	0.0	0.0	0.3
12	0.8	0.0	0.0	1.8
13	0.3	0.0	1.7	0.3
14	0.0	0.0	0.2	0.2
15	0.0	-	0.0	0.5

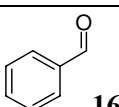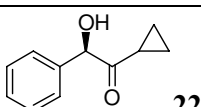- = not determined, 0.0 = activity < 0.05 U/mg
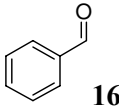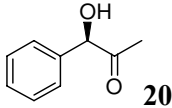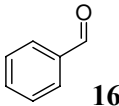
Tab. 4. Kinetic parameters for the decarboxylation of pyruvate by *Ap*PDC compared to *Zm*PDC, *Sc*PDC and *Zp*PDC. Data from Raj et al. [38] were measured in sodium citrate buffer at 25°C; *Zm*PDC data in Mes/KOH-buffer, pH 6.5, at 30°C [17], while data obtained in this work were measured in potassium phosphate buffer, pH 6.5, (*Zp*PDC: pH 7.0) at 30°C.

	*Ap*PDC	*Zm*PDC	*Sc*PDC	*Zp*PDC
V_{max} [U/mg]	110 ± 1.9 97 (pH 5.0) [38] 79 (pH 7.0) [38]	120-150 [17]	112.0 ± 35.6	116 ± 2.0 130 (pH 6.0) [38] 140 (pH 7.0) [38]
K_M [mM]	2.8 ± 0.2 0.35 (pH 5.0) [38] 5.1 (pH 7.0) [38]	1.1 ± 0.1 [17]	$S_{0.5} = 21.6 \pm 7.4$	2.5 ± 0.2 0.24 (pH 6.0) [38] 0.71 (pH 7.0) [38]

Tab. 5. Temperature stability of *Ap*PDC compared to *Sc*PDC and *Zp*PDC. Residual activity was measured using the coupled decarboxylase assay with pyruvate; - = not determined.

T (°C)	half-life $t_{1/2}$ [h] *Ap*PDC pH 6.5	half-life $t_{1/2}$ [h] *Sc*PDC pH 6.5	half-life $t_{1/2}$ [h] *Zp*PDC pH 7
20	193	235	-
25	-	243	-
30	144	78	150
35	-	62	-
40	34	-	40
50	12	-	10
60	2	-	0.4
70	0.4	-	-

Tab. 6. Mixed carboligations of different aliphatic aldehydes with benzaldehyde catalysed by *Ap*PDC. Reaction conditions: 50 mM potassium phosphate buffer, pH 7.0, containing 2.5 mM MgSO$_4$, 0.1 mM ThDP and 20% DMSO, 30 °C. Product detection was done by NMR after 3 d, the *ee* was determined by chiral phase HPLC.

no.	donor	acceptor	product	ee
1	17	16	20	91% (R)
2	18	16	21	82% (R)
3	19	16	22	81% (R)

Tab. 7. Residues lining the substrate channel and the active site of various ThDP-dependent enzymes. Corresponding residues are found in similar structural positions, if not otherwise indicated. For the positions of the respective areas see the schematic presentation (Fig. 6). Residues in parenthesis are situated in a similar position but are only partially involved in lining the respective site.

BAL	BFD	*Ap*PDC	*Zm*PDC	KdcA	*Sc*PDC	remarks
			pdb codes			
2ag0	1mcz, 1bfd	2vbi	1zpd	2vbf	1pyd	

Residues lining the *S*-pocket (area A1-Fig. 6 I+II)

BAL	BFD	*Ap*PDC	*Zm*PDC	KdcA	*Sc*PDC	remarks
L25	(N23)	V24	V24	V23	L25	
H26	P24	A25	A25	P24	P26	
G27	G25	G26	G26	G25	G27	
A28	S26	D27	D27	D26	D28	
F484	F464	I472	I476	I465	I480	
(W487)	(Y458)	Y466	Y470	Y459	Y474	
-	L461	E469	E473	E462	E477	

Entrance to the *S*-pocket (area A2-Fig. 6 I + II)

BAL	BFD	*Ap*PDC	*Zm*PDC	KdcA	*Sc*PDC	remarks
A480	A460	I468	I472	V461	I476	

Further important residues lining the donor binding site (area B-Fig. 6 I)

BAL	BFD	*Ap*PDC	*Zm*PDC	KdcA	*Sc*PDC	remarks
L112	L109	H113	H113	H112	H114	backbone-displacement, BFDL109/L110 correspond to H113/H114/Y290 in *Ap*PDC
Q113	L110	H114	H114	H113	H115	
H29	H70					proton relay system; in PDCs H113, H114, respectively H115 function as proton rely system
Q113						
(L398)	(T380)	W388	W392	F381	A392	residue defining the size of the binding pocket for 2-keto acid and donor aldehyde
H415	C398					
(C414)	F397					
Y397						
M421	L403	I411	I415	I404	I415	stabilising the V-conformation of ThDP

C-terminal α-helix covering the entrance to the substrate channel (area C-Fig. 6 I)

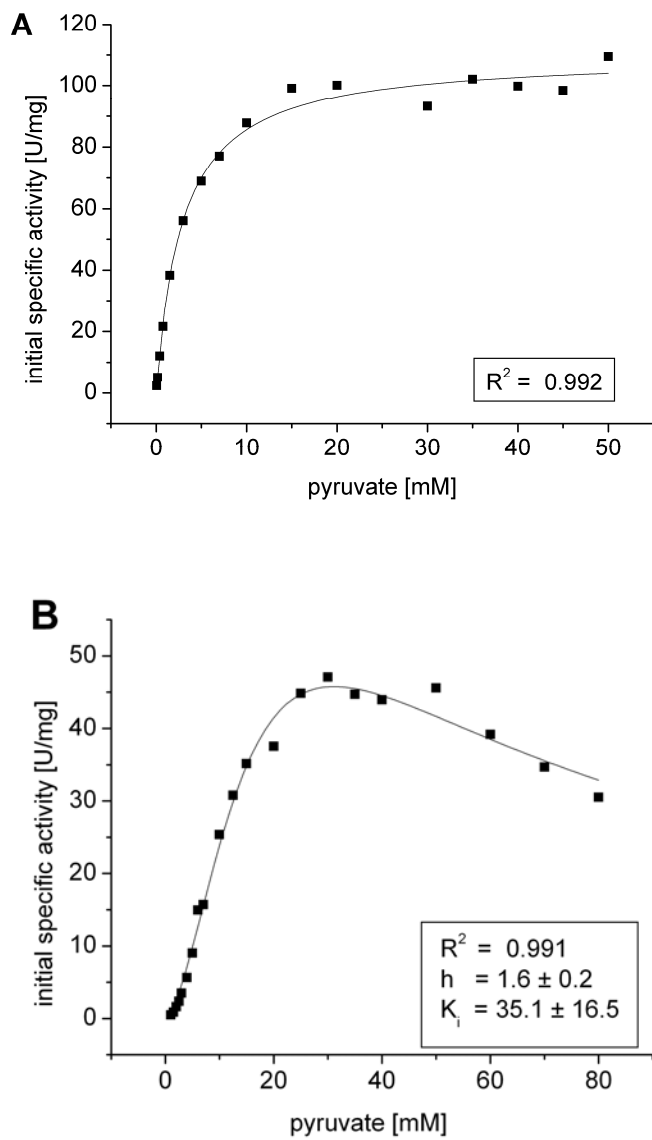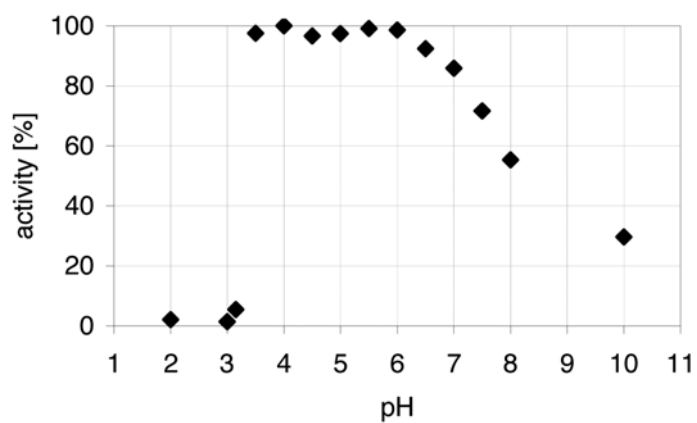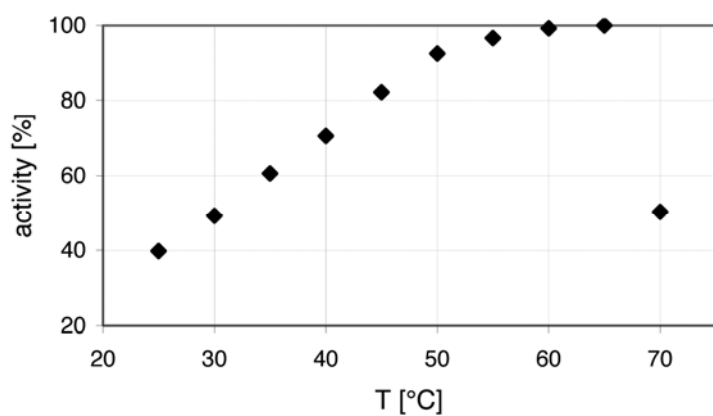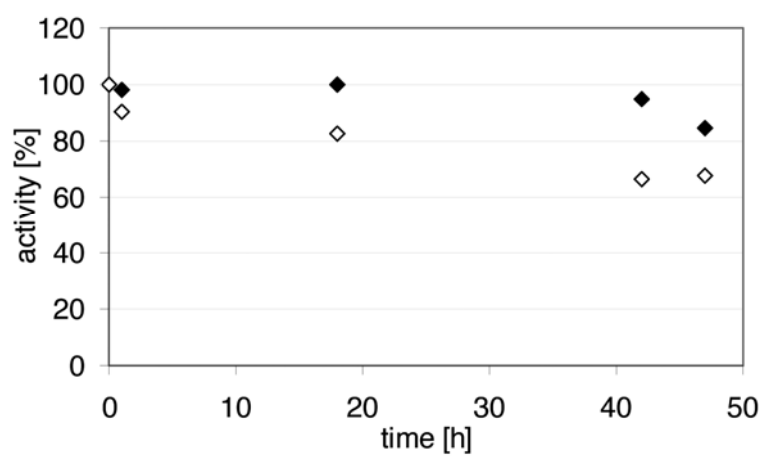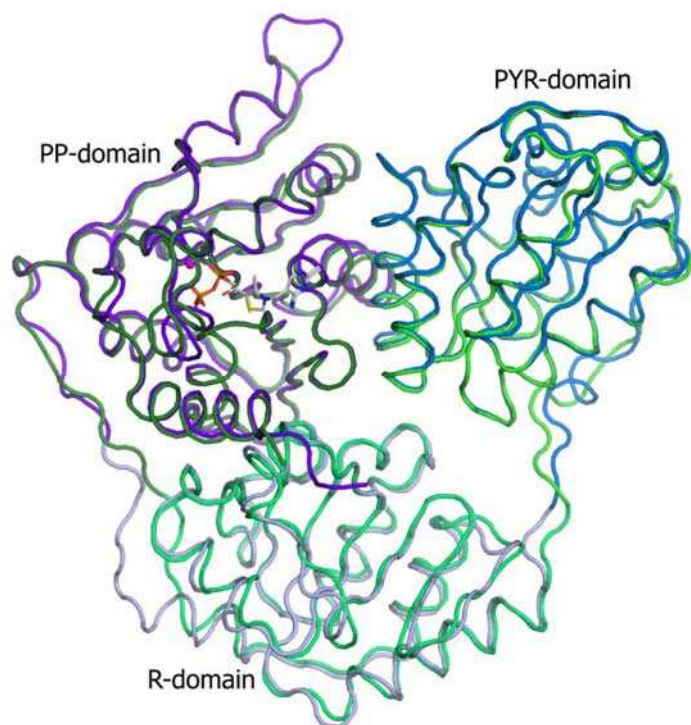| BAL	BFD	*Ap*PDC	*Zm*PDC	KdcA	*Sc*PDC	remarks
aa 550-555 α-helix, 1.5 turns	no helix	aa 538-555 α-helix, 4 turns	aa 544-566 α-helix, > 4 turns	aa 532-547 α-helix, 4 turns	aa 547-556 α-helix, 2-3 turns	

Fig. 1

Fig. 2

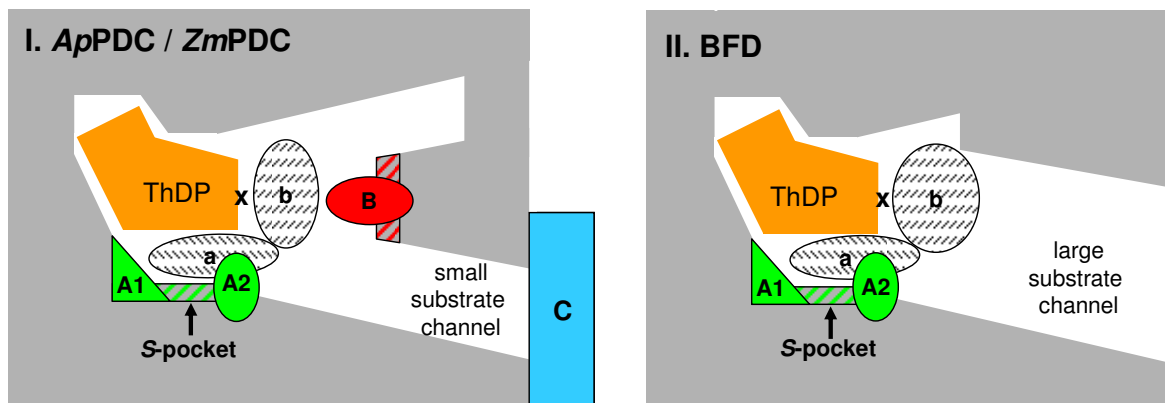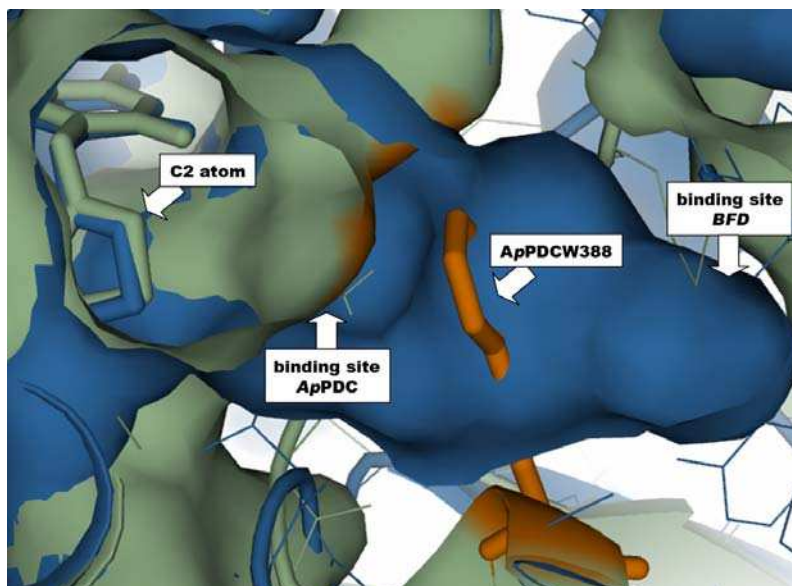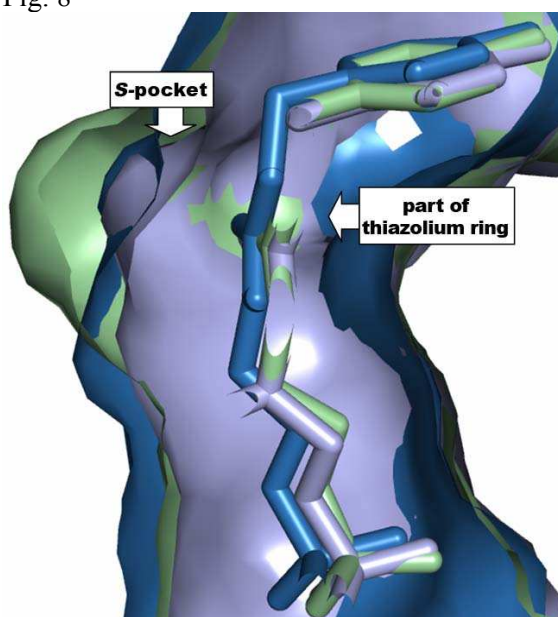Fig. 3

Fig. 4

Fig.5

Fig. 6

Fig. 7

Fig. 8

Supplementary Material

Pyruvate Decarboxylase from *Acetobacter pasteurianus*: BIOCHEMICAL AND STRUCTURAL CHARACTERISATION

Dörte Gocke, Catrine L. Berthold, Thorsten Graf, Helen Brosi, Ilona Frindi-Wosch, Michael Knoll, Thomas Stillger, Lydia Walter, Michael Müller, Jürgen Pleiss, Gunter Schneider, Martina Pohl

1. Alignment of the *Ap*PDC protein sequence in this work and the published sequence at NCBI.

Differences are marked in yellow.

ApPDC_Pohl   = *Ap*PDC sequence cloned by Gocke *et al.*

ApPDC_NCBI = *Ap*PDC sequence cloned by Raj *et al.* (1), accession code at NCBI: AAM21208
→ 98 % identity (9 different amino acids, 1 gap)

```
ApPDC_Pohl      MTYTVGMYLAERLVQIGLKHHFAVAGDYNLVLLDQLLLNKDMKQIYCCNELNCGFSAEGY 60
ApPDC_NCBI      MTYTVGMYLAERLVQIGLKHHFAVGGDYNLVLLDQLLLNKDMKQIYCCNELNCGFSAEGY 60
                ************************.*+**********************************

ApPDC_Pohl      ARSNGAAAAVVTFSVGAISAMNALGGAYAENLPVILISGAPNSNDQGTGHILHHTIGKTD 120
ApPDC_NCBI      ARSNGAAAAVVTFSVGAISAMNALGGAYAENLPVILISGAPNSNDQGTGHILHHTIGKTD 120
                ************************************************************

ApPDC_Pohl      YSYQLEMARQVTCAAESITDAHSAPAKIDHVIRTALRERKPAYLDIACNIASEPCVRPGP 180
ApPDC_NCBI      YSYQLEMARQVTCAAESITDAHSAPAKIDHVIRTALRERKPAYLDIACNIASEPCVRPGP 180
                ************************************************************

ApPDC_Pohl      VSSLLSEPEIDHTSLKAAVDATVALLEKSASPVMLLGSKLRAANALAATETLADKLQCAV 240
ApPDC_NCBI      VSSLLSEPEIDHTSLKAAVDATVALLKNRPAPVMLLGSKLRAANALAATETLADKLQCAV 240
                *************************::  .:*****************************

ApPDC_Pohl      TIMAAAKGFFPEDHAGFRGLYWGEVSNPGVQELVETSDALLCIAPVFNDYSTVGWSAWPK 300
ApPDC_NCBI      TIMAAAKGFFPEDHAGFRGLYWGEVSNPGVQELVETSDALLCIAPVFNDYSTVGWSGMPK 300
                *******************************************************.  **

ApPDC_Pohl      GPNVILAEPDRVTVDGRAYDGFTLRAFLQALAEKAPARPASAQKSSVPTCSLTATSDEAG 360
ApPDC_NCBI      GPNVILAEPDRVTVDGRAYDGFTLRAFLQALAEKAPARPASAQKSSVPTCSLTATSDEAG 360
                ************************************************************

ApPDC_Pohl      LTNDEIVRHINALLTSNTTLVAETGDSWFNAMRMTLPRGARVELEMQWGHIGWSVPSAFG 420
ApPDC_NCBI      LTNDEIVRHINALLTSNTTLVAETGDSWFNAMRMTLA-GARVELEMQWGHIGWSVPSAFG 419
                ************************************.  *********************

ApPDC_Pohl      NAMGSQDRQHVVMVGDGSFQLTAQEVAQMVRYELPVIIFLINNRGYVIEIAIHDGPYNYI 480
ApPDC_NCBI      NAMGSQDRQHVVMVGDGSFQLTAQEVAQMVRYELPVIIFLINNRGYVIEIAIHDGPYNYI 479
                ************************************************************

ApPDC_Pohl      KNWDYAGLMEVFNAGEGHGLGLKATTPKELTEAIARAKANTRGPTLIECQIDRTDCTDML 540
ApPDC_NCBI      KNWDYAGLMEVFNAGEGHGLGLKATTPKELTEAIARAKANTRGPTLIECQIDRTDCTDML 539
                ************************************************************

ApPDC_Pohl      VQWGRKVASTNARKTTLA 558
ApPDC_NCBI      VQWGRKVASTNARKTTLA 557
                ******************
```

2. Alignment of *Ap*PDC and pyruvate decarboxylase from *Zymomonas mobilis* (*Zm*PDC)
ZmPDC_NCBI: *Zm*PDC sequence cloned by Pohl *et al.* (2)
ApPDC_Pohl: *Ap*PDC sequence cloned by Gocke *et al.*
➡ *Zm*PDC used for molecular replacement, homology 62 %

```
ZmPDC_NCBI    MSYTVGTYLAERLVQIGLKHHFAVAGDYNLVLLDNLLLNKNMEQVYCCNELNCGFSAEGY 60
ApPDC_Pohl    MTYTVGMYLAERLVQIGLKHHFAVAGDYNLVLLDQLLLNKDMKQIYCCNELNCGFSAEGY 60
              *:**** *****************************:*****:*:*:**************

ZmPDC_NCBI    ARAKGAAAAVVTYSVGALSAFDAIGGAYAENLPVILISGAPNNNDHAAGHVLHHALGKTD 120
ApPDC_Pohl    ARSNGAAAAVVTFSVGAISAMNALGGAYAENLPVILISGAPNSNDQGTGHILHHTIGKTD 120
              **::*********:****:**::*:******************.**:.:**:***::****

ZmPDC_NCBI    YHYQLEMAKNITAAAEAIYTPEEAPAKIDHVIKTALREKKPVYLEIACNIASMPCAAPGP 180
ApPDC_Pohl    YSYQLEMARQVTCAAESITDAHSAPAKIDHVIRTALRERKPAYLDIACNIASEPCVRPGP 180
              * ******::*.***:*   ...*********:*****:**.**:******* **. ***

ZmPDC_NCBI    ASALFNDEASDEASLNAAVEETLKFIANRDKVAVLVGSKLRAAGAEEAAVKFADALGGAV 240
ApPDC_Pohl    VSSLLSEPEIDHTSLKAAVDATVALLEKSASPVMLLGSKLRAANALAATETLADKLQCAV 240
              .*:*::.    *.:**:***: *: :: :   . .:*:*******.* *: .:** * **

ZmPDC_NCBI    ATMAAAKSFFPEENPHYIGTSWGEVSYPGVEKTMKEADAVIALAPVFNDYSTTGWTDIPD 300
ApPDC_Pohl    TIMAAAKGFFPEDHAGFRGLYWGEVSNPGVQELVETSDALLCIAPVFNDYSTVGWSAWPK 300
              : *****.****::. : *  ***** ***:: :: :**::.:*********.**:.  *.

ZmPDC_NCBI    PKKLVLAEPRSVVVNGIRFPSVHLKDYLTRLAQKVSKKTGALDFFKSLNAGELKKAAPAD 360
ApPDC_Pohl    GPNVILAEPDRVTVDGRAYDGFTLRAFLQALAEKAPARP-ASAQKSSVPTCSLTATS--- 356
              ::::****   *.*:*   : .. *: :*   **:*.. :. *     .*: : .*. ::

ZmPDC_NCBI    PSAPLVNAEIARQVEALLTPNTTVIAETGDSWFNAQRMKLPNGARVEYEMQWGHIGWSVP 420
ApPDC_Pohl    DEAGLTNDEIVRHINALLTSNTTLVAETGDSWFNAMRMTLPRGARVELEMQWGHIGWSVP 416
              .* *.* **.*::***.***:*********** **.** ***** ************

ZmPDC_NCBI    AAFGYAVGAPERRNILMVGDGSFQLTAQEVAQMVRLKLPVIIFLINNYGYTIEVMIHDGP 480
ApPDC_Pohl    SAFGNAMGSQDRQHVVMVGDGSFQLTAQEVAQMVRYELPVIIFLINNRGYVIEIAIHDGP 476
              :*** *:*: :*:::.:****************** :********** **.**: *****

ZmPDC_NCBI    YNNIKNWDYAGLMEVFNGNGGYDSGAGKGLKAKTGGELAEAIKVALANTDGPTLIECFIG 540
ApPDC_Pohl    YNYIKNWDYAGLMEVFNAG----EGHGLGLKATTPKELTEAIARAKANTRGPTLIECQID 532
              ** **************..     .* * ****.* **:*** * *** ******* *.

ZmPDC_NCBI    REDCTEELVKWGKRVAAANSRKPVNK 566
ApPDC_Pohl    RTDCTDMLVQWGRKVASTNARKTTLA 558
              * ***: **:**::**::*:**..
```

3. Alignment of our *Zp*PDC protein sequence and the published sequence at NCBI
ZpPDC_Pohl = *Zp*PDC sequence cloned by Gocke *et al.*
ZpPDC_NCBI = *Zp*PDC sequence cloned by Raj *et al.* (3)
→ 99 % identity (2 different amino acids, 1 insertion)

```
ZpPDC_Pohl    MGYTVGMYLAERLAQIGLKHHFAVAGDYNLVLLDQLLLNKDMEQVYCCNELNCGFSAEGY 60
ZpPDC_NCBI    -MYTVGMYLAERLAQIGLKHHFAVAGDYNLVLLDQLLLNKDMEQVYCCNELNCGFSAEGY 59
               ********************************************************

ZpPDC_Pohl    ARARGAAAAIVTFSVGAISAMNAIGGAYAENLPVILISGSPNTNDYGTGHILHHTIGTTD 120
ZpPDC_NCBI    ARARGAAAAIVTFSVGAISAMNAIGGAYAENLPVILISGSPNTNDYGTGHILHHTIGTTD 119
              ************************************************************

ZpPDC_Pohl    YNYQLEMVKHVTCAAESIVSAEEAPAKIDHVIRTALRERKPAYLEIACNVAGAECVRPGP 180
ZpPDC_NCBI    YNYQLEMVKHVTCARESIVSAEEAPAKIDHVIRTALRERKPAYLEIACNVAGAECVRPGP 179
              **************.*********************************************

ZpPDC_Pohl    INSLLRELEVDQTSVTAAVDAAVEWLQDRQNVVMLVGSKLRAAAAEKQAVALADRLGCAV 240
ZpPDC_NCBI    INSLLRELEVDQTSVTAAVDAAVEWLQDRQNVVMLVGSKLRAAAAEKQAVALADRLGCAV 239
              ************************************************************
```

```
ZpPDC_Pohl     TIMAAAKGFFPEDHPNFRGLYWGEVSSEGAQELVENADAILCLAPVFNDYATVGWNSWPK 300
ZpPDC_NCBI     TIMAAEKGFFPEDHPNFRGLYWGEVSSEGAQELVENADAILCLAPVFNDYATVGWNSWPK 299
               *****  *****************************************************

ZpPDC_Pohl     GDNVMVMDTDRVTFAGQSFEGLSLSTFAAALAEKAPSRPATTQGTQAPVLGIEAAEPNAP 360
ZpPDC_NCBI     GDNVMVMDTDRVTFAGQSFEGLSLSTFAAALAEKAPSRPATTQGTQAPVLGIEAAEPNAP 359
               ***********************************************************

ZpPDC_Pohl     LTNDEMTRQIQSLITSDTTLTAETGDSWFNASRMPIPGGARVELEMQWGHIGWSVPSAFG 420
ZpPDC_NCBI     LTNDEMTRQIQSLITSDTTLTAETGDSWFNASRMPIPGGARVELEMQWGHIGWSVPSAFG 419
               ***********************************************************

ZpPDC_Pohl     NAVGSPERRHIMMVGDGSFQLTAQEVAQMIRYEIPVIIFLINNRGYVIEIAIHDGPYNYI 480
ZpPDC_NCBI     NAVGSPERRHIMMVGDGSFQLTAQEVAQMIRYEIPVIIFLINNRGYVIEIAIHDGPYNYI 479
               ***********************************************************

ZpPDC_Pohl     KNWNYAGLIDVFNDEDGHGLGLKASTGAELEGAIKKALDNRRGPTLIECNIAQDDCTETL 540
ZpPDC_NCBI     KNWNYAGLIDVFNDEDGHGLGLKASTGAELEGAIKKALDNRRGPTLIECNIAQDDCTETL 539
               ***********************************************************

ZpPDC_Pohl     IAWGKRVAATNSRKPQALEHHHHHH 565
ZpPDC_NCBI     IAWGKRVAATNSRKPQA-------- 556
               *****************
```

4. Calibration G200 size exclusion chromatography column
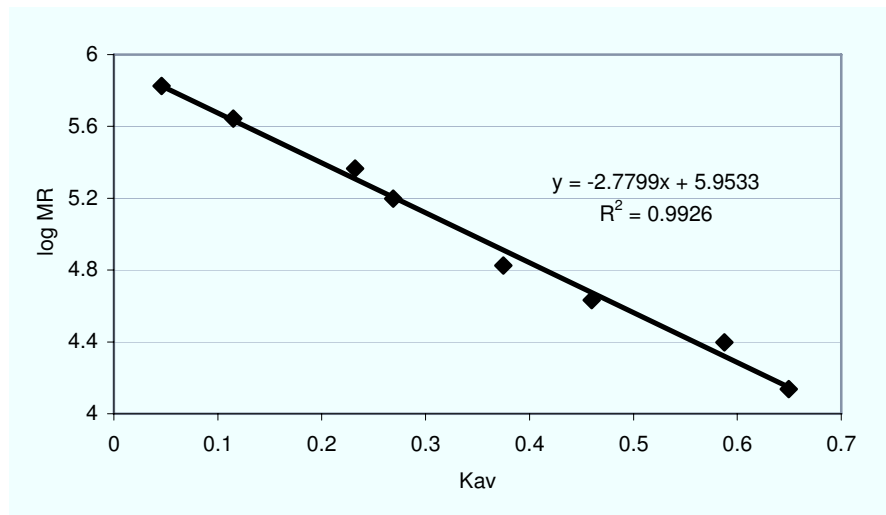
FIGURE 1. Calibration G200 size exclusion chromatography column.
Buffer: 50 mM potassium phosphate buffer, 0.1 mM ThDP, 2.5 mM MgSO$_4$, 150 mM KCl, pH 6.5.

For *Ap*PDC log M$_R$ = 5.35.

5. Hill-equation for determination of kinetic constants of *Sc*PDC

$$V = \frac{V_{max} \cdot [S]^h}{S_{0.5}{}^h + S^h \left(1 + \dfrac{S}{K_i}\right)}$$

6. Activation energy for the *Ap*PDC catalysed decarboxylation of pyruvate
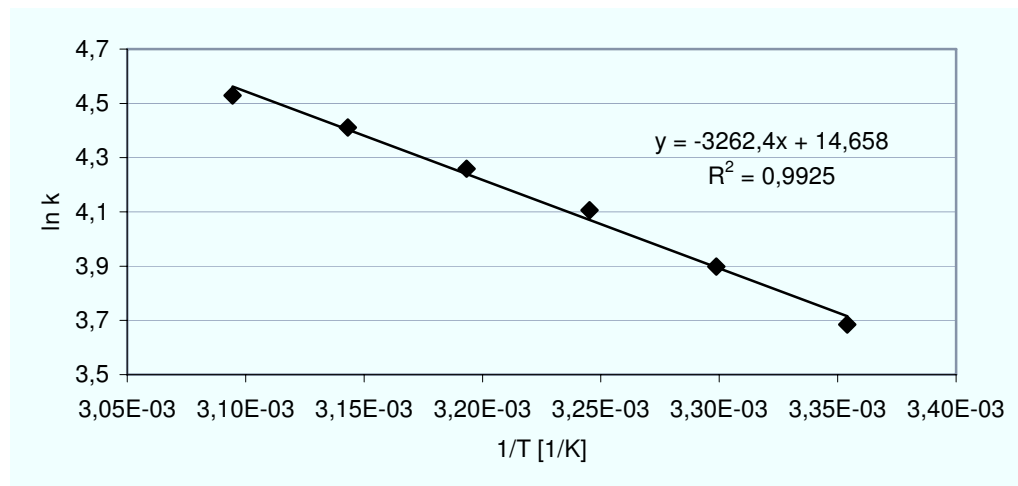
FIGURE 2. $\ln V_{max}$/[1/T]-plot to determine the activation energy of the decarboxylase reaction. Values used in the linear area of 20-50°C only.

REFERENCES

1. Raj, K. C., Ingram, L. O., Maupin-Furlow, J. A. (2001) *Arch. Microbiol.* **176**, 443-451
2. Pohl, M., Mesch, K., Rodenbrock, A., Kula, M.R. (1995) *Biotechnol. Appl. Biochem.* **22**, 95-105
3. Raj, K. C., Talarico, L. A., Ingram, L. O., Maupin-Furlow, J. A. (2002) *Appl. Environ. Microbiol.* **68**, 2869-2876

Publication V

Branched-chain keto acid decarboxylase from *Lactococcus lactis* (KdcA), a valuable thiamine diphosphate-dependent enzyme for asymmetric C-C bond formation.

Gocke, D., Nguyen, C. L., Pohl, M.,

Stillger, T., Walter, L. & Müller, M.

Branched-chain ketoacid decarboxylase from *Lactococcus lactis* (KdcA), a valuable thiamin diphosphate-dependent enzyme for asymmetric C-C bond formation

Dörte Gocke[a], Cong Luan Nguyen[a], Martina Pohl[a]*, Thomas Stillger[b], Lydia Walter[b], Michael Müller[b]

Both groups contributed equally to the publication:
[a] Institute of Molecular Enzyme Technology, Heinrich-Heine University Duesseldorf, Research Center Jülich, 52426 Jülich (Germany), Fax: (+49)2461-612490, *corresponding author: ma.pohl@fz-juelich.de
[b] Institute of Pharmaceutical Sciences, Albert-Ludwigs-University Freiburg, Albertstr. 25
79104 Freiburg (Germany), Fax: (+49)761-2036351

Keywords: Bioorganic chemistry, biotransformation, C-C coupling, pyruvate decarboxylase, benzoylformate decarboxylase, benzaldehyde lyase.

Abstract

The thiamin diphosphate-dependent branched-chain 2-ketoacid decarboxylase from *Lactococcus lactis* sup. cremoris B1157 (KdcA) is a new valuable enzyme for the synthesis of chiral 2-hydroxy ketones. The gene was cloned and the enzyme was expressed as an N-terminal hexahistidine fusion protein in *E. coli*. It has a broad substrate range for the decarboxylation reaction including linear and branched-chain aliphatic and aromatic ketoacids as well as phenylpyruvate and indole-3-pyruvate. The dimeric structure of recombinant KdcA is in contrast to the tetrameric structure of other 2-ketoacid decarboxylases. The enzyme is stable between pH 5-7 with a pH optimum of pH 6-7 for the decarboxylation reaction. While KdcA is sufficiently stable up to 40 °C it rapidly looses activity at higher temperatures.
In this work the carboligase activity of KdcA is demonstrated for the first time. The enzyme shows an exceptional broad substrate range and, most strikingly, it catalyzes the carboligation of different aromatic aldehydes as well as CH-acidic aldehydes such as phenylacetaldehyde and indole-3-acetaldehyd with aliphatic aldehydes such as acetaldehyde, propanal, cyclopropylcarbaldehyde, yielding chiral 2-hydroxy ketones in high enantiomeric excess. Noteworthy, the donor-acceptor selectivity is strongly influenced by the nature of the respective substrate combination.

Introduction

Thiamin diphosphate-dependent enzymes such as pyruvate decarboxylase (PDC, E.C. 4.1.1.1), benzoylformate decarboxylase (BFD, E.C. 4.1.1.7) and benzaldehyde lyase (BAL, E.C. 4.1.2.38) have been intensively studied with respect to their carboligation activity, offering an easy access to chiral 2-hydroxyketones from aldehydes.[1-8] In order to enlarge the range of accessible 2-hydroxyketones, we studied the carboligase properties of a recently described branched-chain keto acid decarboxylase (E.C. 4.1.1.72). Two highly homologous enzymes have been found in different *Lactococcus lactis* strains: *Lactococcus lactis* sup. cremoris B1157[9, 10] (KdcA) and *Lactococcus lactis* IFPL730[11] (Kivd). These enzymes are involved in the process of cheese ripening due to their decarboxylation activity of 2-ketoacids which are formed through transamination of the corresponding branched-chain amino acids.[9] Recently the substrate binding site of KdcA has been probed by site-directed mutagenesis studies based on a homology model using the structure of pyruvate decarboxylase from *Zymomonas mobilis* as a template.[12]

In this study we investigated the carboligase properties of KdcA for the first time. Compared to other decarboxylases KdcA has a broader substrate range and accepts besides acetaldehyde several aliphatic aldehydes such as propanal, butanal, isobutyraldehyde and cyclopropylcarbaldehyde as donor and/or acceptor in the enzyme-catalyzed acyloin condensation. Moreover, KdcA catalyzes the carboligation of enolizable CH-acidic aldehydes such as indole-3-acetaldehyde and phenylacetaldehyde. The substrate range of the decarboxylase and the carboligase reaction, the pH and temperature dependent stability and activity, as well as the stereoselectivity of KdcA are reported.

Results and Discussion

1. Cloning, overexpression and purification

The coding gene (kdca, gene bank CAG34226) was cloned and overexpressed in *E. coli* as an N-terminal hexahistidine fusion protein. Addition of the His-tag led to an N-terminal elongation of the protein by 23 amino acids: MGSS**HHHHHH**SSGLVPRGSHMAS.

A 15 L fed-batch fermentation of the recombinant *E. coli* strain yielded in 720 kU (decarboxylase activity) KdcA in 1.2 kg cells. The enzyme was purified by immobilized metal chelate chromatography yielding about 1.7 g KdcA from 1.2 kg cells.

2. Determination of the native molecular weight
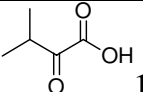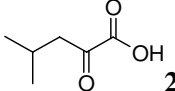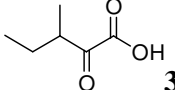
The native molecular weight of the recombinant KdcA was determined by size-exclusion chromatography to a molecular weight of about 146 kDa (for logMr/Kav-plot see supplementary material). As the calculated size of the monomeric subunit is 63,337 kDa, the observed data correlate best with a dimeric structure of KdcA in the native state. This is in contrast to other tetrameric 2-ketoacid decarboxylases, such as PDC[13,14] and BFD.[15] However, other ThDP-dependent enzymes are known to be active as a dimer, such as acetohydroxyacid synthase (AHAS)[16] and transketolase (TK).[17] Meanwhile the three-dimensional structure of KdcA has been determined proving the dimeric structure of the enzyme (unpublished results).
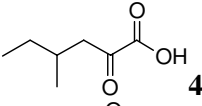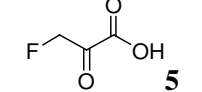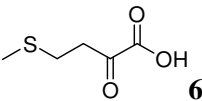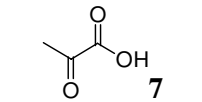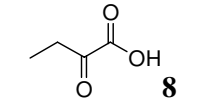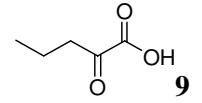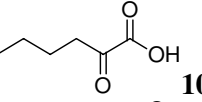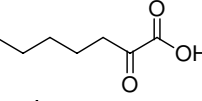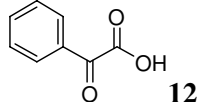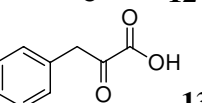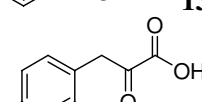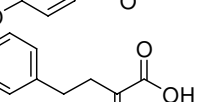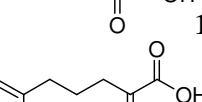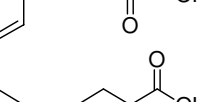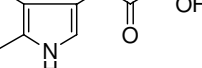
3. Decarboxylase activity

The investigation of the substrate range of decarboxylation is of interest to deduce information about the acyldonor spectrum for carboligase activity of KdcA. If a respective 2-ketoacid is a substrate for the enzyme the binding of the corresponding aldehyde to the C2-atom of ThDP located in the active center is most likely. In the case of carboligation the ThDP-bound aldehyde is ligated to a second acceptor aldehyde molecule. Therefore, the aldehydes which are products of the decarboxylation reaction can in principle be used at least as acyldonors for the carboligation reaction.

KdcA has an exceptionally broad substrate range for the decarboxylation reaction as determined by Smit *et al*.[10] and Yep *et al*.[12] (Tab. 1). In contrast to other decarboxylases like PDC or BFD KdcA accepts different aliphatic branched-chain as well as enolizable 2-ketoacids such as phenylpyruvate and indole-3-pyruvate. Although the natural branched-chain aliphatic substrate **1** as well as its analogue **2** (Tab. 1) show the highest reaction velocities, the K_M-values are significantly higher than those of linear aliphatic 2-ketoacids, phenylpyruvate and indole-3-pyruvate.

Table 1: Substrate range and kinetic data of the decarboxylation reaction. Data are compared with the literature. Smit et al. used crude extracts of KdcA without any tag. Yep et al. used a purified variant with a C-terminal hexahistidine tag and own data were obtained with a purified variant containing an N-terminal hexahistidine tag. Conditions: Activity towards different 2-ketoacids was measured using the coupled continuous assay. All substrates were applied in a concentration of 30 mM, except **17**, which was applied with 1 mM. Relative activities for **2**, **13**, and **17** have been calculated based on the maximal velocities.

2-Ketoacid	Relative activity (%) own data compared to [Smit et al.][10] (Yep et al.)[12]	Kinetic data own data (Yep et al.)[12]	
		V_{max} [U/mg]	K_M [mM]
1	100 [100] (100)	181.64±0.21 (47.3)	5.02±0.21 (2.8)
2	26.3 [31] (100)	34.3±0.2 (48.2)	0.264±0.011 (3.7)
3	38.1 [27.8]	n.d.	n.d.

Compound			
4	19	n.d.	n.d.
5	0	n.d.	n.d.
6	(18)	n.d. (8.6)	n.d. (1.3)
7	1.2 [1.3]	3.67±0.31	29.77±6.44
8	9.3 [7.3]	n.d.	n.d.
9	15 [19] (20.6)	n.d. (9.8)	n.d. (1.3)
10	13 [25.3] (26.5)	n.d. (12.8)	n.d. (0.6)
11	0	n.d.	n.d.
12	8.4 (15.2)	n.d. (7.2)	n.d. (7.5)
13	8.6 [7] (56.3)	15.69±0.21 (26.6)	0.127±0.007 0.21
14	(6)	n.d. (2.9)	n.d. (0.63)
15	1.6	n.d.	n.d.
16	1.1	n.d.	n.d.
17	0.85 [3.5]	1.55±0.6	0.234±0.024

KdcA has an exceptionally broad substrate range for the decarboxylation reaction as determined by Smit et al.[10] and Yep et al.[12] (Tab. 1). In contrast to other decarboxylases like PDC or BFD KdcA accepts different aliphatic branched-chain as well as enolizable 2-ketoacids such as phenylpyruvate and indole-3-pyruvate. Although the natural branched-chain aliphatic substrate **1** as well as its analogue **2** (Tab. 1) show the highest reaction velocities, the K_M-values are significantly higher than those of linear aliphatic 2-ketoacids, phenylpyruvate and indole-3-pyruvate.

Kinetic constants were determined for the physiological substrate 3-methyl-2-oxobutanoic acid (**1**), 4-methyl-2-oxopentanoic acid (**2**) as well as phenylpyruvate (**13**) and indole-3-pyruvate (**15**) (Tab. 1). With substrates **1** and **2** hyperbolic v/[S]-plots have been observed up to 50 mM. With phenylpyruvate

the v/[S]-plot is hyperbolic up to 12 mM (Fig. 1a) followed by a progressive decay of activity up to 30 mM and a subsequent rapid complete loss of activity (not shown). In accordance with the observation of Yep et al.[12] KdcA exhibits an extremely low K_M-value for phenylpyruvate (0.13 mM), which is about 40-times lower compared to the natural substrate **1**. However, the maximal velocity in the presence of phenylpyruvate is only about 9% of the velocity obtained with **1**. Our data correlate well with those obtained by Smit et al.[10] who investigated KdcA without any hexahistidine tag (Tab. 1). However, we could not reproduce the high relative activity of 56.3% compared to the natural substrate **1** which was found by Yep et al. with a KdcA-variant carrying a C-terminal hexahistidine tag (Tab. 1). In the case of indole-3-pyruvate kinetic investigation has been hampered by the low solubility (about 1.3 mM) and the strong absorbance of the substrate in aqueous buffer. However, as the K_M-value is also very low (0.23 mM) saturation is nearly achieved at 1 mM (Fig. 1b). Compared to 3-methyl-2-oxobutanoic acid (**1**) the KdcA shows 0.85% relative activity with indole-3-pyruvate.
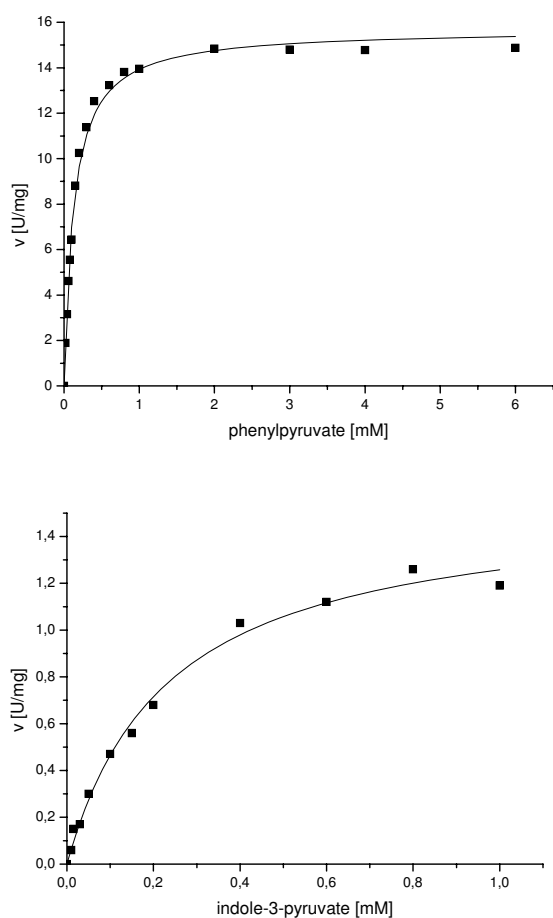
Figure 1: v/[S]-plot of the KdcA-catalyzed decarboxylation of (A) phenylpyruvate (**13**) and (B) indole-3-pyruvate (**17**). Buffer: 50 mM potassium phosphate, 2.5 mM MgSO₄, 0.1 mM ThDP, pH 6.8, 30 °C. Data have been obtained in triplicates using the coupled decarboxylase assay.

4. Optimal cofactor concentration

As in all ThDP-dependent enzymes the cofactors in KdcA are bound non-covalently to the active site. The main contribution to the binding arise from the coordinative interaction of the diphosphate moiety via Mg^{2+} to the protein as well as from hydrophobic and ionic interactions between protein site chains and the thiazol- and pyrimidine ring of ThDP.

For stability of the holoenzyme most ThDP-dependent enzymes require the addition of cofactors to the buffer. In case of KdcA the addition of 2.5 mM MgSO₄ and 0.1 mM ThDP to the buffer is sufficient to keep the enzyme stable and active.

5. pH-dependent activity and stability

KdcA shows a pH-optimum of pH 6-7 for the decarboxylation of **1** (Fig. 2). This activity optimum overlaps very well with the stability optimum of KdcA in potassium phosphate buffer, where no loss of activity can be detected at pH 5-7 within 60 h. Rapid inactivation occurs at pH 4 (< 2 h) and slower inactivation at pH 8 (half-life: 40 h).
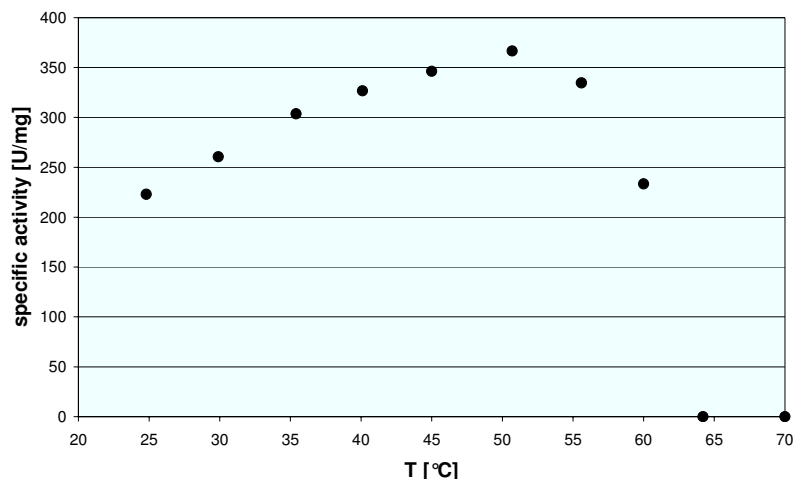

Figure 2: Determination of the temperature optimum and the midpoint of thermal inactivation of the KdcA-catalyzed decarboxylation of 3-methyl-2-oxobutanoic (**1**). Data were obtained using the direct decarboxylase assay.

6. Temperature optimum for activity and stability

Under initial rate conditions (90 sec) the temperature optimum was observed at 50 °C. Further increase in temperature resulted in a fast decay of activity with a midpoint of thermal inactivation (Tm) at about 62 °C (Fig. 2).
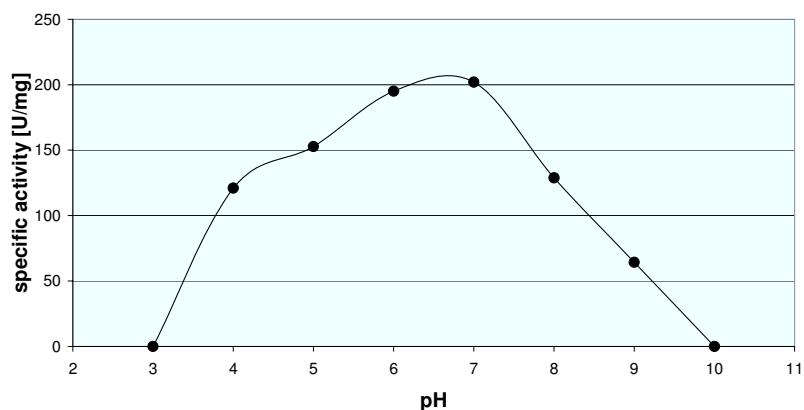
Figure 3: Determination of the pH-optimum of the KdcA-catalyzed decarboxylation of 3-methyl-2-oxo-butanoic acid (**1**). Data were obtained using the coupled decarboxylase assay.

From these data the activation energy of the decarboxylase reaction was calculated from a $\ln V_{max}/[1/T]$-plot in the range of 25-40 °C as 8.5 kJ/mol, which is relatively low compared to other 2-ketoacid decarboxylases. For benzoylformate decarboxylase from *Pseudomonas putida* and for pyruvate decarboxylase from *Zymomonas mobilis* an activation energy of 38 kJ/mol[18] and 43 kJ/mol, respectively, (unpublished) have been determined.

In order to optimize the conditions for the application of KdcA in enzymatic syntheses, the temperature stability of the enzyme in 50 mM potassium phosphate buffer, pH 6.8 in the presence of 2.5 mM $MgSO_4$ and 0.1 mM ThDP was determined. While the enzyme is sufficiently stable up to 40

°C (half-life: 80 h) it rapidly looses activity at higher temperatures (50 °C: half-life 9 h, 55 °C: half-life 4 h).

7. Organic solvents

The biotransformation of aromatic aldehydes is often hampered by their low solubility in aqueous systems. Since the addition of either 20% (v/v) DMSO or 15% (v/v) PEG-400 have been successfully applied for BAL and BFD catalyzed carboligase reactions,[1,3,19-20] these organic solvents were tested with KdcA. The enzyme is completely stable in the presence of 20% (v/v) DMSO (half-life: 150 h), whereas it is rapidly inactivated in the presence of 15% (v/v) PEG-400 (half-life: 6 h) (Fig. 4).
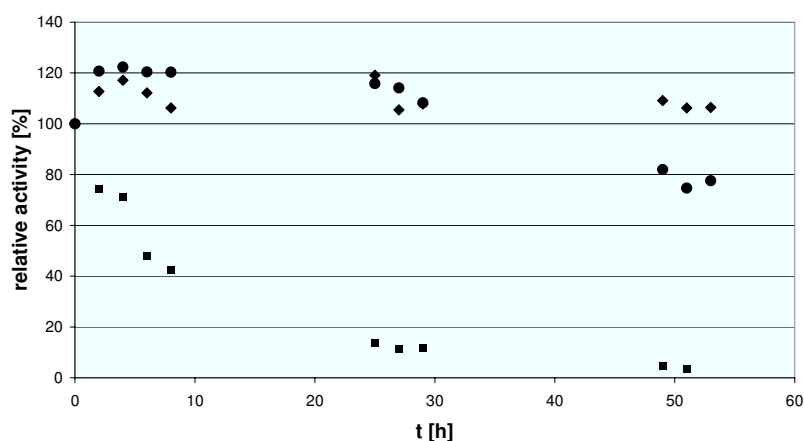
Figure 4: Stability of KdcA in 50 mM potassium phosphate buffer, pH 6.8, 2.5 mM MgSO$_4$, 0.1 mM ThDP (●), in the same buffer plus 20% (v/v) DMSO (♦) and in the same buffer plus 15% (v/v) PEG-400 (■). Activities were related to the starting activity in buffer without additives. The up to 20% higher activity in the case of (●) and (♦) after short time incubation is due to normal fluctuations in decarboxylase activity observed after solution of lyophilized enzyme.

8. Carboligase activity

8.1 Carboligation of aliphatic aldehydes

Since KdcA is involved in the production of branched-chain aliphatic aldehydes, isovaleraldehyde (**18**) and isobutyraldehyde (**19**) were tested at first for carboligation (Table 2, entry 1, 2). As summarized in Table 2 the enzyme catalyzes the carboligation of **18** with a low specific activity of 0.004 U/mg to give (*S*)-5-hydroxy-2,5-dimethyloctan-4-one (**27**) with an enantiomeric excess (ee) of 30-47%. BAL catalyzes this reaction with a specific activity of 0.05 U/mg resulting in the formation of (*R*)-**27** (ee 80%).[21,22] Thus, although the activity of KdcA is distinctly lower for this transformation it opens the way to both enantiomers of aliphatic acyloins.

Surprisingly, KdcA is not able to catalyze the carboligation of isobutyraldehyde (**19**) in detectable amounts, although **19** is the reaction product of the physiological decarboxylation of 3-methyl-2-oxobutanoic acid (**1**) and should therefore fit into the active center. We assumed an inhibition or inactivation of KdcA caused by isobutyraldehyde (**19**). In order to test this hypothesis the enzyme was incubated with increasing concentrations of **19** (5, 20, 40 mM) for 26 h and the decrease of activity of KdcA was followed by measuring the residual decarboxylase activity using the direct decarboxylase assay. The data presented in Fig. 5 show a concentration dependent inactivation of KdcA, which explains the negativ results of the carboligation studies with this aldehyde.
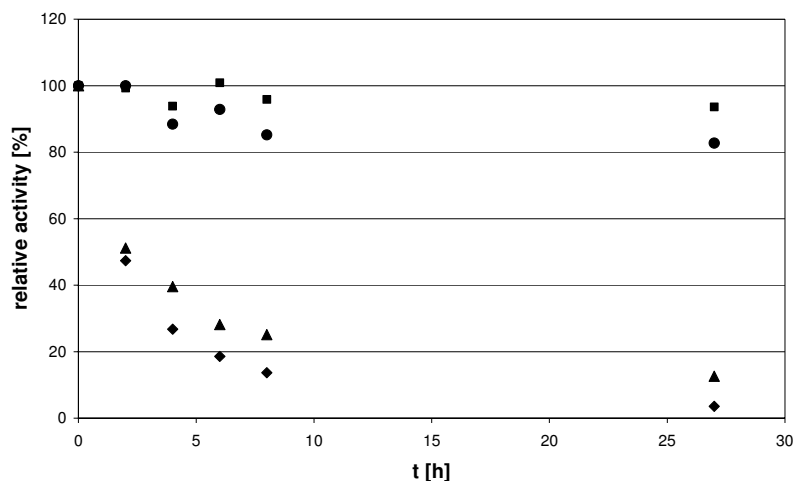
Figure. 5: Inactivation of KdcA by isobutyraldehyde (**19**) (■: 0 mM, ●: 5 mM, ▲: 20 mM ◆: 40 mM). After the indicated time intervals 50 µL samples were withdrawn and residual activity was measured by the direct decarboxylase assay.

To investigate the reversibility of the inactivation isobutyraldehyde (**19**) was removed from the 40 mM sample after complete inactivation of KdcA using ultra filtration. The regain of activity was followed by measuring the decarboxylase activity of the enzyme. After 2 hours 7% and after 16 hours 30% of the original activity were recovered (data not shown), demonstrating that the inactivation of the enzyme is at least partially reversible. Further, kinetic studies using the direct decarboxylase assay in the presence of 40 mM isobutyraldehyde resulted neither in a decay of V_{max} nor did isobutyraldehyde affect the K_M-value for 3-methyl-2-oxobutanoic acid (**1**) (data not shown). We therefore conclude that the observed inactivation in the presence of isobutyraldehyde is a slow process and not detectable under initial rate conditions.

Besides the branched-chain aldehydes other aldehydes such as acetaldehyde, propanal and cyclopropyl carbaldehyde were tested as substrates for the carboligase reaction. In all cases carboligase products were observed by means of GC/MS. The specific activity for acetoin formation (0.052 U/mg) is 13fold higher than the catalytic activity towards the formation of the branched-chain 2-hydroxyketone **27** (Tab. 2, entry 3). The (*R*)-enantiomer of acetoin (**28**) is formed predominantly (ee 46%) as proven by comparison with the product obtained from the PDC (*Saccharomyces cerevisiae*)-catalyzed reaction.
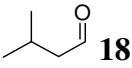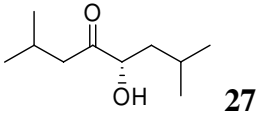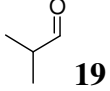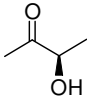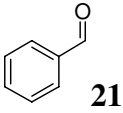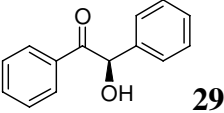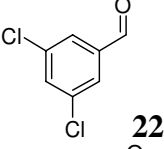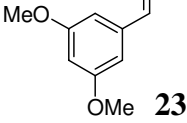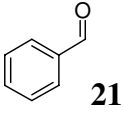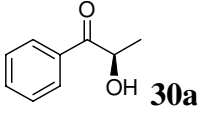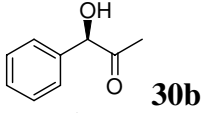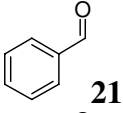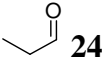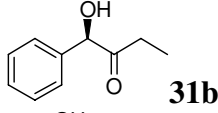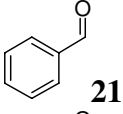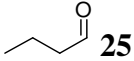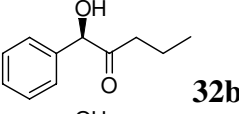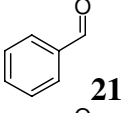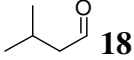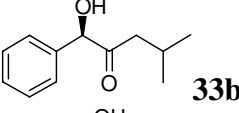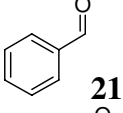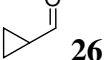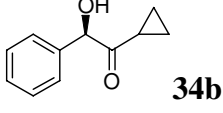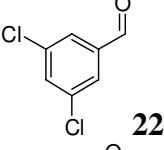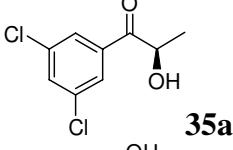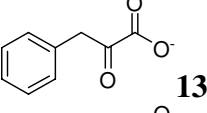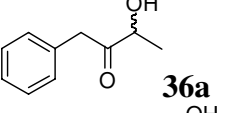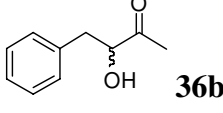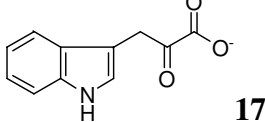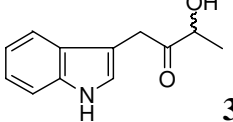
8.2 Carboligation of aromatic aldehydes
We could show that KdcA catalyzes the selfligation of benzaldehyde (**21**) to (*R*)-benzoin (**29**) with high enantiomeric excess (> 98%), whereas 3,5-dichlorobenzaldehyde (**22**) or 3,5-dimethoxybenzaldehyde (**23**) were not ligated to symmetric benzoin derivatives (Tab. 2, entry **4-6**).

8.3 Mixed carboligation of aromatic and aliphatic aldehydes
As shown above KdcA accepts various aliphatic aldehydes as well as benzaldehyde as acyldonor and acceptor. However, it is unknown, whether benzaldehyde derivatives like **22** or **23** can function as selective donors or acceptors for this enzyme. Therefore, we tested the catalytic activity of KdcA with various mixtures of aromatic and aliphatic aldehydes. As demonstrated in Table 2 (entry **7**) the mixed carboligation of acetaldehyde and benzaldehyde resulted in the formation of nearly equal amounts of (*R*)-2-hydroxypropiophenone (ee 93%) (2-HPP, **30a**) and (*R*)-phenylacetylcarbinol (ee 92%) (PAC, **30b**) besides traces of (*R*)-benzoin (**29**) as was deduced from GC/MS and HPLC.

Table 2: Results of KdcA catalyzed carboligation reactions.
(* absolute configuration was determined by comparison of HPLC data and with regard to mechanistic aspects of KdcA catalysis; n.d.: not determined).

No.	Substrate 1	Substrate 2	Product(s)	Enantiomeric exess
1	**18**	-	**27**	30-47% (*S*)
2	**19**	-	-	-
3	**20**	-	**28**	46% (*R*)
4	**21**	-	**29**	>98% (*R*)
5	**22**	-	-	-
6	**23**	-	-	-
7	**21**	**20**	**30a**, **30b**	**30a**: (*R*) 93% **30b**: (*R*) 92% **30a/30b**: 60:40
8	**21**	**24**	**31b**	>98% (*R*)*
9	**21**	**25**	**32b**	96.5% (*R*)
10	**21**	**18**	**33b**	88% (*R*)
11	**21**	**26**	**34b**	98% (*R*)*
12	**22**	**20**	**35a**	96.5% (*R*)
13	**13**	**20**	**36a**, **36b**	ee n.d. **36a/36b**: 80:20
14	**17**	**20**	**37a**	ee n.d.

The combination of benzaldehyde with larger aliphatic aldehydes gave remarkable results: the selectivity of KdcA shifted completely to the PAC-derivatives **31b-34b,** if propanal (**24**), butanal (**25**), isovaleraldehyde (**18**) or cyclopropylcarbaldehyde (**26**) were applied (Tab. 2, entry **8-11**). This is in contrast to similar BAL-catalyzed reactions, which resulted e.g. in mixtures of the HPP (**32a**)- and PAC (**32b**)–derivatives using benzaldehyde and propanal assubstrates (unpublished results).

Subsequently, 3,5-dichloro- (**22**) and 3,5-dimethoxy benzaldehyde (**23**) were tested in KdcA-catalyzed carboligations together with acetaldehyde. Whereas no reaction occurred with **23** the dichloro derivative **22** is clearly accepted as a donor aldehyde, yielding selectively the 2-HPP derivative **35a** with high enantiomeric excess (ee 96.5%) (Tab. 2, entry **12**).

The 2-hydroxyketones **29-35** are formed with high enantioselectivity (ee > 88%). The absolute (*R*)-configuration was assigned unambiguously to benzoin (**29**), 2-hydroxypropiophenone (**30a**), phenylacetylcarbinol (**30b**), **32b**, **33b,** and **35a** using circular dichroism or by comparison with authentic samples. The absolute configuration of the hydroxyketones **31b** and **34b** is assumed to be (*R*), too, according to comparison with products obtained from BAL and PDC catalysis.

8.4 Carboligation of phenylacetaldehyde and indole-3-acetaldehyde with acetaldehyde

CH-acidic aldehydes such as phenylacetaldehyde and indole-3-acetaldehyde are prone to enolization and thus are difficult substrates for enzymatic and also non-enzymatic transformations.[23] Moreover, indole-3-acetaldehyde is very instable and decomposes rapidly, which renders its direct application in biotransformations impossible.[24]

Phenylacetaldehyde is stable and commercially available; however, in aqueous buffer solution aldol reaction occurs spontaneously. Application of this aldehyde in KdcA-catalyzed carboligations with acetaldehyde gives the 2-hydroxyketones **36a,b** besides significant amounts of the aldol product.

To overcome these problems, we made use of KdcA's ability to decarboxylate the corresponding 2-ketoacid of both aldehydes, which offers an easy way to generate these aldehydes *in situ*. Referring to the reaction mechanism, the decarboxylation of a 2-ketoacid results in a reaction intermediate with the corresponding aldehyde bound to the cofactor ThDP. This so called "activated aldehyde" can react with a further acceptor aldehyde. If the KdcA-catalyzed decarboxylation of phenylpyruvate (**13**) and indole-3-pyruvate (**17**) is performed in the presence of acetaldehyde the 2-hydroxyketones **36a**,**b** and **37**, respectively, are formed predominantly in good yield (Tab. 2, entry 13, 14).

Conclusions

KdcA is a valuable new biocatalyst for asymmetric C-C bond formation. The enzyme can be produced easily at large scale in *E. coli* and purified using immobilized metal affinity chromatography. The branched-chain ketoacid decarboxylase is stable up to 40 °C at pH 6.8, and also in the presence of water miscible organic solvents such as 20% (v/v) DMSO.

The enzyme is able to catalyze the carboligation of a broad range of aldehydes to chiral 2-hydroxy ketones. Most interestingly the selectivity of the products formed in mixed acyloin synthesis of benzaldehyde and derivatives with different aliphatic acceptor aldehydes strongly depends on the substrate combination. The product composition can be shifted from a selectively formed PAC derivative of the aromatic aldehyde in the presence of large aliphatic aldehydes, over a nearly equal ratio of the HPP- and PAC-derivative in the presence of acetaldehyde to a selectively formed HPP derivative in the presence of 3,5-dichlorobenzaldehyde (Tab. 2). A further important achievement is the selective asymmetric synthesis of 2-hydroxyketones containing phenylacetaldehyde and indole-3-acetaldehyde. Good results have been obtained by *in situ* production of these enolizable (CH-acidic) aldehydes by enzymatic decarboxylation of the corresponding 2-ketoacid in the presence of acetaldehyde, with both reactions being catalyzed by KdcA (Tab. 2).

Experimental Section

Cloning of KdcA

The KdcA gene was amplified by PCR from the genomic DNA of *Lactococcus lactis* subspecies cremoris B1157 which was kindly provided by NIZO food research B. V. (Ede, NL). The strain was grown in M17' medium and isolation of genomic DNA was performed using the DNA-Tissue Kit (Qiagen). The following primers have been used for cloning:

LlKdcA_hisN-up: 5′– ATAT*GCTAGC*ATGTATACAGTAGGAGATTAC– 3′
                            *Nhe*I   start
LlKdcA_hisN-down: 5′– ATAT*GAATTC*TTATTTATTTTGCTCAGCAAATAA –3′
                               *Eco*RI   stop

After digestion of the amplified gene with *Eco*RI the gene was first ligated into the vector pBluescript in order to allow restriction by *Nhe*I and *Eco*RI. Subsequently the kdca gene (1.6 kb) was ligated into vector pET28a (Novagen) carrying already the information for an N-terminal His-tag yielding the final construct pET28a-KdcA-His-N. Transformation of the expression host *E. coli* BL21(DE3) was performed by electroporation. Identity of the gene with the published sequence[10] was confirmed by DNA sequencing (Sequiserve, Germany).

Expression and purification

A fed-batch cultivation according to the method of Korz et al.[25] was performed in a 40 L Techfors reactor (Infors AG, CH) at 30 °C, pH 7.0. Dissolved oxygen saturation was regulated between 30-40% saturation by increasing the stirrer speed and the air flow rate and subsequently by a pressure increase during the high oxygen demanding overexpression. After a preliminary growth phase protein expression was induced by the addition of 1.5 mM IPTG at an $OD_{600} \sim 60$. From 15 L culture 1.2 kg *E. coli* cells with a specific activity of 21 U/mg in the crude cell extract were gained. Cells were harvested by a separator (Westfalia Separator AG) and stored at -20 °C.

Purification of KdcA to homogeneity (> 95%) was performed by immobilized metal ion chromatography followed by size exclusion chromatography using a purification protocol previously developed for benzoylformate decarboxylase[18] with the following alterations:

The pH of all buffers used was set to 6.8. Buffers contained 2.5 mM $MgSO_4$ and 0.1 mM ThDP. Proteins unspecifically bound were eluted with 50 mM imidazole. KdcA was eluted with 250 mM imidazole. After purification the enzyme was either freeze dried or diluted with 50% (v/v) glycerol and stored at -20 °C.

Decarboxylase activity

One unit of decarboxylase activity is defined as the amount of KdcA which catalyzes the decarboxylation of 1 μmol 3-methyl-2-oxobutanoic acid (**1**) per minute under standard conditions (pH 6.8, 30 °C).

To investigate the decarboxylase activity of KdcA two continuous decarboxylase assays were used. In the *coupled decarboxylase assay* horse liver alcohol dehydrogenase (HL-ADH) is used to reduce the aldehydes obtained from the KdcA-catalyzed decarboxylation. Thereby NADH is oxidized to NAD^+. The decay of NADH is followed spectrophotometrically for 90 sec at 340 nm.

Assay composition: 700 μL buffer A (50 mM potassium phosphate buffer, pH 6.8, 2.5 mM $MgSO_4$, 0.1 mM ThDP), 100 μL 3-methyl-2-oxobutanoic acid (or another 2-ketoacid) 300 mM in buffer A (final concentration 30 mM; except indole-3-pyruvate: 1 mM), 100 μL NADH 2.5 mM in buffer A (final concentration 0.25 mM), 50 μL HL-ADH (Sigma-Aldrich) 5 Units/mL in buffer A (final concentration 0.25 U/mL). The assay was started by addition of 50 μL KdcA.

Further, a *direct decarboxylase assay* following the direct decay of 3-methyl-2-oxobutanoic acid (**1**) was developed in order to measure KdcA activity under NADH degradating conditions.

Assay composition: 950 μL 3-methyl-2-oxobutanoic acid (**1**) (60 mM in buffer A). The reaction was started by addition of 50 μL KdcA solution. To avoid a background by absorption of isobutyraldehyde, the decay of 3-methyl-2-oxobutanoic acid (λmax 320 nm) (**1**) was followed at 340 nm ($\varepsilon = 0.017$ L mmol^{-1} cm^{-1}).

Kinetic constants were calculated my non-linear regression using Origin 7G SR4 (OriginLab Coop., Northampton, USA).

Protein determination

Protein determination was performed according to Bradford[26] using BSA as a standard.

Determination of molecular mass

Size-exclusion chromatography was performed using a Superdex G200 prep grade column (total volume 122 mL (∅ 1,6 cm)) (Amersham) and 50 mM potassium phosphate buffer, pH 6.5, including

2.5 mM MgSO$_4$, 0.1 mM ThDP and 150 mM KCl. The coefficient of available volume (Kav = (Ve-Vo)/(Vt-Vo), Ve: elution volume of the respective protein, Vt: total volume, elutions volume of blue dextran) for KdcA and the standard proteins have been determined twice with a standard deviation of 0.2%.

Calibration was performed using ribonuclease A (13.7 kD, Kav = 0.66), chymotrypsinogen A (25 kD, Kav = 0.6), ovalbumin (43 kD; Kav = 0.48), BSA (67 kD, Kav = 0.4), aldolase (158 kD, Kav = 0.29), catalase (232 kD, Kav = 0.27), Ferritin (440 kD, Kav = 0.15), and Thyroglobulin (669 kD, and Kav = 0.07). The Kav-coefficient of KdcA was determined as 0.33. Data were plotted as logMr over Kav resulting in a linear correlation with a R^2 = 0.9919

Flow: 1 mL/min

Sample: 2 mL KdcA (1 mg/mL) in 50 mM potassium phosphate buffer, pH 6.5, including 2.5 mM MgSO$_4$, 0.1 mM ThDP

Determination of pH and temperature optima

The pH-optimum was measured using the coupled decarboxylase assay. For determination of the temperature optimum the direct decarboxylase assay was used.

Stability investigations

For investigation of the stability towards pH, temperature, and organic solvents KdcA was incubated under the reaction conditions given in the figure legends and residual activity was assayed with the coupled decarboxylase assay.

Analytical performance

The conversion was followed by GCMS, employing a HP 6890 series GC-system fitted with a HP 5973 mass selective detector (Hewlett Packard; column HP-5MS, 30 m·250 µm; T_{GC}(injector) = 250 °C, T_{MS}(ion source) = 200 °C, time program (oven): $T_{0\,min}$ = 60 °C, $T_{3\,min}$ = 60 °C, $T_{14\,min}$ = 280 °C (heating rate 20 °C·min^{-1}), $T_{19\,min}$ = 280 °C). The enantiomeric excess was determined by chiral GC, employing a Shimadzu GC 2010, fitted with a FS Lipodex D column (50 m x 0.25 mm) and a FID detector, or by chiral HPLC employing a HP 1100 HPLC System (Agilent) fitted with a diode-array detector. NMR spectra were recorded on a Bruker DPX-400. Chemical shifts are reported in ppm relative to CHCl$_3$ (^{1}H NMR: δ = 7.27) and CDCl$_3$ (^{13}C NMR: δ = 77.0) as internal standards. CD-spectra were recorded on a JASCO J-810 Spectropolarimeter using acetonitrile as solvent.

Carboligase activity

Representative example for the KdcA-catalyzed synthesis of 2-hydroxyketones:

(R)-1-Hydroxy-1-phenylpentan-2-one (32b)

Benzaldehyde (29 mg, 0.27 mmol) was dissolved in a mixture of dimethyl sulfoxide (3 mL) and potassium phosphate buffer [12 mL, 50 mM, pH 6.8, containing MgSO$_4$ (2.5 mM) and ThDP (0.1 mM)]. To this solution 19 mg (0.27 mmol) butanal was added. After addition of KdcA (3.3 mg Protein) the reaction was stirred slowly at 30 °C for 104 h. The reaction mixture was extracted with diethylether (25 mL) and the organic layer washed with brine and dried over Na$_2$SO$_4$. Evaporation of the solvent and purification of the crude product, which contained besides **32b** small amounts of substrates and 5-hydroxyoctan-4-one, by flash column chromatography afforded 15.6 mg (32%, 96.5% ee) (*R*)-1-hydroxy-1-phenylpentan-2-one (**32b**) as low-viscous-oil; HPLC: (Chiral OM n-hexane / 2-propanol 98:2, 0.5 mL min^{-1}, 40°C) R_t(S) = 30.1 min, R_t(R) = 36.5 min; $[\alpha]_D^{23\,°C}$= -130.4 (c = 0.1 g/ 100 ml CDCl$_3$); CD (acetonitrile): λ (Δε) [nm] (mol. CD) =199 (+18.02), 211 (+9.5), 218 (+9.7), 231 (+0.9), 238 (-0.1), 283 (-7.3); ^{1}H NMR (400 MHz, CDCl$_3$, ppm) δ = 0.81 (t, 3 H, *J* = 7.5 Hz, CH$_3$), 1.46-1.65 (m, 2 H, CH$_2$), 2.25-2.41 (m, 2 H, CH$_2$), 4.37 (bs, OH), 5.08 (s, 1 H, *CH*OH), 7.3-7.4 (m, 5 H, ArH) – contains 2 mol% acyloin; ^{13}C NMR (100 MHz, CDCl$_3$, ppm) δ = 13.5 (CH$_3$), 17.1, 39.7 (CH$_2$), 79.7 (CHOH), 127.4, 128.6, 128.9 (CH$_{ar}$), 138.08 (C$_q$), 209.4 (C=O); GCMS R_t = 9.02 min, MS (70 eV, EI): *m/z* (%):178 (1%) [M$^+$], 107 (100%), 79 (42%).

(R)*-1-Hydroxy-1-phenylbutan-2-one (31b)
HPLC (Chiral OM, n-hexane / 2-propanol 95:5, 0.5 mL min^{-1}, 40°C): ee > 98.5%, R_t = (S) 21.6 min, R_t (R) =25.0 min; ^{1}H NMR (400 MHz, CDCl$_3$, ppm) δ = 1.01 (t, ^3J(H,H) = 7.4 Hz, 3 H, CH$_3$), 2.29-2.46 (m, 2 H, CH$_2$), 4.3-4.4 (bs, 1 H, OH), 5.1 (s, 1 H, C*H*OH), 7.3-7.41 (m, 5H, Ar-H) - contains 3 mol% benzaldehyde and acylion; ^{13}C NMR (100 MHz, CDCl$_3$, ppm) δ = 7.6 (CH$_3$), 31.1 (CH$_2$), 79.4 (CHOH), 127.3 (2 x Ar), 128.6, 128.9 (CH$_{ar}$), 138.3 (C$_q$), 210.1 (C=O); GCMS R_t = 8.44 min, MS (70 eV, EI): *m/z* (%): 164 (0.1%) [M$^+$], 107 (100%), 79 (81%).

(R)-1-Hydroxy-1-phenyl-4-methylpentan-2-one (33b)
Isolated yield: 25%; ee = 88% ; HPLC, Chiralcel OD-H, n-hexane / 2-propanol 98:2, 0.5 mL min^{-1}, 25° C, R_t = (S) 22.9 min, R_t (R) =29.5 min; [α]$_D^{23}$ $^{°C}$= -256.2 (c = 0.2 g/ 100 ml CDCl$_3$); CD (acetonitrile): λ (Δε) [nm] (mol. CD) =199 (+20.9), 218 (+11.05), 231 (+0.9), 237 (+0.04), 283 (-8.5); ^{1}H NMR (400 MHz, CDCl$_3$, ppm) δ = 0.75 (d, 3 H, J = 6.5 Hz, CH$_3$), 0.89 (d, 3 H, J= 6.5 Hz, CH$_3$), 2.04-2.31 (m, 3 H, CH, CH$_2$), 4.39 (d, J = 4.1 Hz, OH), 5.05 (d, 1 H, J=4.1, C*H*OH), 7.29-7.41 (m, 5 H, ArH) – contains 20 mol% acyloin and substrate; ^{13}C NMR (100 MHz, CDCl$_3$, ppm) δ = 22.2 (CH$_3$), 22.4 (CH$_3$), 24.6 (CH), 46.6 (CH$_2$), 80.0 (CHOH), 127.4, 128.6, 128.9 (CH$_{ar}$), 137.9 (C$_q$), 209.0 (C=O); GCMS R_t = 9.30 min, MS (70 eV, EI): *m/z* (%): 192 (1%) [M$^+$], 136 (1%) , 107 (100%), 79 (58%);

(R)*-2-Cyclopropyl-1-hydroxy-1-phenylethanone (34b)
Isolated yield 14%; ee = 98%; HPLC (Chiracel-OD-H, n-hexane / 2-propanol 95:5, 0.5 mL min^{-1}, 40 °C): R_t = (S) 16.6 min, R_t (R) = 21.15 min; ^{1}H NMR (400 MHz, CDCl$_3$, ppm) δ = 0.78-0.86 (m, 1 H), 0.94-1.2 (m, 3 H), 1.83-1.9 (m, 1 H), 4.39 (bs, 1 H, OH), 5.27 (s, 1 H, C*H*OH), 7.3-7.45 (m, 5 H, Ar-H) – contains 20 mol% acyloin; ^{13}C NMR (100 MHz, CDCl$_3$, ppm) δ = 12.1, 12.7 (CH$_2$), 17.6 (CH), 80.1 (CHOH), 127.7, 128.6, 128.9 (CH$_{ar}$), 138.1 (C$_q$), 209.6 (C=O); GCMS R_t = 9.45 min, MS (70 eV, EI): *m/z* (%): 176 (3%) [M$^+$], 107 (100%), 79 (71%).

(R)-1-(3,5-Dichlorophenyl)-2-hydroxypropan-1-one (35a)
HPLC (Chiralpak AD, n-hexane / 2-propanol 98:2, 0.75 ml min^{-1}, 20°C): ee = 96.5%, R_t = (S) 20.4 min, R_t (R) = 25.4 min; CD (acetonitrile): λ (Δε) [nm] (mol. CD) = 212 (+4.9), 211 (+9.5), 245 (-2.0), 289 (+1.65); ^{1}H NMR (400 MHz, CDCl$_3$, ppm) δ = 1.46 (d, 3H, J=7.1, CH$_3$), 3.56 (bs, 1 H, OH), 5.07 (q, 1 H, J=7.1, C*H*OH), 7.61 (t, 1 H, J=2.07, ArH), 7.78 (d, 2 H, J=2.07, ArH) - contains 8 mol% 3,5-dichlorobenzaldehyde;, ^{13}C NMR (100 MHz, CDCl$_3$, ppm) δ =21.9 (CH$_3$), 69.6 (CHOH), 126.9, 133.5 (CH$_{ar}$), 135.9, 136.0 (C$_q$), 200.2 (C=O); GCMS R_t = 10.04 min, MS (70 eV, EI): *m/z* (%): 218 (7%) [M$^+$], 175 (100%), 147 (43%), 111 (71%).

3-Hydroxy-1-phenylbutan-2-one (36a) and 1-Phenyl-2-hydroxybutan-3-one (36b)
36a: yield: 23%; ee = not determined; ^{1}H NMR (400 MHz, CDCl$_3$, 300 K, ppm) δ = 1.4 (d, 3 H, J = 7.1 Hz, CH$_3$), 3.55 (d, 1 H, J = 4.6 Hz, OH), 3.77 (d, 1 H, J = 15.8 Hz, CH$_2$), 3.83 (d, 1 H, J = 15.8 Hz, CH$_2$), 4.3-4.36 (m, 1 H, C*H*OH), 7.18-7.34 (m, 5 H, ArH) – contains 19 mol% **36b**; ^{13}C NMR (400 MHz, CDCl$_3$, 300 K, ppm) δ = 19.7 (CH$_3$), 44.5 (CH$_2$), 72.2 (CHOH), 127.2, 128.7, 129.4, 133.1, 210.04 (C=O); GCMS R_t = 8.6 min, MS (70 eV, EI): *m/z* (%): 164 (1%) [M$^+$], 146 (30%), 121 (41%), 103 (38%), 91 (100%).
36b: yield: 23%; ee = not determined; ^{1}H NMR (400 MHz, CDCl$_3$, 300 K, ppm) δ = 2.19 (s, 3 H, CH$_3$), 2.87 (dd, 1 H, J = 14.2 Hz, J = 7.5 Hz, CH$_2$), 3.12 (dd, 1 H, J = 14.2 Hz, J = 4.6 Hz, CH$_2$), 3.52 (d, 1 H, J = 5.3 Hz, OH), 4.38-4.42 (m, 1 H, CHOH), 7.18-7.34 (m, 5 H, ArH) – contains **36a**; ^{13}C NMR (400 MHz, CDCl$_3$, 300 K, ppm) δ = 25.8 (CH$_3$) 39.8 (CH$_2$) 77.6 (CHOH), 126.8, 128.5, 129.2, 209.4 (C=O); GCMS R_t = 8.6 min, MS (70 eV, EI): m/z (%): 164 (1%) [M$^+$], 146 (30%), 121 (41%), 103 (38%), 91 (100%).

3-Hydroxy-1-(3-indolyl)-butan-2-one (37a)
Yield: 23 %; ee = not determined; ^{1}H NMR (400 MHz, CDCl$_3$, 300 K, ppm) δ = 1.46 (d, 3 H, J = 7.1 Hz, CH$_3$), 3.43 (d, 1 H, J = 4.9 Hz ,OH), 3.96 (s, 2 H, CH$_2$), 4.38-4.45 (m, 1 H, C*H*OH), 7.14-7.18 (m, 1 H, ArH + 1 H, NCH), 7.21-7.25 (m, 1 H, ArH), 7.39 (d, 1 H, ArH),7.53 (d, 1 H, ArH),8.06-8.19 (bs, 1 H, NH); GCMS R_t = 12.54 min, MS (70 eV, EI): *m/z* (%): 203 (9.5%) [M$^+$], 130 (100%).
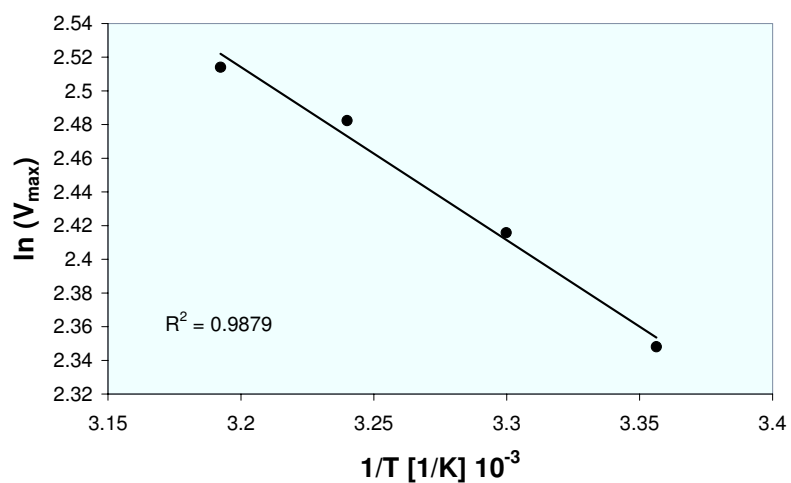
Acknowledgements

References

[1]   T. Stillger, M. Pohl, C. Wandrey, A. Liese, *Org. Proc. Res. Develop.* **2006**, *10*, 1172-1177.
[2]   M. Knoll, M. Müller, J. Pleiss, M. Pohl, *Chembiochem* **2006**, *7*, 1928-1934.
[3]   E. Janzen, M. Müller, D. Kolter-Jung, M. M. Kneen, M. J. McLeish, M. Pohl, *Bioorg. Chem.* **2006**, *34*, 345-361.
[4]   P. Domingez de Maria, T. Stillger, M. Pohl, S. Wallert, K. Drauz, H. Gröger, H. Trauthwein, A. Liese, *J. Mol. Catal. B: Enzym.* **2006**, *38*, 43-47.
[5]   P. Siegert, M. J. McLeish, M. Baumann, H. Iding, M. M. Kneen, G. L. Kenyon, M. Pohl, *Protein Eng. Des. Sel.* **2005**, *18*, 345-357.
[6]   M. Pohl, B. Lingen, M. Müller, *Chem. Eur. J.* **2002**, *8*, 5288-5295.
[7]   A. S. Demir, O. Sesenoglu, E. Eren, B. Hosrik, M. Pohl, E. Janzen, D. Kolter, R. Feldmann, P. Dünkelmann, M. Müller, *Adv. Syn. Catal.* **2002**, *344*, 96-103.
[8]   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.
[9]   G. Smit, B. A. Smit, W. J. Engels, *FEMS Microbiol. Rev.* **2005**, *29*, 591-610.
[10]   B. A. Smit, J. Vlieg, W. J. M. Engels, L. Meijer, J. T. M. Wouters, G. Smit, *Appl. Environ. Microbiol.* **2005**, *71*, 303-311.
[11]   M. de la Plaza, P. Fernandez de Palencia, C. Pelaez, T. Requena, *FEMS Microbiol. Lett.* **2004**, *238*, 367-374.
[12]   A. Yep, G. L. Kenyon, M. J. McLeish, *Bioorg. Chem.* **2006**, *34*, 325-336.
[13]   P. Arjunan, T. Umland, F. Dyda, S. Swaminathan, W. Furey, M. Sax, B. Farrenkopf, Y. Gao, D. Zhang, F. Jordan, *J. Mol. Biol.* **1996**, *256*, 590-600.
[14]   D. Dobritzsch, S. König, G. Schneider, G. Lu, *J. Biol. Chem.* **1998**, *273*, 20196-20204.
[15]   M. S. Hasson, A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A. Gerlt, G. L. Kenyon, G. A. Petsko, D. Ringe, *Biochemistry* **1998**, *37*, 9918-9930.
[16]   A. K. Chang, R. G. Duggleby, *Biochem. J.* **1997**, *327*, 161-169.
[17]   Y. Lindqvist, G. Schneider, U. Ermler, M. Sundstrom, *Embo J.* **1992**, *11*, 2373-2379.
[18]   H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir, M. Pohl, *Chem. Eur. J.* **2000**, *6*, 1483-1495.
[19]   F. Hildebrand, S. Kühl, M. Pohl, D. Vasic-Racki, M. Müller, C. Wandrey, S. Lütz, *Biotechnol. Bioeng.* **2006**, *96*, 835-843.
[20]   P. Domingez de Maria, T. Stillger, M. Pohl, S. Wallert, K. Drauz, H. Gröger, H. Trauthwein, A. Liese, *J. Mol. Catal. B: Enzym.* **2005**, *38*, 43-47.
[21]   E. Janzen, M. Müller, D. Kolter-Jung, M. M. Kneen, M. J. McLeish, M. Pohl, *Bioorg. Chem.* **2006**, *34*, 345-361.
[22]   P. Domingez de Maria, M. Pohl, D. Gocke, H. Gröger, H. Trauthwein, L. Walter, M. Müller, *Eur. J. Org. Chem.* **2007**, *in press*.
[23]   A. Schütz, R. Golbik, K. Tittmann, D. I. Svergun, M. H. Koch, G. Hübner, S. König, *Eur. J. Biochem.* **2003**, *270*, 2322-2331.
[24]   J. Koga, T. Adachi, H. Hidaka, *J. Biol. Chem.* **1992**, *267*, 15823-15828.
[25]   D. J. Korz, U. Rinas, K. Hellmuth, E. A. Sanders, W.-D. Deckwer, *J. Biotechnol.* **1995**, *39*, 59-65.
[26]   M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248-254.

Supplementary Material:

Determination of native molecular weight by size-exclusion chromatography: KdcA: logMr = 0.33.
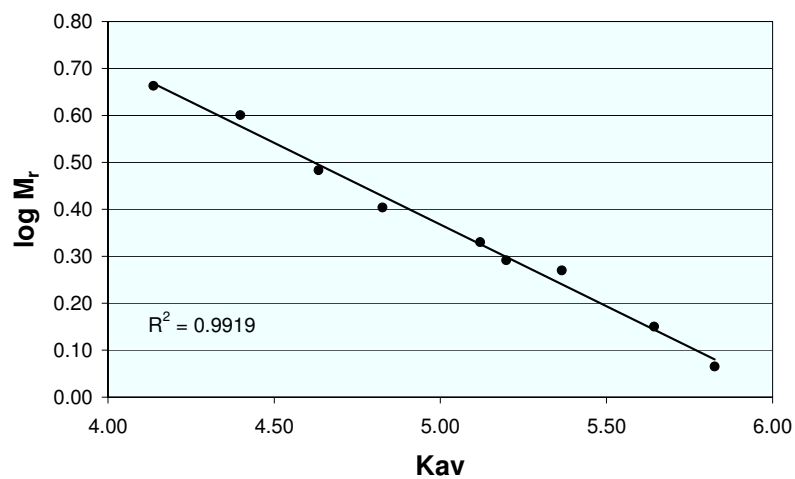
Determination of the activation energy of the decarboxylation of according to Arrhenius (see Fig. 4)

Publication VI

Crystal structure of the branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* provides insights into the structural basis for the chemo- and enantioselective carboligation reaction.

Berthold, C. L., Gocke, D., Wood, M. D., Leeper, F. J., Pohl, M. & Schneider, G

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

Catrine L. Berthold,[a] Dörte Gocke,[b] Martin D. Wood,[c] Finian J. Leeper,[c] Martina Pohl[b] and Gunter Schneider[c]*

[a]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden, [b]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Research Center Jülich, 52426 Jülich, Germany, and [c]Cambridge University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

Correspondence e-mail: gunter.schneider@ki.se

The thiamin diphosphate (ThDP) dependent branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* catalyzes the decarboxylation of 3-methyl-2-oxobutanoic acid to 3-methylpropanal (isobutyraldehyde) and CO_2. The enzyme is also able to catalyze carboligation reactions with an exceptionally broad substrate range, a feature that makes KdcA a potentially valuable biocatalyst for C—C bond formation, in particular for the enzymatic synthesis of diversely substituted 2-hydroxyketones with high enantioselectivity. The crystal structures of recombinant holo-KdcA and of a complex with an inhibitory ThDP analogue mimicking a reaction intermediate have been determined to resolutions of 1.6 and 1.8 Å, respectively. KdcA shows the fold and cofactor–protein interactions typical of thiamin-dependent enzymes. In contrast to the tetrameric assembly displayed by most other ThDP-dependent decarboxylases of known structure, KdcA is a homodimer. The crystal structures provide insights into the structural basis of substrate selectivity and stereoselectivity of the enzyme and thus are suitable as a framework for the redesign of the substrate profile in carboligation reactions.

1. Introduction

ThDP-dependent enzymes appear to be particularly promising as catalysts in bioorganic synthesis owing to their large diversity in substrates and diverse stereospecificity (Müller & Sprenger, 2004). The ability of many ThDP-dependent enzymes to catalyze the ligation of aldehydes into chiral 2-hydroxyketones as a side reaction, the so-called carboligation reaction, has been used in large-scale organic synthesis for decades (Sprenger & Pohl, 1999). The mode of action of the cofactor in ThDP-dependent decarboxylases is well understood (Kern *et al.*, 1997) and the carboligation reaction is most likely to proceed through the same mechanism (Fig. 1; Lobell & Crout, 1996; Siegert *et al.*, 2005).

A general feature of the family of ThDP-dependent enzymes is a high variation of the active-site residues, with hardly any strict conservation, despite the conserved overall reaction mechanism. This provides the basis for the considerable variation in the chemoselectivity of the carboligation reaction, which is ultimately dependent on the size and composition of the substrate-binding pocket. It is also noteworthy that most of the enzymes only catalyze the synthesis of (*R*)-enantiomers of 2-hydroxyketones and that the corresponding (*S*)-products are not readily accessible. Exceptions

have been reported concerning the ligation of benzaldehyde derivatives with acetaldehyde using benzoylformate decarboxylase (BFD) from *Pseudomonas putida* as catalyst, resulting in (*S*)-hydroxyketones (Iding *et al.*, 2000).

The stereochemical outcome of the carboligation reaction is dependent on whether the *Re* or *Si* face of the acceptor

substrate approaches the α-carbanion/enamine intermediate; an approach by the *Si* face results in the (*R*)-enantiomer and the *Re* face yields the (*S*)-product (Iding *et al.*, 2000). Steric reasons have been put forward to explain the preference for (*R*)-enantioselectivity of the product (Knoll *et al.*, 2006). Analysis of the structures of several ThDP-dependent decarboxylases suggested that during the C—C bond-forming step the O atoms of the two substrates point roughly in the same direction to allow proton transfer facilitated by a histidine residue and that the presence or absence of a pocket, denoted the S-pocket, determines the enantioselectivity in the carboligation reaction (Knoll *et al.*, 2006; Fig. 2). Formation of the (*S*)-enantiomer in high enantiomeric excess only occurs when the acceptor aldehyde has an optimal fit in this S-pocket, otherwise the (*R*)-product is formed. Rational redesign of the S-pocket in BFD indeed resulted in an (*S*)-specific enzyme variant accepting longer aliphatic aldehydes, demonstrating the structure-guided approach for engineering enantioselectivity in ThDP-dependent enzymes (Gocke *et al.*, unpublished results).

The branched-chain keto acid decarboxylase (KdcA) from *Lactococcus lactis* is a recently discovered member of the ThDP-dependent enzyme family which was originally identified as taking part in flavour formation during cheese ripening (Smit *et al.*, 2005). A sequence comparison suggests that KdcA belongs to the ThDP-dependent decarboxylases (Yep *et al.*, 2006), also named the POX group (Duggleby, 2006) after pyruvate oxidase, the first member of this subgroup of ThDP-dependent enzymes of known structure (Muller & Schulz, 1993). KdcA catalyses the decarboxylation of 3-methyl-2-oxobutanoic acid into isobutyraldehyde, but other substrates are also accepted for decarboxylation (Smit *et al.*, 2005; Yep *et al.*, 2006; Gocke *et al.*, 2007). A recent study of the carboligation activity of KdcA revealed an exceptionally broad substrate range, including various aromatic aldehydes and CH-acidic aldehydes as well as alipathic aldehydes (Gocke *et al.*, 2007). KdcA is the only enzyme so far identified that accepts substrates as large as indole-3-acetaldehyde, produced *in situ* by decarboxylation of
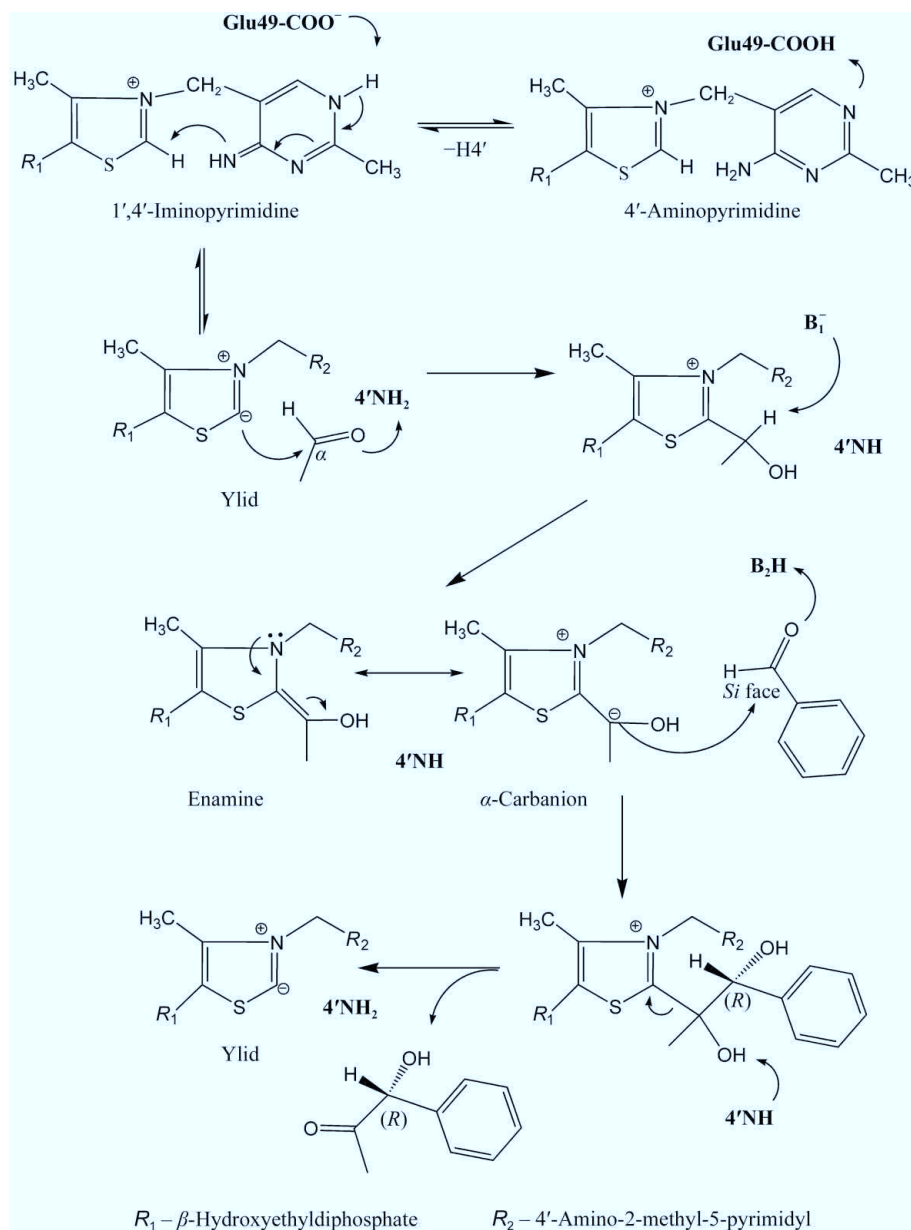
Figure 1
The mechanism of the carboligation reaction of ThDP-dependent enzymes with acetaldehyde as acyl donor and benzaldehyde as the acceptor substrate. KdcA catalyzes the formation of the (*R*)-enantiomer of phenylacetylcarbinol in 92% enantiomeric excess (Gocke *et al.*, 2007). Initially, the cofactor is activated by the enzyme by a conserved glutamic acid residue that stabilizes the imino form of the pyrimidine ring, in which $4'$NH can abstract a proton from the C2 atom of the thiazolium ring and create the reactive ylid. Next follows activation of the acyl donor aldehyde by covalent linkage to the C2 atom of ThDP. During this step $4'$NH$_2$ is thought to stabilize the negative charge generated on the substrate carbonyl group. After deprotonation of C$^\alpha$, which forms the enamine/α-carbanion intermediate, ligation with the acceptor aldehyde follows and then product release. Note that the roles of donor and acceptors can be reversed (*i.e.* benzaldehyde is the acyl donor and acetaldehyde is the acceptor substrate), which results in the formation of (*R*)-2-hydroxypropiophenone.

indole-3-pyruvate, as an acyl donor for carboligation. These features make KdcA a valuable addition to the toolbox of ThDP-dependent enzymes for C—C bond formation (Müller & Sprenger, 2004), in particular the enzymatic synthesis of diversely substituted 2-hydroxyketones with high enantio-selectivity. Here, we provide high-resolution crystal structures of recombinant holo-KdcA and of a complex with an inhibitory ThDP analogue that mimics a reaction intermediate. Based on these two structures, we compare and discuss the carboligation activity of KdcA and other homologous enzymes for which the reaction has been characterized.

2. Methods

2.1. Protein production

Full details of the cloning, expression and purification of KdcA from *L. lactis* have been described elsewhere (Gocke *et al.*, 2007). In brief, the gene coding for KdcA was cloned into the pET28a (Novagen) vector containing a hexahistidine tag and recombinant protein was produced in *Escherichia coli* strain BL21(DE3). The protein contains a 23-amino-acid N-terminal extension, MGSSHHHHHHSSGLVPRGSHMAS, and could be purified to homogeneity by a single step of Ni–NTA chromotography.

2.2. Preparation of apoenzyme and inhibition by a deazaThDP analogue

The apo form of KdcA was produced by overnight dialysis against 50 mM Tris–HCl pH 8.5, 1 mM EDTA and 1 mM DTT. The buffer was exchanged to 20 mM MES buffer pH 6.8 containing 1 mM DTT prior to protein concentration by centrifugation. No decarboxylase activity was observed for the apoenzyme using the assay described previously (Gocke *et al.*, 2007), while activity was restored upon addition of ThDP and Mg^{2+}. Incubation of the apoenzyme with 0.1 mM 2-[(R)-1-hydroxyethyl]deazaThDP prior to addition of ThDP and Mg^{2+} did not result in active holoenzyme, demonstrating the strong binding of the inactive deazaThDP analogue.

2.3. Crystallization

Crystallizations were performed using the vapour-diffusion method. A protein solution of 3–5 mg^{-1} KdcA in 20 mM MES buffer pH 6.8 containing 2.5 mM MgSO$_4$, 0.1 mM ThDP and 1 mM DTT was used for crystallization of the holoenzyme. Initial conditions were found using commercial crystallization screens; the small irregular crystals obtained from these could be optimized to single 0.1 × 0.1 × 0.2 mm crystals by additive screening. The optimized crystals were produced in 24-well hanging-drop plates with a reservoir solution containing 25–26%(v/v) PEG 200 in 0.1 M MES buffer pH 6.2. The drops were set up by mixing 2 µl protein solution with 2 µl well solution to which 0.5 µl of the additive spermine tetra-hydrochloride (0.1 M) had been added. The crystals grew to full size within 1 d and were flash-frozen in liquid nitrogen without any additional cryoprotectant.

The complex of KdcA with 2-[(R)-1-hydroxyethyl]deaza-ThDP was crystallized using the same condition as for the holoenzyme. The deazaThDP derivative was synthesized as described by Leeper *et al.* (2005). A solution of the apoenzyme (5 mg ml^{-1} in 20 mM MES buffer pH 6.8) was incubated for several hours with 1 mM 2-[(R)-1-hydroxyethyl]deazaThDP and 2.5 mM MgSO$_4$ in order to form the complex. These crystals grew more slowly and heavy precipitant appeared among the few sporadically formed crystals.

2.4. Data collection and structure determination

X-ray data were collected in a nitrogen stream on beamlines ID14-EH1 and ID14-EH3 at the European Synchrotron Research Facility, Grenoble, France. A summary of the data statistics can be found in Table 1. The data were processed with *MOSFLM* (Leslie, 1992) and scaled with *SCALA* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Molecular replacement was carried out with *MOLREP* (Vagin & Teplyakov, 1997). The monomer of the closest homologue of known structure, indolepyruvate decarboxylase (IPDC) from *Enterobacter cloacae* (PDB code 1ovm; Schütz *et al.*, 2003), was used as a search model for the initial structure determination of the holoenzyme. Two monomers were found in the asymmetric unit, comprising the biological dimer. The biological sequence was fitted to the electron density with the program *ARP/wARP* (Perrakis *et al.*,

Figure 2
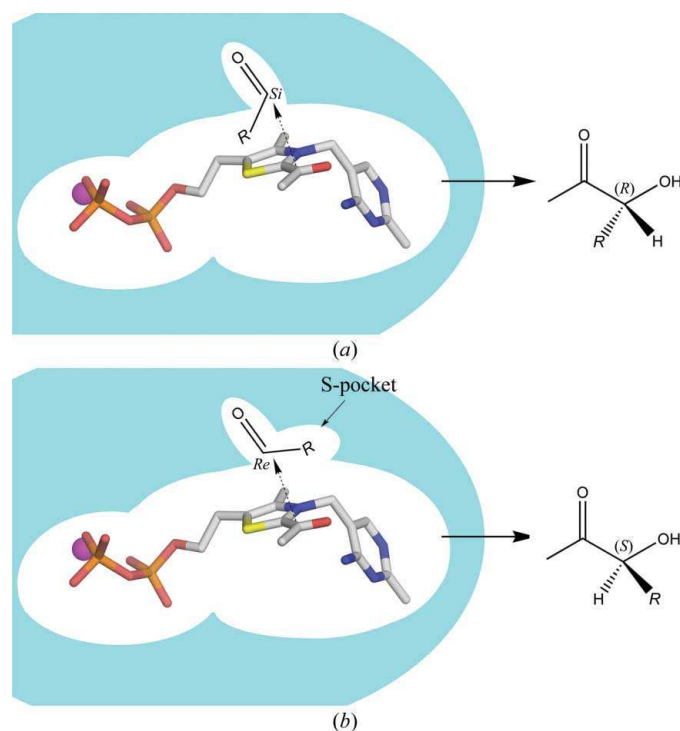A schematic representation of the S-pocket. (*a*) When there is no S-pocket or an S-pocket that is too small to fit the acceptor substrate side chain *R*, it will bind with the *Si* face towards the enamine and the (*R*)-enantiomer will be formed. (*b*) Perfect fit of the acceptor substrate side chain in the S-pocket above the thiazolium ring will allow the substrate to align, resulting in the (*S*)-enantiomer.

1999). Refinement was carried out with *REFMAC*5 (Murshudov *et al.*, 1997); *Coot* (Emsley & Cowtan, 2004) was employed for model building as well as assignment of water molecules. The structure of KdcA with the deazaThDP inhibitor was determined using the dimer of the holoenzyme as a search model. Restraints for the inhibitor molecule, obtained using the Dundee *PRODRG*2 server (Schüttelkopf & van Aalten, 2004), were loosened towards the end of the refinement. The validation tools in *Coot* as well as *PROCHECK* (Laskowski *et al.*, 1993) were used to monitor the geometry of the structures. All protein-molecule figures were produced using *PyMOL* (W. L. DeLano; http://www.pymol.org).

2.5. Modelling of carboligation substrates

Indole-3-pyruvate, which is the largest known substrate that (after decarboxylation) can undergo ligation to an acceptor aldehyde in KdcA, was modelled into the holoenzyme. Alignment of the C^α atom and the carboxyl group was performed by superimposing the structure of KdcA with those of substrate complexes of a homologous decarboxylase determined previously (Berthold *et al.*, 2007). The planar enamine intermediates were modelled based on the structure of the 2-[(R)-1-hydroxyethyl]deazaThDP–KdcA complex.

3. Results and discussion

3.1. Three-dimensional structure of the KdcA subunit

The structure of holo-KdcA from *L. lactis* was determined by molecular replacement to a resolution of 1.6 Å. The crystals belonged to space group $P2_12_12_1$ and contained one dimer per asymmetric unit. All residues (1–547) could be modelled into the electron-density map, with the exception of the loop 182–187, for which no interpretable electron density could be detected. In addition, residues 538–547 of the C-terminal helix were not well defined in the electron-density map and were refined with occupancies of 0.5 in one of the monomers. The bound cofactors ThDP and Mg^{2+} are well defined in both monomers, which allowed a precise definition of their interactions with enzymic groups.

The KdcA monomer shares the conserved fold of three domains of α/β topology common to members of the ThDP-dependent enzyme family (Muller *et al.*, 1993; Fig. 3a). ThDP is bound in a cleft formed at the interface between the PYR (residues 1–181) and PP (residue 346–547) domains from two different subunits and displays the characteristic V-conformation, stabilized by a conserved large hydrophobic residue (Ile404) positioned below the thiazolium ring of the cofactor (Lindqvist & Schneider, 1993). The R domain (residues 182–345) in KdcA does not display the nucleotide-binding site found in other members of the family (Muller & Schulz, 1993; Pang *et al.*, 2002; Berthold *et al.*, 2005) and the corresponding binding pocket is instead filled by large aromatic residues. The closest homologue IPDC (Schütz *et al.*, 2003) shows 41% sequence identity to KdcA and superposition of the subunits results in a root-mean-square deviation of 1.5 Å over 516 C^α atoms. The crystal structure of KdcA was further compared

Table 1
Crystallographic data-collection and refinement summary.

Values in parentheses are for the highest resolution shell.

	Holo-KdcA	KdcA–deazaThDP
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters		
a (Å)	65.6	66.0
b (Å)	108.4	108.6
c (Å)	146.6	146.5
Molecules in ASU	2	2
Resolution (Å)	1.60 (1.69-1.60)	1.8 (1.9-1.8)
R_{sym}	0.067 (0.476)	0.087 (0.376)
Mean $I/\sigma(I)$	11.3 (2.2)	13.1 (2.1)
Completeness (%)	92.4 (63.8)	97.4 (85.2)
Wilson B factor (Å^2)	18	16
Refinement		
Reflections in working set	114611	85927
Reflections in test set	6436	4792
R factor (%)	16.3	16.3
R_{free} (%)	20.0	20.5
Atoms modelled	9806	9634
No. of amino acids in model	1092	1090
No. of ligands	2	2
No. of magnesium ions	2	2
No. of water molecules	1108	950
Average B factor (Å^2)		
Protein	19.5	15.3
Ligands	12.7	9.7
Magnesium ions	13.2	10.8
Waters	32.3	26.0
R.m.s. deviations from ideals		
Bonds (Å^2)	0.009	0.008
Angles (°)	1.207	1.131
Ramachandran distribution (%)		
Most favoured	91.2	91.4
Additionally allowed	8.6	8.6
Generously allowed	0.2	0
Disallowed	0	0

with those of pyruvate decarboxylase from *Zymomonas mobilis* (*Zm*PDC; Dobritzsch *et al.*, 1998), BFD from *P. putida* (Hasson *et al.*, 1998) and benzaldehyde lyase from *P. fluorescens* (BAL; Mosbacher *et al.*, 2005), which are homologues with well characterized carboligation activity (Knoll *et al.*, 2006; Janzen *et al.*, 2006; Siegert *et al.*, 2005). Superposition of the KdcA monomer results in a root-mean-square deviation of 2.1 Å over 513 C^α atoms with *Zm*PDC (PDB code 1zpd), 2.5 Å over 442 C^α atoms with BFD (PDB code 1bfd) and 2.2 Å over 438 C^α atoms with BAL (PDB code 2ag0).

After completion of the model, two stretches of unassigned electron density remained in surface pockets located between the R and PYR domains from different subunits in the dimer (Fig. 4). These electron densities fitted well to the sequence stretch HSSGL (residues −14 to −10) of the N-terminal tag, but no electron density was observed between residue −10 and the N-terminus of the enzyme subunit. The distances between the N-termini of the two subunits in the KdcA dimer and these two peptide stretches are too large to be covered by the remaining nine residues of the tag, *i.e.* the bound peptide has to come from another molecule related by crystallographic symmetry. Indeed, the N-terminus of a symmetry-related adjacent molecule is located at a distance (∼22 Å) that could be bridged by the nine disordered residues of the tag. These

interactions between adjacent molecules in the lattice suggest that the peptide tags might contribute to crystal packing.

3.2. Quaternary structure

The smallest catalytic unit of ThDP-dependent enzymes is a dimer owing to the location of the active site at the interface of two subunits (Muller *et al.*, 1993). Most members of the POX subclass of ThDP-dependent enzymes, including IPDC, form tetramers in solution and in the crystal, with the packing being

best described as a dimer of dimers (Duggleby, 2006). The asymmetric unit of the crystals of KdcA contains two subunits which are related by a molecular noncrystallographic twofold axis. This molecular twofold axis generates a dimeric molecule with the packing common to other ThDP-dependent enzymes (Fig. 3*b*). However, there are no packing interactions in the crystals that might suggest a higher oligomeric form for KdcA. The crystal structure is thus consistent with size-exclusion chromatography studies indicating that KdcA is a dimer in solution (Gocke *et al.*, 2007), in contrast to most other members of this enzyme family. A comparison of the areas involved in the tetramer interface in IPDC with KdcA shows very low sequence conservation; many of the smaller hydrophobic residues in IPDC are replaced by glutamate or lysine residues in KdcA, with flexible charged side chains commonly being found on protein surfaces.

3.3. Active site

The active-site residues of KdcA are highly conserved with respect to IPDC, including the two catalytic histidines (His112 and His113), an aspartic acid (Asp26) and a glutamic acid (Glu462) positioned close to the thiazolium ring (KdcA numbering). The active-site pocket, which is lined by polar residues on one side and hydrophobic side chains on the other, only displays one distinct difference compared with the IPDC structure: a leucine residue in the latter is replaced by a phenylalanine (Phe542) in KdcA. Replacement by this bulkier amino acid results in closure of the entrance to the active site, which in the IPDC holo structure is open and connects the active site to the surface.

3.4. Inhibitor complex

2-[(R)-1-Hydroxyethyl]deazaThDP is an inhibitor of many ThDP-dependent enzymes and mimics the covalent reaction intermediate 2-[(R)-1-hydroxy-ethyl]ThDP that occurs in the decarboxylation of pyruvate (Leeper *et al.*, 2005). Binding of this deazaThDP derivative totally abolished the activity of KdcA as a consequence of the impaired ylid formation of this compound. We determined the structure of the complex of KdcA with this analogue of the reaction intermediate to 1.8 Å resolution in order to obtain a template that was more suitable than the holoenzyme
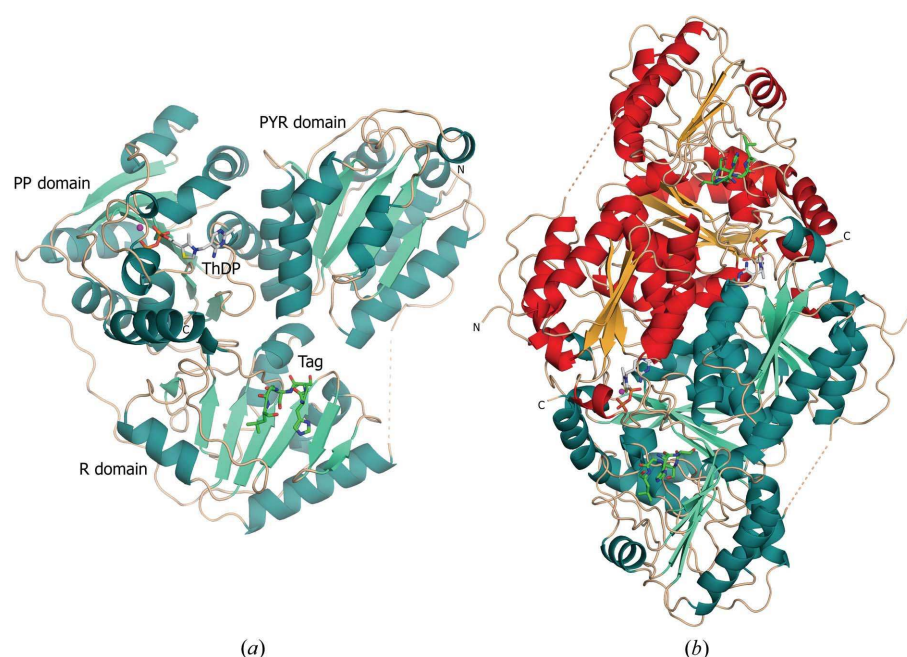
Figure 3
KdcA architecture: (*a*) monomer, (*b*) dimer. ThDP and the ordered N-terminal tag are shown as stick models with C atoms in grey and bright green, respectively. The Mg^{2+} ion is coloured magenta.
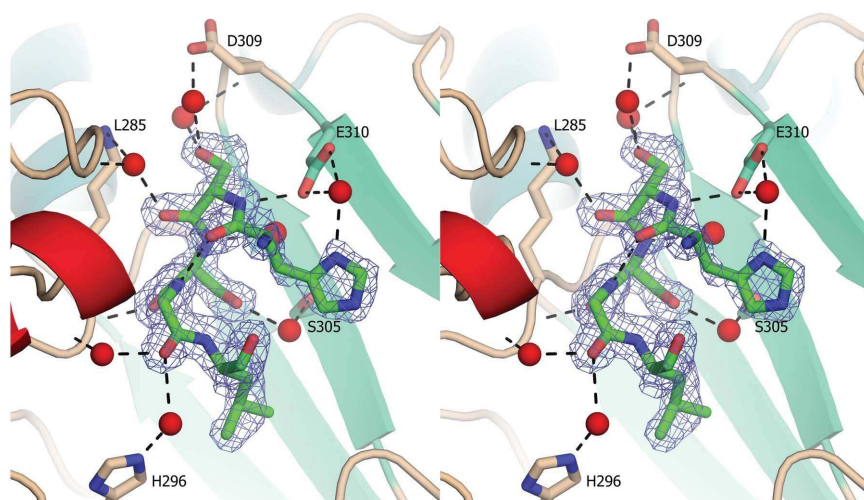
Figure 4
A stereoview of the interactions between the N-terminal tag and a crystallographically related molecule. Residues involved in hydrogen-bonding interactions with residues from the N-terminal tag are shown as stick models. All but two hydrogen bonds (indicated by dashed black lines) are mediated by ordered water molecules (red spheres). The $2F_o - F_c$ electron-density map is contoured at 1σ.

for modelling of substrate/intermediate complexes (Fig. 5). The overall structure of the enzyme in this complex is very similar to that of the holoenzyme with ThDP, with an r.m.s. deviation of 0.11 Å based on 541 C$^\alpha$ atoms. The only exceptions are the last residues of the C-terminal helix, which appear to be better ordered in the complex.

The position of the bound 2-[(R)-1-hydroxyethyl]deazaThDP ligand superimposes well with that of ThDP in the holoenzyme. The covalent hydroxyethyl moiety of the ligand extends into the active-site channel and the α-hydroxyl group forms a hydrogen bond to the 4′-amino group of the pyrimidine ring of ThDP. The geometry around the C$^\alpha$ atom of the hydroxyethyl moiety is slightly distorted from tetrahedral, with a smaller C2—C$^\alpha$—OH angle (99–100°) and a larger C2—C$^\alpha$—C$^\beta$ angle (117–119°). Furthermore, the C$^\alpha$ atom of the ligand lies slightly out of the thiazolium plane (26°). The out-of-plane feature of the intermediate has previously been observed in freeze-trapping experiments in which substrates were trapped as postdecarboxylation intermediates in the decarboxylase subunit of the human α-ketoacid dehydrogenase complex (Machius *et al.*, 2006) and oxalyl-CoA decarboxylase (Berthold *et al.*, 2007).

3.5. Carboligation features and modelling of substrates

The carboligation reactions catalyzed to various extents by ThDP-dependent decarboxylases are of particular interest for biocatalytic purposes since they allow the chemoenzymatic formation of C—C bonds in a stereoselectively defined manner. The identification of the features of these enzymes that determine substrate selectivity and stereoselectivity is therefore essential if the substrate spectrum and enantioselectivity of these enzymes are to be modified by rational design. The structure of KdcA provides insights into the structural basis of several unique properties of the enzyme with respect to carboligation.

The position of the bound aldehyde substrate that also determines the stereochemistry of the product is determined by two factors. One is the catalytically relevant orientation of the aldehyde carbonyl group, which has to point towards catalytic residues implied in proton-transfer steps in ThDP-dependent decarboxylases such as ZmPDC (Schenk *et al.*, 1997; Dobritzsch *et al.*, 1998),

BAL (Kneen *et al.*, 2005), IPDC (Schütz *et al.*, 2003) and BFD (Polovnikova *et al.*, 2003). In KdcA, His112 (corresponding to His115 in IPDC, His113 in ZmPDC and His70 in BFD) is structurally conserved and is most likely to have a similar
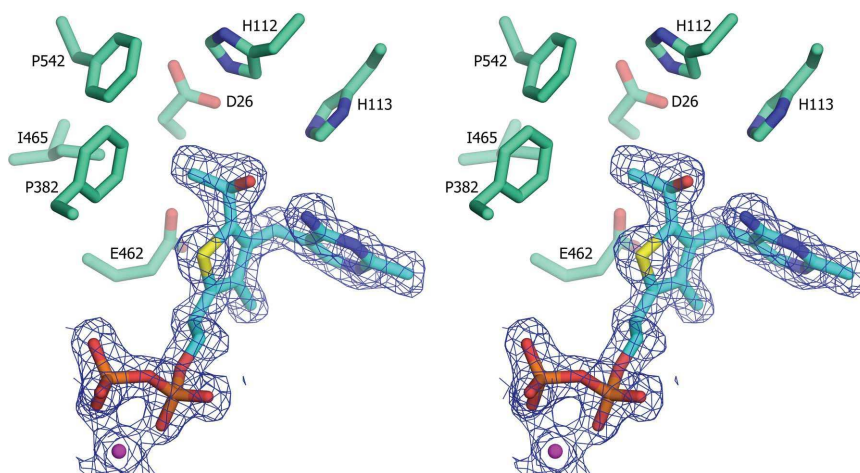
Figure 5
Part of the $2F_o - F_c$ electron-density map of the complex of KdcA with the bound analogue of the reaction intermediate 2-[(R)-1-hydroxyethyl]deazaThDP. The electron-density map at the positions of the bound ligand and Mg^{2+} ion is contoured at 1σ. Active-site residues in the proximity of the hydroxyethyl moiety are shown as stick models.
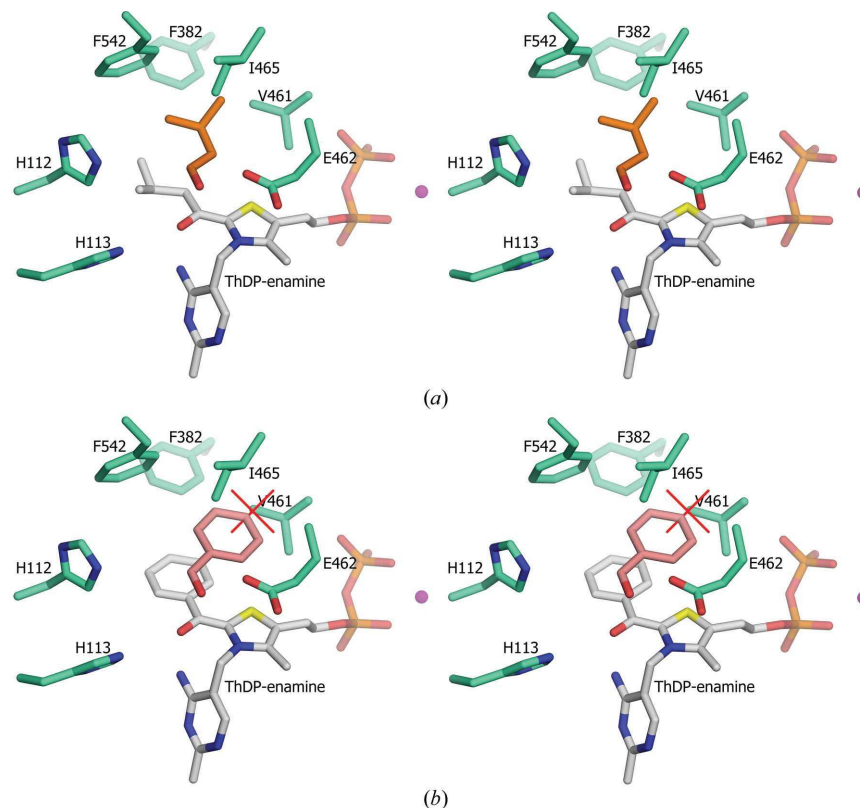
Figure 6
The S-pocket in KdcA controls enantioselectivity. (*a*) The self-ligation of isovaleraldehyde results in a slight excess of the (*S*) enantiomers (30–47%; Gocke *et al.*, 2007). Here, the acyl donor is modelled covalently bound to ThDP in the form of the enamine (grey) and the acceptor (orange) is fitted into the S-pocket with the *Re* face towards the enamine, resulting in the unusual (*S*)-enantiomer. (*b*) Self-ligation of benzaldehyde, on the other hand, results in pure (*R*)-benzoin. Modelling of the acceptor substrate (pink) into the S-pocket results in unavoidable clashes.

function in catalysis. Secondly, the aldehyde side chain can point (i) towards the substrate channel, resulting in an attack on the *Si* face of the carbonyl group (giving *R*-selectivity), or (ii) bind in the S-pocket, leading to attack on the *Re* face [*i.e.* (*S*)-enantiomer] (Fig. 2). Comparison of the S-pockets of various ThDP-dependent enzymes, defined in KdcA by residues Phe542, Ile465, Glu462 and Val461, and their accessibilities shows that KdcA has an average-sized S-pocket. BAL, on the other hand, has no S-pocket, BFD has a small pocket with a large entrance (substitution of Val461 in KdcA by alanine) and *Zm*PDC has an S-pocket about as large as KdcA, but the entrance is closed by a bulky isoleucine residue (Ile472). For *Zm*PDC it was shown that by replacing Ile472 at the entrance by site-directed mutagenesis the S-pocket could be made more accessible and the enantioselectivity was shifted from (*R*) to (*S*) (Siegert *et al.*, 2005). In a similar way, BFD was engineered to increase the size of the S-pocket, which is only accessible to acetaldehyde in the wild-type enzyme. The resulting BFD variants prefer propanal over acetaldehyde and produce (*S*)-2-hydroxy-1-phenylbutan-1-one (Gocke *et al.*, unpublished results). In KdcA, isovaleraldehyde is the only substrate for which the (*S*)-enantiomer of the corresponding 2-hydroxyketone can be produced to some extent (Gocke *et al.*, 2007). Modelling shows that the side chain of this substrate fits well into the S-pocket of the enzyme with a number of hydrophobic interactions (Fig. 6*a*), whereas the pocket is restricted in size owing to residues Ile465, Val461

and Phe542, which would clash with larger substrates such as benzaldehyde (Fig. 6*b*).

The substrate channel in KdcA, which is lined by Phe542 and Met538, is narrower than the wide and straight BFD channel but larger than that in *Zm*PDC, where a tryptophan residue (Trp392) limits the access to the active site. The small donor pocket of *Zm*PDC is known to optimally fit an acetaldehyde molecule, while KdcA on the other hand can adopt larger donor aldehydes (Gocke *et al.*, 2007). Indole-3-pyruvate, which after decarboxylation KdcA can ligate to acetaldehyde, was fitted into the structure of the holoenzyme as a model of the Michaelis complex. The substrate could be modelled straightforwardly such that the substrate C$^\alpha$ atom was aligned for nucleophilic attack by the cofactor, while at the same time favourable stacking interactions of the indole moiety were made to the side chains of Phe542 and Phe382 (Fig. 7). The shape of the substrate-binding pocket allows an optimal fit of the indole moiety, which makes KdcA the only biocatalyst identified so far that accepts substrates as large as indole-3-acetaldehyde in carboligation reactions.

KdcA displays the intriguing feature that aliphatic substrates larger than acetaldehyde only act as acyl donors when ligated with benzaldehyde as acceptor. In ligation reactions with acetaldehyde, however, benzaldehyde acts as an acyl donor in about 60% of the cases, whereas the larger substituted benzaldehyde 3,5-dichlorobenzaldehyde can only act as an acyl donor (Gocke *et al.*, 2007). These features of the carboligation reaction involving benzaldehyde and benzaldehyde derivatives can be explained by the structure of KdcA. Fig. 8 shows a model of the covalent ThDP–benzaldehyde adduct with a second benzaldehyde molecule bound in the acceptor position. The stacking interactions between the benzene rings favour binding of the second benzaldehyde molecule, *i.e.* the carboligation reaction between two benzaldehyde molecules yielding benzoin. Binding of a substituted benzene ring in the acceptor position, however, would result in several unfavorable interactions with surrounding hydrophobic residues, which prevents these compounds from acting as acceptors in the carboligation reaction. Indeed, the formation of benzoin derivatives from substituted benzaldehydes has not been observed with KdcA. In mixed carboligations with aliphatic and aromatic aldehydes, the nonplanarity of the aliphatic chains would result in clashes with either the donor benzene ring and/or the two bulky phenylalanine residues 382 and 542 and thus disfavour carboligation with aliphatic aldehydes as acceptors. Larger aliphatic aldehydes are therefore predominantly bound as acyl donors with benzaldehyde as acceptor.

Figure 7
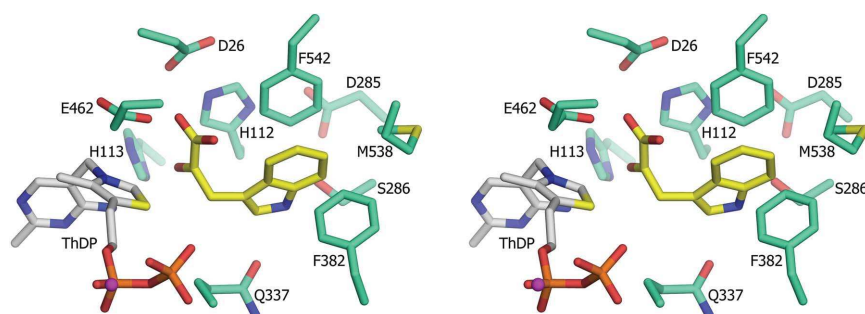Structural basis of chemoselectivity. Stereoview of the Michaelis complex of KdcA with the bound donor substrate indole-3-pyruvate. The substrate-binding pocket allows an optimal fit of the indole moiety.
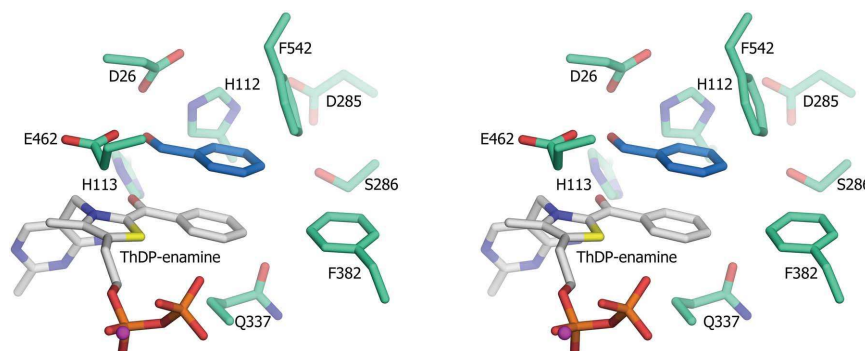

Figure 8
Benzaldehyde modelled as acyl donor aldehyde covalently bound to ThDP in the form of the enamine (grey) and as acceptor substrate (blue) with the *Si* side facing the enamine.

research papers

References

Berthold, C. L., Moussatche, P., Richards, N. G. & Lindqvist, Y. (2005). *J. Biol. Chem.* **280**, 41645–41654.

Berthold, C. L., Toyota, C. G., Moussatche, P., Wood, M. D., Leeper, F., Richards, N. G. & Lindqvist, Y. (2007). *Structure*, **15**, 853–861.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dobritzsch, D., König, S., Schneider, G. & Lu, G. (1998). *J. Biol. Chem.* **273**, 20196–20204.

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

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Gocke, D., Nguyen, C., Pohl, M., Stillger, T., Walter, L. & Müller, M. (2007). *Adv. Synth. Catal.* **349**, 1425–1435.

Hasson, M. S., Muscate, A., McLeish, M. J., Polovnikova, L. S., Gerlt, J. A., Kenyon, G. L., Petsko, G. A. & Ringe, D. (1998). *Biochemistry*, **37**, 9918–9930.

Iding, H., Dünnwald, T., Greiner, L., Liese, A., Müller, M., Siegert, P., Grötzinger, J., Demir, A. S. & Pohl, M. (2000). *Chemistry*, **6**, 1483–1495.

Janzen, E., Müller, M., Kolter-Jung, D., Kneen, M. M., McLeish, M. J. & Pohl, M. (2006). *Bioorg. Chem.* **34**, 345–361.

Kern, D., Kern, G., Neef, H., Tittmann, K., Killenberg-Jabs, M., Wikner, C., Schneider, G. & Hübner, G. (1997). *Science*, **275**, 67–70.

Kneen, M. M., Pogozheva, I. D., Kenyon, G. L. & McLeish, M. J. (2005). *Biochim. Biophys. Acta*, **1753**, 263–271.

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

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leeper, F. J., Hawksley, D., Mann, S., Perez Melero, C. & Wood, M. D. (2005). *Biochem. Soc. Trans.* **33**, 772–775.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Lindqvist, Y. & Schneider, G. (1993). *Curr. Opin. Struct. Biol.* **3**, 896–901.

Lobell, M. & Crout, D. H. G. (1996). *J. Am. Chem. Soc.* **118**, 1867–1873.

Machius, M., Wynn, R. M., Chuang, J. L., Li, J., Kluger, R., Yu, D., Tomchick, D. R., Brautigam, C. A. & Chuang, D. T. (2006). *Structure*, **14**, 287–298.

Mosbacher, T. G., Müller, M. & Schulz, G. E. (2005). *FEBS J.* **272**, 6067–6076.

Müller, M. & Sprenger, G. A. (2004). *Thiamine: Catalytic Mechanisms in Normal and Disease States*, edited by F. Jordan & M. S. Patel, pp. 77–91. New York: Marcel Dekker.

Muller, Y. A., Lindqvist, Y., Furey, W., Schulz, G. E., Jordan, F. & Schneider, G. (1993). *Structure*, **1**, 95–103.

Muller, Y. A. & Schulz, G. E. (1993). *Science*, **259**, 965–967.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Pang, S. S., Duggleby, R. G. & Guddat, L. W. (2002). *J. Mol. Biol.* **317**, 249–262.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Polovnikova, E. S., McLeish, M. J., Sergienko, E. A., Burgner, J. T., Anderson, N. L., Bera, A. K., Jordan, F., Kenyon, G. L. & Hasson, M. S. (2003). *Biochemistry*, **42**, 1820–1830.

Schenk, G., Leeper, F. J., England, R., Nixon, P. F. & Duggleby, R. G. (1997). *Eur. J. Biochem.* **248**, 63–71.

Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.

Schütz, A., Sandalova, T., Ricagno, S., Hübner, G., König, S. & Schneider, G. (2003). *Eur. J. Biochem.* **270**, 2312–2321.

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

Smit, B. A., van Hylckama Vlieg, J. E., Engels, W. J., Meijer, L., Wouters, J. T. & Smit, G. (2005). *Appl. Environ. Microbiol.* **71**, 303–311.

Sprenger, G. A. & Pohl, M. (1999). *J. Mol. Catal. B*, **6**, 145–159.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Yep, A., Kenyon, G. L. & McLeish, M. J. (2006). *Bioorg. Chem.* **34**, 325–336.

electronic reprint

Publication VII

Rational protein design of ThDP-dependent enzymes: engineering stereoselectivity.

Gocke, D., Walter, L., Gauchenova, E., Kolter, G., Knoll, M., Berthold, C. L., Schneider, G., Pleiss, J., Müller, M. & Pohl, M.

(Engineering stereoselectivity)

Rational Protein Design of ThDP-Dependent Enzymes: Engineering Stereoselectivity

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

Benzoylformate decarboxylase (BFD) from *Pseudomonas putida* is an exceptional thiamin diphosphate (ThDP)-dependent enzyme, as it catalyzes the formation of (S)-2-hydroxy-1-phenyl-propan-1-one from benzaldehyde and acetaldehyde. This is the only currently known S-selective reaction (92% ee) which is catalyzed by this otherwise R-selective class of enzymes. Here we describe the molecular basis to introduce S-selectivity into ThDP-dependent decarboxylases. By shaping the active site of BFD using rational protein design, structural analysis and molecular modeling, optimal sterical stabilization of the acceptor aldehyde in a structural element called S-pocket, as the predominant interaction to adjust stereoselectivity, was identified. Our studies revealed Leucine 461 as a hot spot for stereoselectivity in BFD. Exchange for alanine and glycine resulted in variants, which catalyze the S-stereoselective addition of larger acceptor aldehydes such as propanal with benzaldehyde and derivatives thereof; a reaction which is not catalyzed by the wild type enzyme. Crystal structure analysis of the variant BFDL461A supports the modeling studies.

Introduction

The potential of thiamin diphosphate (ThDP)-dependent enzymes to catalyze a benzoin condensation-like carboligation of aldehydes to chiral 2-hydroxy ketones with high stereoselectivity is well established.[1] Our goal is to generate a toolbox of various ThDP-dependent enzymes to create a platform for diversely substituted and enantio-complementary 2-hydroxy ketones.

With the current set of enzymes including benzoylformate decarboxylase (BFD), benzaldehyde lyase (BAL), branched-chain 2-ketoacid decarboxylase (KdcA), different pyruvate decarboxy-lases (PDC), and variants thereof, the carboligation of various aliphatic and aromatic aldehydes is possible yielding symmetrical and mixed (R)-2-hydroxy ketones with predominantly high enantioselectivity. However, the corresponding (S)-products are hardly accessible by these enzymes. Exceptions are the kinetic resolution of benzoin derivatives by BAL from *Pseudomonas fluorescens*,[2] and the carboligation of benzaldehyde derivatives and acetaldehyde yielding (S)-2-hydroxypropiophenone derivatives using BFD from *Pseudomonas putida* as a catalyst.[3]

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

Here, we verify the S-pocket approach by site-directed mutagenesis of amino acid residues lining this part of the active centre and probing the resulting variants in carboligation reactions with different aliphatic aldehydes as acyl acceptors. Further, we provide evidences that similar S-pockets are also

[a*]   D. Gocke, G. Kolter, PD Dr. M. Pohl
       Institute of Molecular Enzyme Technology
       Heinrich-Heine University Düsseldorf
       D-52426 Jülich (Germany)
       Fax: (+49) 2461-612940
       E-mail: ma.pohl@fz-juelich.de
       Homepage: www.iet.uni-duesseldorf.de
[b]   L. Walter, Dr. E. Gauchenova, Prof. Dr. M. Müller
       Institute of Pharmaceutical Sciences
       Albert-Ludwigs University Freiburg
       Albertstraße 25, D-79104 Freiburg (Germany)
[c]   M. Knoll, Prof. Dr. J. Pleiss
       Institute of Technical Biochemistry
       University of Stuttgart
       Allmandring 31, D-70569 Stuttgart (Germany)
[d]   C. L. Berthold, Prof. Dr. G. Schneider
       Department of Medical Biochemistry and Biophysics
       Karolinska Institutet
       Tomtebodavägen 6, S-17177 Stockholm (Sweden)

       Supporting information for this article is available on the WWW
       under http://www.chembiochem.org.

present in other ThDP-dependent decarboxylases, which opens access to a broad range of (S)-2-hydroxy ketones as valuable building blocks for compounds such as the taxol side chain and 5'-methoxyhydnocarpin.[6]
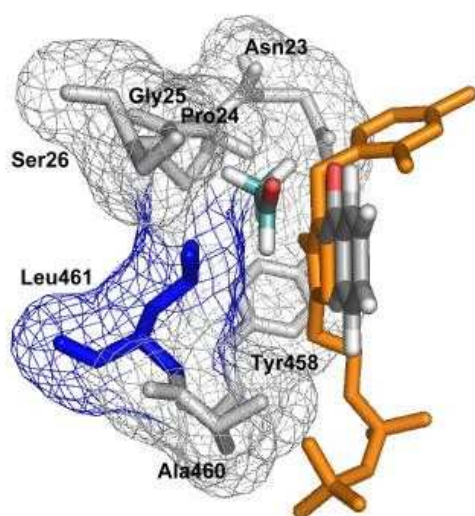

Figure 1. Active site of BFD with the cofactor ThDP (orange). The donor benzaldehyde (grey) and the acceptor acetaldehyde (light blue) are modeled inside. The side chain of acetaldehyde is bound to the S-pocket, which is mainly defined by Leu 461 (blue). The diametrically opposed orientation of the side chains of donor and acceptor results in the formation of (S)-products.[6]

Results and Discussion

The S-pocket in BFD from *Pseudomonas putida* is formed by the side chains of P24, A460, and predominantly by L461 (Figure 1). These residues were replaced by smaller amino acid residues in order to dissect the impact of the respective amino acid on an increase of the S-pocket. All variants were produced by site-directed mutagenesis. After cloning, overexpression, and purification the variants were investigated with respect to their decarboxylase and carboligase activity.

Decarboxylase activity of BFD-variants

All variants were still able to catalyze the physiological function of BFD: the decarboxylation of benzoylformate (Tab. 1), showing hyperbolic v/[S]-plots as wild type BFD (BFDwt) but a decreased decarboxylase activity toward benzoylformate. Most pronounced effects were obtained with the mutation in position Leu461 yielding variants with 7 to 8-fold decreased specific decarboxylase activity compared to BFDwt. It is important to note that the K_M-values do not parallel the V_{max} values. Whereas BFDP24A and BFDL461A show K_M-values in the same range as the wild type enzyme, the variant BFDL461V has a 6-fold higher apparent affinity for the substrate, while the K_M-values of the two glycine variants (A460G, L461G) are about 4 to 5-fold higher than that of BFDwt.

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

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

Data were measured in 50 mM potassium phosphate buffer, pH 6.5. Kinetic constants were calculated according to Michaelis-Menten using Origin 7.0 (Origin Lab Corporation, Northampton, MA).

The substrate range of the decarboxylase reaction has not been affected by the mutations, with benzoylformate being the main substrate for all variants (see supporting information).

Carboligase activity of BFD-variants

All variants were investigated with respect to their carboligase activity toward the self-ligation of benzaldehyde to benzoin and acetaldehyde to acetoin (Tab. 2). Mutations in the putative S-pocket effect both, the acetoin and the benzoin synthesis, with the P24A- and the L461V variant being most effective concerning catalysis of the acetoin synthesis and the L461V variant catalyzing also the benzoin synthesis 1.6 times faster than BFDwt. The mutations affect the stereoselectivity of the acetoin formation significantly (Tab. 2): While BFDwt catalyzes predominantly the formation of (R)-acetoin (**1**, Tab. 3) (ee 34%),[7] the (S)-enantiomer (ee 65%) is formed in excess with BFDL461A, supporting the relevance of L461 for the shape of the S-pocket.

Table 2. Space-time yields and enantioselectivities of different BFD variants concerning the formation of acetoin from acetaldehyde and benzoin from benzaldehyde.
n.d.: not determined.

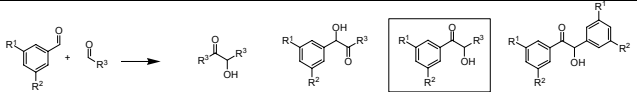| enzyme	acetoin [g/L/d]	ee [%]	benzoin [g/L/d]	ee [%]
BFDwt	0.11	34 (R)	1.08[b]	n.d.
BFDP24A	0.19	3 (R)	0.58[b]	n.d.
BFDA460G	0.07	38 (R)	0.23[b]	n.d.
BFDL461V	0.1	33 (R)	1.75[a]	n.d.
BFDL461A	0.07	64 (S)	0.14[a]	n.d.
BFDL461G	n.d.	n.d.	n.d.	n.d.

Space-time-yields were calculated within the linear range of [a] 3 h, [b] 2 h.

Further studies with mixed carboligations were focused on the glycine and alanine variants in position of leucine 461. In mixed carboligations with benzaldehyde and acetaldehyde the stereoselectivity of the (S)-2-hydroxy-1-phenyl-propan-1-one ((S)-HPP, **3**) synthesis was improved by both mutations (ee 98%) relative to BFDwt (92% (Table 3, A).[3] HPP is formed with benzaldehyde as the donor and acetaldehyde as the acceptor aldehyde (Figure 1). As predicted from modeling studies the most pronounced effect of the increased S-pocket should become apparent, if acetaldehyde is replaced by the larger propanal, which was proved by analytical studies (Table 3, B). Both variants in position 461 catalyze the synthesis of the desired product (S)-2-hydroxy-1-phenyl-butan-1-one ((S)-**7**) with high stereoselectivity, showing even higher enantioselectivities at pH 7.9 in comparison to standard conditions (pH 7.0), whereas variation of the substrate concentration and reaction temperature had no significant effect on the ee-values (data not shown).

According to the predictions based on the structure of BFDL461A a further increase of the acceptor aldehyde should decrease the stereoselectivity again. This assumption was experimentally confirmed by application of butanal in mixed carboligation reactions with benzaldehyde yielding **10** with decreased stereoselectivity (ee 63%, (S)) (Table 3, C) and activity. Besides propanal, monomethoxyacetaldehyde is an acceptor aldehyde in presence of benzaldehyde for BFDL461A, resulting in a yield of about 20% in analytical biotransformations (Table 3, D). By contrast, cyclopropylcarbaldehyde results in lower amounts yet high selectivity of the mixed carboligation product (Table 3, E).

Table 3. Relative product distribution and enantiomeric excesses (ee) obtained in analytical scale carboligation reactions of benzaldehyde derivatives and various aliphatic aldehydes catalyzed by BFD variants. Relative product distribution is given in mol% (NMR); ee-values were determined by HPLC-analysis.
All studies were performed with equimolar concentrations of both aldehydes (18 mM) in potassium phosphate buffer (50 mM, pH 7, 2.5 mM $MgSO_4$, 0.1 mM ThDP, 20 vol% DMSO) 0.3 mg/mL purified enzyme, 30 °C, if not otherwise indicated.

A — $R^1 = R^2 = H$, $R^3 = CH_3$

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

B — $R^1 = R^2 = H$, $R^3 = C_2H_5$

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

C — $R^1 = R^2 = H$, $R^3 = C_3H_7$

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

D — $R^1 = R^2 = H$, $R^3 = OCH_3$

D	products			
	11	12	13	4
	BFDL461A			
	-	-	9.5%, 21%[b] 93% (S)	<1% ee n.d.

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

E	products			
	15	16	17	4
	BFDL461A			
	-	-	5% 97.5% (S)	-

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

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

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

G	products			
	21	22	23	20
	BFDwt			
	19% ee n.d.	26% ee n.d.	-	-
	BFDL461A			
	-	-	-	-
	BFDL461G			
	2% ee n.d.	26% ee n.d.	4% ee n.d.	-

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

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

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

I	products			
	1	27	28	29
	BFDL461A			
	-	-	22.5%, 60%[b] 98% (S)	

[a] Carboligations have been performed at pH 7.9.
[b] Carboligations have been performed with a 3-fold excess of the aliphatic aldehyde (54 mM).

Compared to benzaldehyde the stereo control in mixed carboligations with propanal is even better with 3,5-dimethoxy-benzaldehyde and 3-cyanobenzaldehyde (Table 3, F, I). Very interesting results have been obtained in carboligations of 3,5-dimethoxybenzaldehyde with butanal and pentanal, respectively. Whereas no product was obtained with the variant BFDL461A, the variant BFDL461G is able to catalyze the mixed carboligation with pentanal even better than with butanal (Table 3, G, H).

The carboligation of benzaldehyde and propanal was investigated in more detail on a preparative scale. After 70 h BFDL461A yielded 35% (w/w) and BFDL461G 31% (w/w) product **7**, respectively, while negligible amounts of **4** and **5** were formed demonstrating very high chemoselectivity for both variants. Confirming the prediction of the modeling studies, the desired product (*S*)-2-hydroxy-1-phenyl-butan-1-one ((*S*)- **7**) was formed with very high stereoselectivity of 93-97% using the variants BFDL461A and BFDL461G, which is in contrast to BFDwt yielding (*R*)-**7** (*ee* 21%) (Table 4). Although carboligation with BFDwt resulted in a total conversion of 37%, a mixture of **4, 6,** and **7** was formed in almost equal amounts, demonstrating the low chemo- and stereoselectivity of BFDwt for this reaction.

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

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

[a] Numbers refer to Table 3. Product composition is given in mol% (NMR); *ee*-values were determined by chiral HPLC.

Structural investigation of BFDL461A

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

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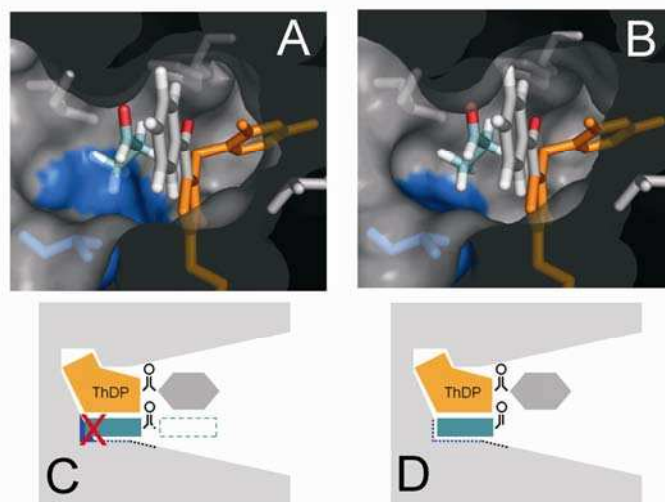

Figure 2. Crystal structure of the active site of BFDwt (A) and BFDL461A (B) with benzaldehyde and propanal modeled inside. The side chain of the amino acid residue in position 461 is marked in blue. Benzaldehyde (grey) is bound to the C2-atom of the thiazolium ring (orange) and is coplanarly arranged due to steric and electronic demands. Propanal (light blue) is located in the *S*-pocket. Models show a perfect stabilization of the acceptor aldehyde in the *S*-pocket of the variant (B, D), while L461 causes steric hindrance with propanal in BFDwt (A, C). Consequently, in BFDwt the propanal approaches predominantly parallel to benzaldehyde (dotted square, C) yielding mainly the (*R*)-enantiomer while in BFDL461A the perfect fitting allows an antiparallel arrangement (D) leading to an excess of the (*S*)-product.

Conclusion

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

A superimposition of the crystal structures of BFDwt[5], BAL from *Pseudomonas fluorescens,*[8] PDC from *Zymomonas mobilis* (*Zm*PDC),[9a] *Saccharomyces cerevisiae* (*Sc*PDC)[9b] as well as the recently solved structures of PDCs from *Acetobacter pasteurianus* (*Ap*PDC)[10] and KdcA from *Lactococcus lactis* [11, 12, 13] gave profound insights. While there is no *S*-pocket visible in BAL, the *S*-pockets of the other enzymes increase in the series *Pp*BFD < *Ll*KdcA < *Zm*PDC / *Sc*PDC < *Ap*PDC (Figure 3).

However, the entrance of the *S*-pocket in KdcA and in both PDCs is restricted by bulky residues, such as isoleucine or valine, which could explain, why all these enzymes are strictly *R*-selective. Consequently, (*S*)-2-hydroxy ketones could be formed by improving the access to this pocket. This has successfully been shown with the variant *Zm*PDCI472A which catalyzes the formation of (*S*)-HPP (**3**) (*ee* 70%) while exclusively (*R*)-phenylacetylcarbinol (**2**) (*ee* > 98%) is formed with the wild type enzyme using benzaldehyde and acetaldehyde as substrates.[14]
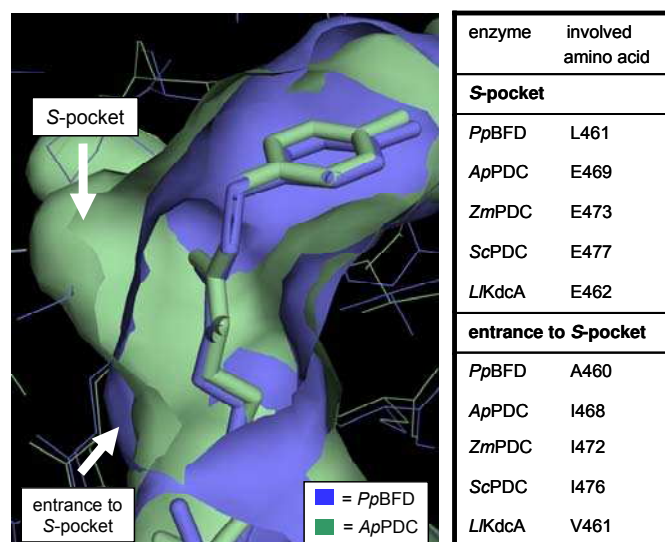
enzyme	involved amino acid
S-pocket	
*Pp*BFD	L461
*Ap*PDC	E469
*Zm*PDC	E473
*Sc*PDC	E477
*Ll*KdcA	E462
entrance to S-pocket	
*Pp*BFD	A460
*Ap*PDC	I468
*Zm*PDC	I472
*Sc*PDC	I476
*Ll*KdcA	V461

Figure 3. Superimposition of the *S*-pockets of BFDwt and *Ap*PDC. Compared to BFD *Ap*PDC shows an enlarged *S*-pocket but the entrance is blocked by residue I468 (left). The amino acids mainly bordering the *S*-pocket as well as those defining the entrance to the pocket are given for BFD, *Ap*PDC and additionally for *Zm*PDC, *Sc*PDC and KdcA (right).

The predominant *R*-selectivity of ThDP-dependent enzymes is therefore soundly explained mainly from their structures. However, many of these enzymes have the latent inherent property of *S*-selectivity, since such *S*-pockets are visible in almost all 3D structures mentioned above, although they are not accessible in many cases. Our results pave the way to expand the shaping strategy of the *S*-pockets to a broad range of other 2-ketoacid decarboxylases. This is a powerful tool to enlarge the toolbox of enzymatically accessible 2-hydroxy ketones by *S*-enantiomers, therewith providing a valuable platform for chemo-enzymatic synthesis.

Experimental Section

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

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

For enzyme crystallization of BFDL461A the purification protocol was kept but buffers were changed (Ni-NTA-chromatography: disintegration buffer (Mes/NaOH, 50 mM, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP), washing buffer (Mes/HCl, 50 mM, pH 7.0, 50 mM imidazole), elution buffer (Mes/HCl, 50 mM, pH 7.0, 250 mM imidazole); G25-chromatography (Mes/NaOH, 20 mM, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP)). Concentration of enzyme solutions was performed in vivaspin 20 centrifuge columns (Sartorius, cut off 10 kDa) up to 130 mg/mL. Superdex G200-size (GE Healthcare) exclusion chromatography showed a purity of 97.5% of the tetrameric BFDL461A (2% dimeric BFDL461A, 0.5% impurity). For storage the enzyme solution was shock-frozen in liquid nitrogen and kept at -20 °C.

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

Protein determination was performed according to Bradford[16] using bovine serum albumin (BSA) as the standard.

Benzoin syntheses: *Reaction conditions*: benzaldehyde (20 mM), DMSO (20 vol%), BFD variant (0.3 mg/mL), potassium phosphate buffer (50 mM, pH 7.5, 2.5 mM MgSO$_4$ and 0.1 mM ThDP), 30°C, 100 rpm. To avoid evaporation of the aldehydes the reaction batch was divided into GC-vials with a volume of 400 µL, after taking the starting sample. The reaction was stopped by addition of acetonitrile (400 µL) followed by intense vortexing and centrifugation of the precipitate. Calibration curves with benzoin were prepared in the same way. Conversions were determined by HPLC, employing a Dionex HPLC (Germering), equipped with a 250 x 4,6 Multohyp ODS-5 µ (CS-Chromatography) and an UV-detector (mobile phase 60% (v/v) H$_2$O: 40% (v/v) acetonitrile, flow 1.1 mL/min, pressure 130 bar, 20 µL injection volume, detection λ = 250 nm), R_t(benzoin) = 32.2 min.

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

Mixed carboligations of benzaldehyde and different aliphatic aldehydes:

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

Preparative scale synthesis: *Reaction conditions*: 15 mL scale: benzaldehyde (29 mg, 0.27 mmol) and propanal (16 mg, 0.27 mmol) were dissolved in DMSO (3 mL). After addition of potassium phosphate buffer (50 mM, 12 mL, pH 7.9, 2.5 mM MgSO$_4$, 0.1 mM ThDP) the reaction was started with the purified BFD variant (4.5 mg)

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

Analysis of (*S*)-2-hydroxy-1-phenylbutan-1-one ((*S*)-7).
HPLC: (chiral OD-H, n-hexane: 2-propanol, 95:5, 0.5 mL min^{-1}, 40°C): R_t (*S*) =11.8 min, R_t (*R*) = 13.9 min; $[\alpha]_D^{21.4}$ = -11.52 (c = 0.4, $CHCl_3$); CD (acetonitrile): λ [nm] (mol. CD) = 298 (-0.3250), 281 (-1.7045), 239 (6.2025), 206 (-7.6004); ^{1}H NMR (400 MHz, $CDCl_3$, 300 K, ppm): ꭥ= 0.96 (t, 3H, ^{3}J (H, H) = 7.4 Hz, CH_3),1.63 (dqd, 1H, ^{2}J (H, H) = 14.3 Hz, ^{3}J (H, H) = 7.4 Hz, 7.3 Hz, CH_2), 1.98 (dqd, 1H, ^{2}J(H, H) = 14.3 Hz, ^{3}J (H, H) = 7.4 Hz, 3.8 Hz, CH_2), 3.73 (d, 1H, ^{3}J (H, H) = 6.4 Hz, OH), 5.08 (ddd, 1H ,^{3}J (H, H) = 7.3 Hz, 6.4 Hz, 3.8 Hz, CHOH), 7.52 (ddm, 2H, ^{3}J (H, H) = 7.4 Hz, Ar-H), 7.64 (ddm, 1H, ^{3}J(H, H) = 7.4 Hz, Ar-H), 7.93 (dm, 2H, ^{3}J (H, H) = 7.4 Hz, Ar-H); ^{13}C NMR (100 MHz, $CDCl_3$, 300 K, ppm): ꭥ= 8.8 (CH_3), 28.8 (CH_2), 73.9 (CHOH), 128.4 (2 x CH_{Ar}), 128.8 (2 x CH_{Ar}), 133.3 (CH_{Ar}), 202.1 (CO); GCMS R_t = 8.7 min; MS (70 eV, EI): m/z (%): 164 (0.1%) [M$^+$], 105 (100%), 77 (46%).

Crystallization. BFDL461A was crystallized using the hanging drop vapor diffusion method. Droplets were set up for crystallization by mixing of protein solution (2 µL) containing protein (13 g/mL) in Mes/NaOH (20 mM, pH 7.0, 2.5 mM $MgSO_4$, 0.1 mM ThDP) and 2 µL of the reservoir solution. Screening and optimization revealed a reservoir solution of 18-24% (w/v) PEG 2000 MME, 0.1 M sodium citrate, pH 5.2-5.8, 100-150 mM $(NH_4)_2SO_4$ to be best. After equilibration for 3 days diffraction-quality crystals appeared.

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

Structure solution and crystallographic refinement. The structure of BFDL461A was determined by molecular replacement using the program MOLREP.[18,19] The BFDwt structure (pdb accession code: 1BFD)[5] was used as search model to place the four monomers into the asymmetric unit. Atomic positions and B-factors of the model were refined by the maximum likelihood method in REFMAC5[18,20] which was interspersed with rounds of manual model building in COOT.[21] Non-crystallographic symmetry restrains were initially applied but released toward the end of refinement. Water assignment was performed in COOT where the model was validated as well. The quality of the final structure was examined using PROCHECK.[18, 22] Statistics of the refinement and final model are given in the supplementary material. The coordinates of BFDL461A, in addition to the structure factors, have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Databank PDB with the accession code **2v3w**.

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

Resolution [Å]	2.2 (2.32-2.2)*
No. of observations	321349 (46653)
No. of unique reflections	111569 (16561)
Completeness [%]	97.0 (99.2)
Multiplicity [%]	2.9 (2.8)
Mean $\{I/\sigma(I)\}$	8.8 (2.2)
Wilson B factor [Å^2]	27.8
R_{merge} [%]	12.4 (45.8)

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

Acknowledgements

[1]   M. Pohl, G. A. Sprenger, M. Müller, *Curr. Opin. Biotechnol*. **2004**, *15*, 335-342.

[2]   A. S. Demir, M. Pohl, E. Janzen, M. Müller, *J. Chem. Soc. Perkin Trans. 1* **2001**, 633-635.

[3]   a) H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir, M. Pohl, *Chem. Eur. J.* **2000**, *6*, 1483-1495.; b) B. Lingen, D. Kolter-Jung, P. Dünkelmann, R. Feldmann, J. Grötzinger, M. Pohl, and M. Müller **2003** *Chembiochem 4*, 721-726

[4]   M. Knoll, M. Müller, J. Pleiss, M. Pohl, *Chembiochem* **2006**, *7*, 1928-1934.

[5]   a) M. S. Hasson, A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A. Gerlt, G. L. Kenyon, G. A. Petsko, D. Ringe, *Biochemistry* **1998**, *37*, 9918-9930; b) E. S. Polovnikova, M. J. McLeish, E. A. Sergienko, J. T. Burgner, N. L. Anderson, A. K. Bera, F. Jordan, G. L. Kenyon, M. S. Hasson, *Biochemistry* **2003**, *42*, 1820-1830.

[6]   a) K. C. Nicolaou, Z. Yang, J. J. Liu, H. Ueno, P. G. Nantermet, R. K. Guy, C. F. Claiborne, J. Renaud, E. A. Couladouros, K. Paulvannan, E. J. Sorensen, *Nature* **1994**, *367*, 630-634; b) F. R. Stermitz, P. Lorenz, J. N. Tawara, L. A. Zenewicz, K. Lewis, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1433-1437.

[7]   P. Domínguez de María, M. Pohl, D. Gocke, H. Gröger, H. Trauthwein, T. Stillger, L. Walter, M. Müller, *Eur. J. Org. Chem.* **2007**, 2940-2944.

[8] T. G. Mosbacher, M. Müller, G. E. Schulz, *FEBS J.* **2005**, *272*, 6067-6076.

[9] a) D. Dobritzsch, S. König, G. Schneider, G. Lu, *J. Biol. Chem.* **1998**, *273*, 20196-20204; b) P. Arjunan, T. Umland, F. Dyda, S. Swaminathan, W. Furey, M. Sax, B. Farrenkopf, Y. Gao, D. Zhang, F. Jordan, *J. Mol. Biol.* **1996**, *256*, 590-600.

[10] D. Gocke, C. Berthold, T. Graf, H. Brosi, I. Frindi-Wosch, M. Knoll, T. Stillger, L. Walter, M. Müller, J. Pleiss, G. Schneider, M. Pohl, *submitted*.

[11] A. Yep, G. L. Kenyon, M. J. McLeish, *Bioorg. Chem.* **2006**, *34*, 325-336.

[12] D. Gocke, C. L. Nguyen, M. Pohl, T. Stillger, L. Walter, M. Müller, *Adv. Synth. Catal.* **2007**, *349*, 1425-1435.

[13] C. Berthold, D. Gocke, M. D. Wood, F. Leeper, M. Pohl, G. Schneider, *Acta Crystallogr. Sec. D: Biol. Crystallogr., accepted*.

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

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

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

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

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

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

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

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

[22] R. A. Laskowsky, M. W. McArthur, D. S. Moss, J. M. Thornton, *J. Appl. Chrystallogr.* **1993**, *26*, 283-291.

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

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

Supplementary Material:

Rational Protein Design of ThDP-dependent Enzymes: Engineering Stereoselectivity

Dörte Gocke, Lydia Walter, Ekaterina Gauchenova, Geraldine Kolter, Michael Knoll, Catrine L. Berthold, Gunter Schneider, Jürgen Pleiss, Michael Müller, Martina Pohl*

1. Mutagenesis Primers:

	5'-upstream primer -3'	5'-downstream primer-3'
Table 1. Primers used for QuikChange® Site-Directed Mutagenesis. Codons carrying the mutations are printed bold and italic.		
BFDP24A	ACG GTC TTC GGC AAT ***GCC*** GGC TCG AAC GAG CTC	GAG CTC GTT CGA GCC ***GGC*** ATT GCC GAA GAC CGT
BFDA460G	AAC GGC ACC TAC GGT ***GGC*** TTG CGA TGG TTT GCC	GGC AAA CCA TCG CAA ***GCC*** ACC GTA GGT GCC GTT
BFDL461V	GGC ACC TAC GGT GCG ***GTG*** CGA TGG TTT GCC GGC	GCC GGC AAA CCA TCG ***CAC*** CGC ACC GTA GGT GCC
BFDL461A	GGC ACC TAC GGT GCG ***GCC*** CGA TGG TTT GCC GGC	GCC GGC AAA CCA TCG ***GGC*** CGC ACC GTA GGT GCC
BFDL461G	GGC ACC TAC GGT GCG ***GGC*** CGA TGG TTT GCC GGC	GCC GGC AAA CCA TCG ***GCC*** CGC ACC GTA GGT GCC

2. Substrate range for the decarboxylation reaction

The substrate range of the BFD variants concerning the decarboxylase reaction are not affected by the mutations with benzoylformate as the predominant substrate.

1.   pyruvate
2.   2-oxobutyrate
3.   2-oxopentanoate
4.   2-oxohexanoate
5.   2-oxooctanoate
6.   2-oxoisohexanoate
7.   2-oxo-4-methylhexanoate
8.   2-oxo-3-methylpentanoate
9.   2-oxo-3-fluoropropanoate
10. benzoylformate
11. phenylpyruvate
12. 2-oxo-4-phenylbutanoate
13. 2-oxo-5-phenylpentanoate
14. 3-indolepyruvate
15. 2-oxo-3-methylbutanoate
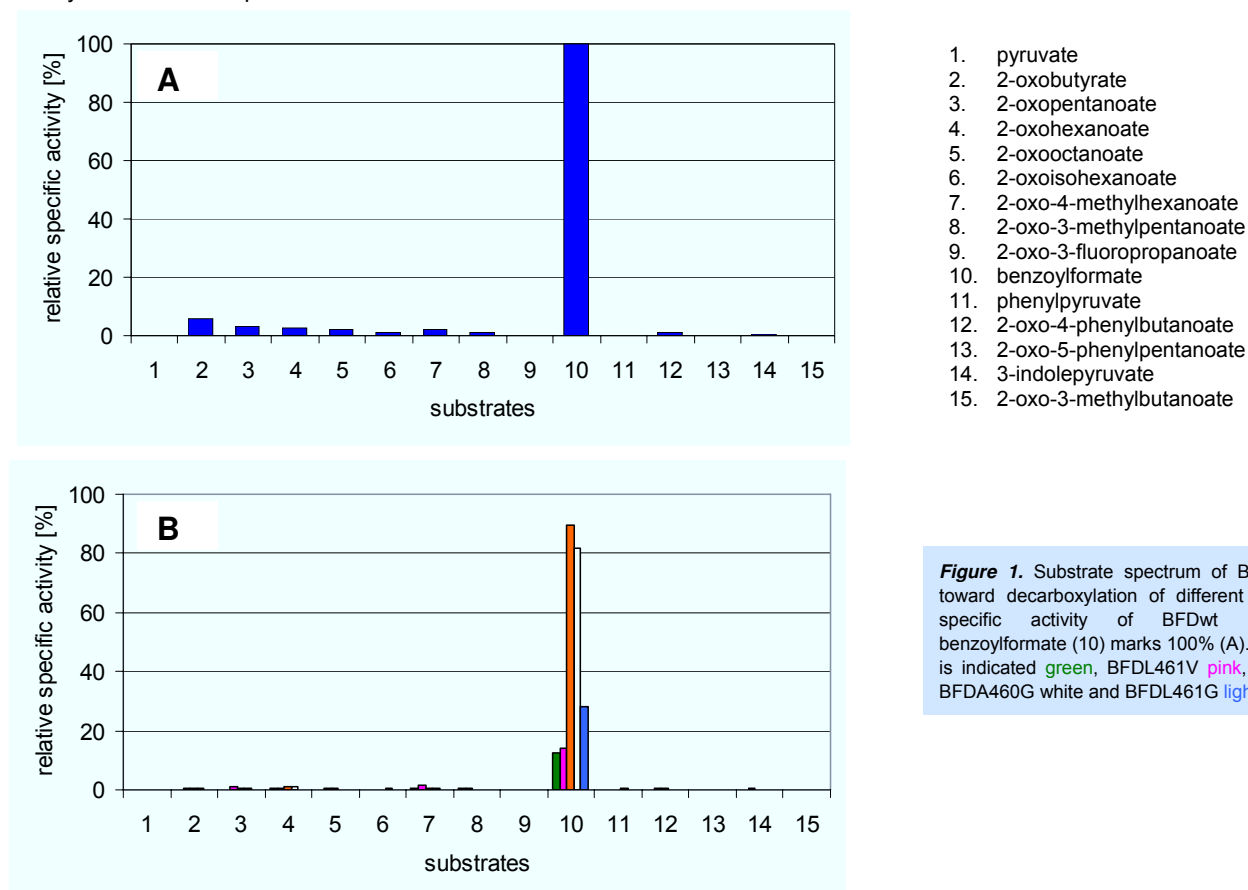
Figure 1. Substrate spectrum of BFDwt and variants toward decarboxylation of different 2-keto acids. The specific activity of BFDwt dark-blue toward benzoylformate (10) marks 100% (A). Variant BFDL461A is indicated green, BFDL461V pink, BFDP24A orange, BFDA460G white and BFDL461G light-blue (B).

3. Structure refinement and final model statistics of BFDL461A.

Table 2. Structure refinement and final model statistics of BFDL461A.	
Refinement program	REFMAC5
Reflections in working set	100 276
Reflections in test set	5 610
R-factor	18.0%
R-free	22.8%
Atoms modeled	16914
Number/Average B-factor	
Amino acids	4 x 526 / 26.6 Å^2
Ligands	4 / 24.9 Å^2
Magnesium ions	6 / 20.1 Å^2
Sulphate ions	4 / 32.2 Å^2
Waters	949 / 29.8 Å^2
RMS Deviations from ideals	
Bonds	0.008 Å
Angles	1.125 °
Ramachandran plot	
Most favored	92%
Additionally allowed	8%
Generously allowed	0%
Disallowed	0%

3 GENERAL DISCUSSION

3.1 An assay for rapid substrate range screenings

In order to identify enzymes with carboligation potential for a selected reaction an appropriate fast and easy assay is essential as a pre-screen prior to detailed instrumental analysis concerning the chemo- and enantioselectivity.

As the goal of this work was to expand the range of enzymatically accessible chiral 2-hydroxy ketones, an appropriate screening assay for a broad range of aliphatic and aromatic 2-hydroxy ketones was developed (***publication I***) based on a previously described high-throughput assay for the detection of phenylacetylcarbinol (PAC) (Breuer *et al.*, 2002, Fig. 3-1). Most importantly it should give information about the substrate range of the respective enzyme and should allow comparison of relative activities.
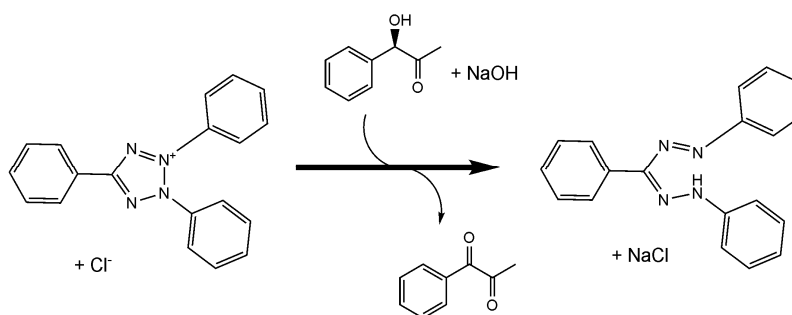
Fig. 3-1: Tetrazolium chloride (TCC) assay for the detection of 2-hydroxy ketones. In alkaline solution (*R*)-phenylacetylcarbinol (PAC) reduces 2,3,4-triphenyltetrazolium chloride (TTC) to the respective red formazane (according to Breuer *et al.*, 2002).

3.1.1 Sensitivity towards various 2-hydroxy ketones

Linear aliphatic and aromatic 2-hydroxy ketones can be detected with high sensitivity
As demonstrated in Fig. 3-2, the detection limits for aromatic 2-hydroxy ketones are lower (0.5-1 mM) than for aliphatic products. As product concentrations of 5-9 mM can be expected under standard conditions (***publication I***), this assay is useful to detect almost all expected products investigated in this work with sufficient sensitivity.

Some branched-chain aliphatic 2-hydroxy ketones require adaptation of the assay
The enzymatic synthesis of branched-chain aliphatic 2-hydroxy ketones was of special interest due to their potential as building blocks (chapter 3.3.1, ***publication II***). As demonstrated in Fig. 3-2 II, the sensitivity of the TTC-assay is reduced for branched-chain products such as 5-hydroxy-(2,7-dimethy)-octan-4-on (DMO), although the detection limit of 5 mM is still tolerable. In contrast, the detection limit of 4-hydroxy-(2,5-dimethyl)-hexane-3-one (DMH) is > 40 mM and thus too high to be detectable by this TTC-assay

(*publication I*) under standard conditions. The problem might be the branch directly positioned next to the hydroxy ketone group, which most likely influences the redox potential of DMH, limiting the reduction of TCC. In order to improve the sensitivity, various tetrazolium salts were tested to identify an appropriate reagent for the detection of DMH. Of the several commercially available tetrazolium salts (Fig. 3-3 I) the lowest detection limits of about 2 mM were observed with INT, MTT and NTB (data not shown). However, in all cases the substrate isobutyric aldehyde (IBA) interfered with the DMH detection resulting in very high background colouration. Thus, nitrotetrazolium chloride (NTC) with a detection limit of 10 mM revealed to be the only possible alternative.
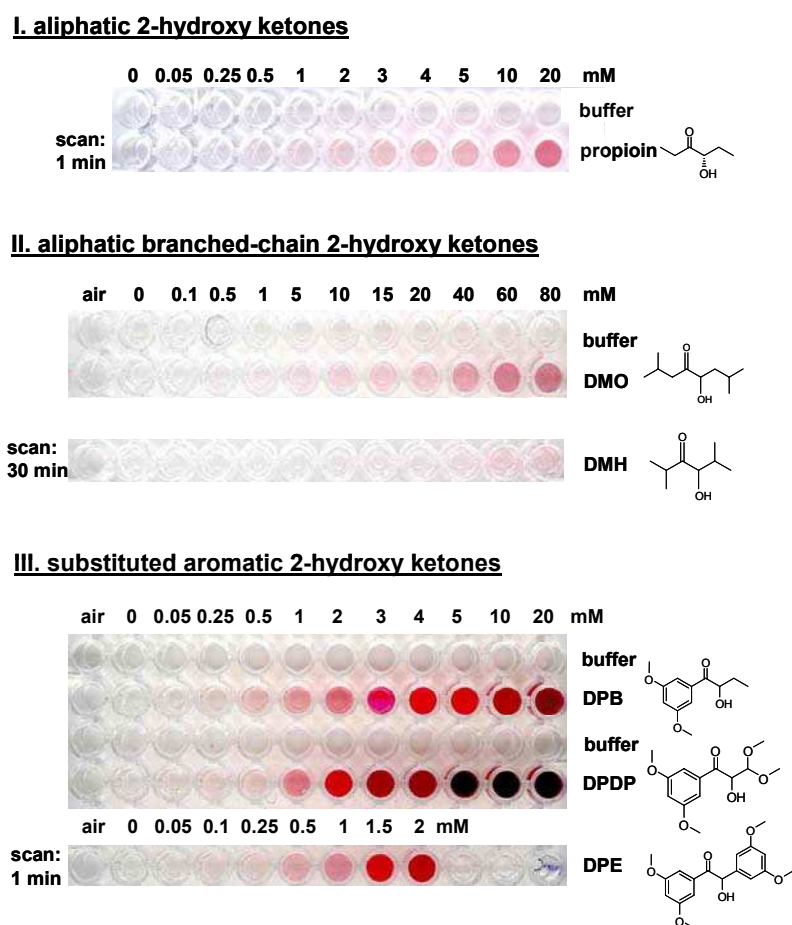
Fig. 3-2: Investigation of the detection limits of various linear aliphatic (I), branched-chain aliphatic (II)* and aromatic 2-hydroxy ketones (III)* in the TTC-assay. Conditions: 2-hydroxy ketones were dissolved in 50 mM potassium phosphate buffer, pH 7, 2.5 mM $MgSO_4$, 0.1 mM ThDP, 20% (v/v) DMSO. The TTC-assay was started by addition of a mixture of 10 µL TTC and 30 µL NaOH, results were scanned 1 min later. Exceptional are branched-chain 2-hydroxy ketones, requiring 30 min for complete colour development. DMO = 5-hydroxy-(2,7-dimethyl)-octan-4-one, DMH = 4-hydroxy-(2,5-dimethyl)-hexan-3-one, DPB = 1-(3,5-dimethoxy-phenyl)-2-hydroxy-butan-1-one, DPE = 1,2-bis-(3,5-dimethoxy-phenyl)-2-hydroxy-ethanone, DPDP = 1-(3,5-dimethoxy-phenyl)-2-hydroxy-3,3-dimethoxy-propan-1-one
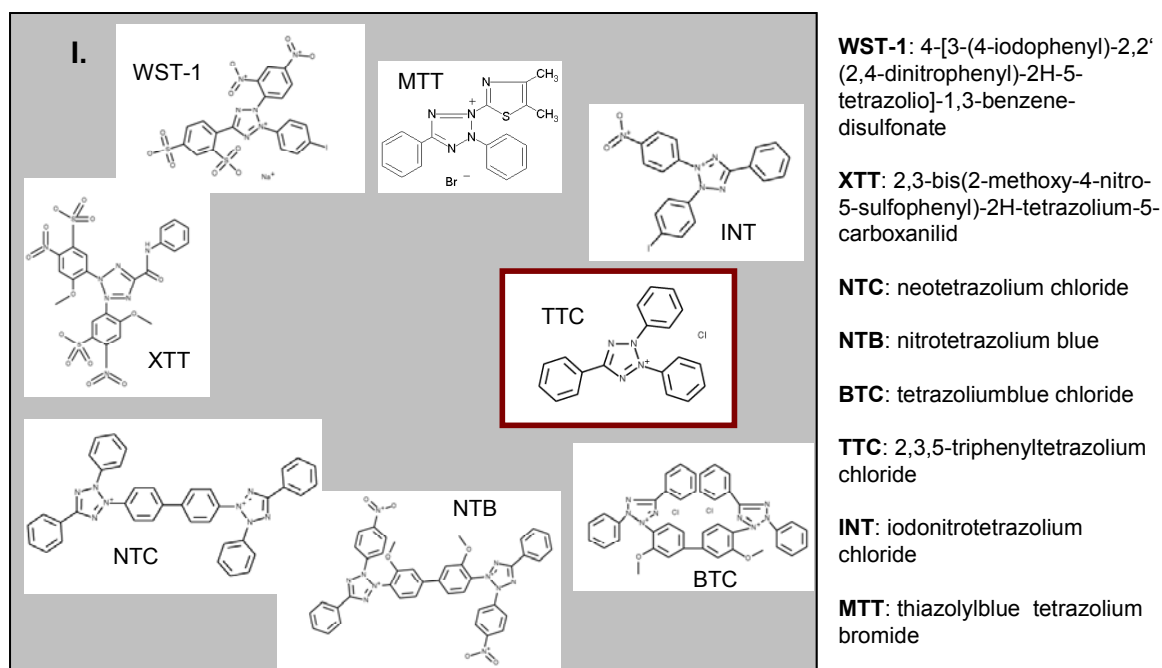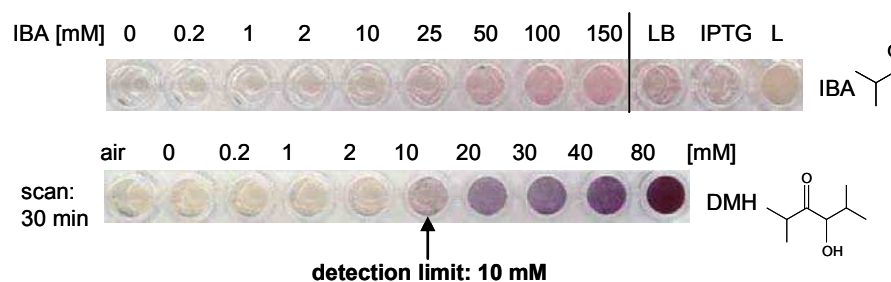

Fig. 3-3: I. Commercially available tetrazolium salts. The predominantly used tetrazolium chloride (TTC) is marked red. **II.** For the detection of 4-hydroxy-(2,5-dimethy)-hexane-3-one (DMH) nitrotetrazolium chloride (NTC) was used, allowing the detection of the product DMH (detection limit 10 mM) without interference of the substrate isobutyric aldehyde (IBA). No background coloration occurred by adding LB-cultivation medium (LB), 1 mg/mL isopropyl-β-D-thiogalactopyranoside (IPTG) or 0.78 mg/mL lysozyme (L), making the assay also useful for the detection of crude cell extracts.

3.1.2 Estimation of product concentrations

In this thesis the TTC-assay is mostly used as an endpoint detection method in order to identify the presence and estimate the concentration of 2-hydroxy ketones produced in analytical scale biotransformations. As demonstrated in Fig. 3-2 the colour intensity depends on the product concentration and is in all cases increasing linearly within a certain concentration range. Therefore relative activities of various enzymes can be distinguished (chapter 3.3.1 / 3.3.3). Beyond the linear range, differences at higher product concentrations can be also measured by following the colour development at 520 nm within a short time periode (e.g. 1 min). The slope increases with the 2-hydroxy ketone concentration. This is a good method to estimate the activities of enzymes with strong carboligase activity. Care has to be taken with poorly soluble 2-hydroxy ketones (e.g.

benzoin), which precipitate during the biotransformation step. In these cases the colour development cannot be followed continuously, but the end point detection is still possible, as both the dissolved and the precipitated product are coloured.

3.1.3 Properties of the TTC-assay

1. *High reliability*: As demonstrated in a comparison with detailed instrumental analysis (representative example given in ***publication I***) the reliability is very high.
2. *Broad detection range after assay adaptation*: Beside a broad range of aliphatic and aromatic 2-hydroxy ketones also branched-chain aliphatic products are detectable.
3. *High sensitivity*: High sensitivity for the detection of aromatic and aliphatic 2-hydroxy ketones with detection limits of 0.5-2 mM could be determined.
4. *Fast and easy to handle*: The assay is easily applicable to large libraries of purified enzymes as well as clarified crude cell extracts with a broad product range.
5. *Detection of relative carboligase activity*: Although the assay is an endpoint assay, it additional yields reliable information about the relative activities of various enzymes.
6. *Substrate inhibition:* Substrate inhibition is also detectable (e.g. if product formation is indicated by a red colour in a vial containing only one substrate, but in a vial with an additional second substrate no coloration occurs.)

3.1.4 Limits of the TTC-assay

1. Although purified enzymes as well as crude cell extracts can be used, high concentrations of cell debris should be avoided and glycerol must be omitted (***publication I***).
2. Only a few aldehydes and 2-keto acids employed as substrates cause background colouration. Care has to be taken with 3-methoxy acetaldehyde (> 1 mM), indole-3-pyruvate (> 1 mM) and phenylpyruvate (> 10 mM) (***publication I***).
3. In carboligation reactions with two different aldehydes the observed colouration can often not unambiguously be correlated with one product, as mixtures may occur. Thus, for the detection of chemo- and stereoselectivity additional instrumental analysis has to follow (chapter 3.3).

> **The improved TTC-assay allows the fast and reliable screening of the enzyme toolbox for a broad range of 2-hydroxy ketones.**

3.2 New enzymes for the enzyme toolbox

The enzyme toolbox was designed by identification of new enzymes with carboligase activity as well as by mutagenesis of already well characterised wild type enzymes (chapter 3.3.4). An overview of all enzymes investigated in this thesis is given in Tab. 3-1.

Tab. 3-1: Overview of recombinant enzymes used in this thesis. Important references concerning the cloning strategies of the various 2-keto acid decarboxylases and benzaldehyde lyases and their characterisation are given. Information about enzyme expression and purification are also found in the cited references.

enzyme (abbreviation)	**vector**[§] (reference)	**His$_6$-tag** (X-terminal position)	**characterisation described in ref.**
2-keto acid decarboxylases			
pyruvate decarboxylase (E.C. 4.1.1.1) from *Saccharomyces cerevisiae* (*Sc*PDC)	pET22b[*]	C	(Hübner *et al.*, 1978) (Jordan *et al.*, 2004)
from *Zymomonas mobilis* (*Zm*PDC)	pBTac2 (Pohl *et al.*, 1998)	C	(Pohl *et al.*, 1998) (Siegert *et al.*, 2005) (Siegert, 2000)
from *Acetobacter pasteurianus* (*Ap*PDC)	pET22b (publication IV)	C	(Raj *et al.*, 2001) publication IV
from *Zymobacter palmae* (*Zp*PDC)	pET28a(+) (publication IV)	C	(Raj *et al.*, 2002) publication IV
benzoylformate decarboxylase (E.C. 4.1.1.7) from *Pseudomonas putida* (*Pp*BFD)	pKK233-2 (Iding *et al.* 2000)	C	(Iding *et al.*, 2000) publication VII
phenylpyruvate decarboxylase(E.C. 4.1.1.74) from *Saccharomyces cerevisiae* (*Sc*PhPDC)	pET17b[*]	C	
from *Azospirillum brasilense* (*Ab*PhPDC)	pET28a(+) (publication III)	N	publication III
branched-chain 2-keto acid decarboxylase (E.C. 4.1.1.72) from *Lactococcus lactis* (*Ll*KdcA)	pET28a(+) (publication V)	N	(Smit *et al.*, 2005a) (Yep *et al.*, 2006a) publication V
benzaldehyde lyases			
benzaldehyde lyase (E.C. 4.1.2.38) from *Pseudomonas fluorescens* (*Pf*BAL)	pkk233-2 (Janzen *et al.*, 2006)	C	(Janzen *et al.*, 2006)
from *Rhodopseudomonas palustris* (*Rp*BAL) (putative)	pET28a(+) (Brosi, 2006)	C	(Brosi, 2006)

[§]The vectors pBTac2 (Roche Diagnostics, Basel, CH) and pKK233-2 (Amersham Biosciences; now GE Healthcare, Piscataway, USA) were cloned into *E. coli* SG13009(prep4) (Qiagen, Hilden, D) while the pET-vectors (Novagen, Madison, USA) were cloned into *E. coli* BL21(DE3) (Novagen, Madison, USA). The putative *Rp*BAL was found based on a sequence alignment with *Pf*BAL showing 38% sequence similarity.

---

[*] The *Sc*PDC gene was provided by Dr. Stephan König, Martin-Luther University Halle-Wittenberg and cloned by Dr. Marion Wendorff into pET22b. A recombinant pET17b plasmid containing the *Sc*PhPDC gene was provided by Dr. Michael McLeish and Dr. Malea Kneen from the University of Michigan (Yep *et al.*, 2006b).

In order to determine optimal reaction conditions, the new enzymes *Ap*PDC, *Zp*PDC, *Ll*KdcA and *Rp*BAL were first characterised with respect to their stabilities and optimal reaction conditions before their decarboxylase- and carboligase potential were investigated in detail. As (except for *Rp*BAL) all new enzymes showed decarboxylase activity, which can be much easier detected and quantified than the ligase activity, the stability investigation and pH- and temperature optima were monitored by a direct or coupled continuous photometric decarboxylase assay (method described in ***publication IV***). Some of the enzymes have already been characterised (Tab. 3-1, ***publications IV*** and ***V***). This thesis gives for the first time a comparative study for all of them under similar conditions. Carboligase activity was not tested for any of these enzymes up to now.

In a second step the substrate range of the decarboxylase reaction was determined in order to deduce information about the acyl donor spectrum for carboligase activity of the new catalyst. If a respective 2-keto acid is a substrate for decarboxylation the binding of the corresponding aldehyde to the C2-atom of ThDP is most likely, which means that this aldehyde may be a possible donor aldehyde in enzyme catalysed carboligation reactions (chapter 1.4.4). Finally, the carboligation potential, the substrate spectrum for carboligation as well as the chemo- and enantioselectivity were investigated. Detailed instrumental analysis of mixed carboligase products (Tab. 3-6/3-7) including GC, HPLC, GCMS and NMR was done by Lydia Walter from the group of Prof. Michael Müller, Albert-Ludwigs-University Freiburg.

3.2.1 Determination of optimal reaction conditions

A comparative survey of the stabilities and optimal reaction conditions for the 2-keto acid decarboxylases is given in Tab. 3-2, completed by the discussion of the most important results concerning similarities and differences between the investigated 2-keto acid decarboxylases.

pH-dependent activity and stability

All 2-keto acid decarboxylases are most stable and most active in the slightly acidic to neutral pH-range (Tab. 3-2). Although *Ap*PDC shows a broader pH-optimum with high initial rate activities even at pH 3.5, the stability in this acidic range is drastically reduced. At weak alkaline pHs > 8 all 2-keto acid decarboxylases are inactive and instable.

Temperature-dependent activity and stability

Under initial rate conditions temperature optima for all bacterial enzymes are between 50-70°C, whereas *Sc*PDC shows maximal activity at only 43°C (Tab. 3-2). Carboligase reactions, which are carried out for 2-3 days, are performed at 30°C. The half-life of all

enzymes is > 3 d. Most activation energies for the decarboxylase reaction of natural substrates range from 27-43 kJ mol^{-1}, with LlKdcA showing exceptionally low activation energy of 8.5 kJ mol^{-1} (Tab. 3-2).

Stability in organic solvents

Since carboligation reactions often contain aromatic aldehydes which are poorly water soluble the influence of the water-miscible organic solvent DMSO toward the enzymes stability was investigated. DMSO has been already successfully implemented in conversions with *Pf*BAL (Janzen *et al.*, 2006) and *Pp*BFD (Demir *et al.*, 1999) (chapter 1.5). While *Sc*PDC and *Ll*KdcA show the same stability in the presence or absence of 20 vol% DMSO, *Ap*PDC exhibits a considerable stabilisation in the presence of 20 and even 30 vol% DMSO (Tab. 3-2).

The investigation of residual activities by measuring the decarboxylase activity proved to be an easy and reliable way to identify suitable reaction conditions also for the carboligase reaction. Anyhow detected optima for ligase reactions might differ from the ones of decarboxylase reactions. As an example Iding proved that an increase of the pH towards 8 enhanced the carboligase activity for *Pp*BFD and several other factors like the temperature influence the chemo- and enantioselectivity of the biocatalysts (Iding *et al.*, 2000). Thus a further investigation of optimal reaction conditions for a desired ligase reaction might be of interest, as described in an example in chapter 3.3.2.

Tab. 3-2: Optima and stabilities of various 2-keto acid decarboxylases. Optima were determined under initial rate conditions within 90 s using a direct decarboxylase assay. For stability investigations the enzymes were incubated under the respective conditions. Samples, withdrawn at appropriate time intervals, were subjected to residual activity determination using the coupled decarboxylase assay (for assay conditions see publication IV). n.d. = not determined

	*Ll*KdcA (publication V, Nguyen, 2006)	*Ap*PDC (publication IV, Graf, 2005)	*Zp*PDC (publication IV, Brosi, 2006)	*Sc*PDC (publication IV)	*Zm*PDC (Pohl *et al.*, 1995)	*Pp*BFD (Iding *et al.*, 2000)
pH-optimum (initial rate conditions)	pH 6-7	pH 3.5-6.5	pH 4.5-8 maximum at pH 7	pH 5-7	pH 5.5-8.0 maximum pH 6-6.5	pH 5.5-7.0 maximum at 6.2
pH-dependent stability [half-life time]	pH 5-7: no activity loss within 60 h pH 8: 40 h pH 4: < 2 h	pH 5-7: no activity loss within 60 h pH 4: 2.3 h	pH 6.5 / 8: stable for several days pH 4 / 9: complete activity loss within 2 h	pH 5: 80 h pH 7: 53 h pH 8: 13 h	n.d. in potassium phosphate buffer	pH 5.5: deactivation 0.3% min^{-1} pH 10: deactivation 0.1% min^{-1}
temperature optimum (initial rate conditions)	50°C	65°C	55°C	43°C	60°C	68°C
temperature dependent stability [half-life time]	40°C: 80 h 50°C: 9 h 55°C: 4 h	20°C: 193 h 30°C: 144 h 40°C: 34 h 50°C: 12 h 60°C: 2 h 70°C: 0.4 h	30°C: 150 h 40°C: 40 h 50°C: 10 h 60°C: 0.4 h	20°C: 235 h 30°C: 78 h 35°C: 62 h	50°C: 24 h	60°C: 2 h 80°C: 15-20 min
activation energy for decarboxylation substrate	8.5 kJ mol^{-1} 3-methyl-2-oxobutanoic acid	27.1 kJ mol^{-1} pyruvate	41 kJ mol^{-1} pyruvate	n.d.	43 kJ mol^{-1} pyruvate (unpublished)	38 kJ mol^{-1} benzoylformate
stability in organic solvents (half-life times measured at 30°C)	stable in 20% (v/v) DMSO (half-life 150 h); unstable in 15% (v/v) PEG 400 (half-life: 6 h)	stability enhanced in 30% (v/v) DMSO (half-life 430 h)	n.d.	stable in 20-30% (v/v) DMSO	n.d.	n.d.
buffer used for kinetic studies and investigation of the substrate range	50 mM potassium phosphate buffer pH 6.8, 2.5 mM MgSO$_4$, 0.1 mM ThDP	50 mM potassium phosphate buffer pH 6.5, 2.5 mM MgSO$_4$, 0.1 mM ThDP	50 mM potassium phosphate buffer pH 6.5, 2.5 mM MgSO$_4$, 0.1 mM ThDP	50 mM potassium phosphate buffer pH 6.5, 2.5 mM MgSO$_4$, 0.1 mM ThDP	100 mM potassium phosphate buffer pH 6.5, 5 mM MgSO$_4$, 0.1 mM ThDP	50 mM potassium phosphate buffer pH 6.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP

3.2.2 Investigation of the decarboxylase activity – substrate and kinetic studies

Bacterial PDCs are highly similar but differ from yeast PDC

All bacterial PDCs (*Zm*PDC, *Zp*PDC, *Ap*PDC) exhibit hyperbolic v/[S]-plots with highly similar kinetic parameters (Tab. 3-4, **publication IV**). Deviations have been observed with *Sc*PDC, where the phenomenon of a sigmoidal plot has been deduced to activation of the enzyme by its substrate pyruvate (König, 1998). Here activator binding causes the formation of a more compact structure, representing the active form of the enzyme (chapter 1.5.2, Lu *et al.*, 2000; Lu *et al.*, 1997). Just recently the mechanism for the allosteric activation of yeast PDC could be elucidated (Kutter *et al.*, 2008). The substrate range of all PDCs is limited to short-chained aliphatic 2-keto acids (Tab. 3-3, **publication IV**). One interesting exception is *Ap*PDC which is able to decarboxylate benzoylformate (**11**) to a certain extend (1.1 U/mg), suggesting a comparatively larger binding site compared to *Zm*PDC. The impact on the carboligation reaction is discussed in chapter 3.4.5.
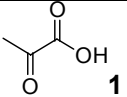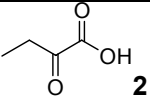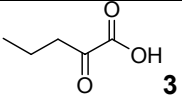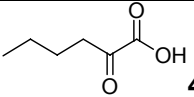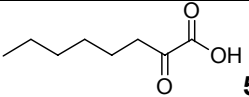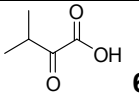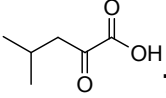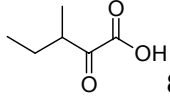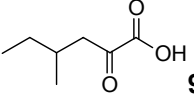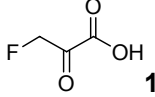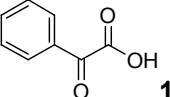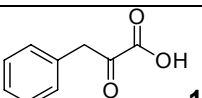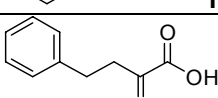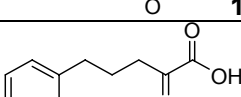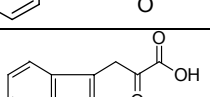
LlKdcA, an enzyme with a very broad substrate range

*Ll*KdcA has an exceptionally broad substrate range encompassing linear and branched-chain aliphatic and aromatic 2-keto acids (**publication V**, Smit *et al.*, 2005a; Yep *et al.*, 2006a), showing the highest maximal velocity with 3-methyl-2-oxobutanoic acid (Tab. 3-3, **6**, Tab 3-4). Although the maximal activity in the presence of phenylpyruvate (**12**) is only about 9% of the velocity towards the natural substrate **6**, the K_M is 40-times lower. This high affinity for phenylpyruvate and also for indole-3-pyruvate is of special interest for certain carboligase reactions as discussed in chapter 3.3.2.
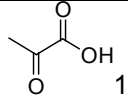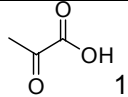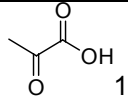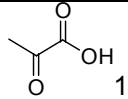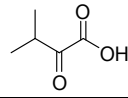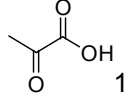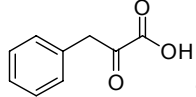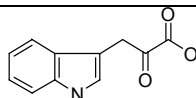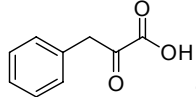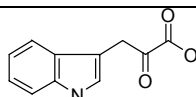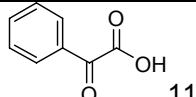
AbPhPDC, a new enzyme with variable kinetic properties

The v/[S]-plots of *Ab*PhPDC for varying substrates differ remarkably (Tab. 3-4). As pointed out in detail in **publication III** the enzyme, which was originally named indole-3-pyruvate decarboxylase, is now renamed as a phenylpyruvate decarboxylase (PhPDC) due to its significant higher specific activity towards phenylpyruvate (**12**) relative to indole-3-pyruvate (**15**). Notably, a classical Michaelis-Menten plot is obtained with phenylpyruvate, while the v/[S]-plot with indole-3-pyruvate is sigmoidal, with a Hill-coefficient of 1.8-1.9. These data hint to an unusual substrate dependent allosteric behaviour. The crystal structure of the *Ab*PhPDC shows a compact tetramer, with no indication for allosteric effects ((Versées *et al.*, 2007), pdb-code: 2nxw). A bound substrate in the crystal structure could shed light on the underlying molecular mechanism; however such data are not yet available.

Tab. 3-3: Substrate range of the decarboxylation reaction of various 2-keto acid decarboxylases. Activity was measured using the coupled decarboxylase assay at 30°C. Substrate concentration: 30 mM (**1-14**), 1 mM (**15**). Buffer conditions are given in Tab. 3-2. The natural substrates (bold face) resulted in highest specific activity and were used for further kinetic studies (Tab. 3-4). n.d. = not determined

2-keto acid	specific activity [U/mg]					
	*Ll*KdcA (publication V)	*Ap*PDC (publication IV)	*Zm*PDC (Siegert, 2000)	*Sc*PDC (publication IV)	*Zp*PDC (publication IV)	*Pp*BFD (publication VII)
1	1.9	**89.3**	**120**	**43.4**	**147**	0.4
2	14.1	60	79	16.9	85.1	21.4
3	22.7	12.9	13	18.8	20.7	12.5
4	20.1	4.2	0.2	5.3	6.2	9.2
5	0.0	1.1	n.d.	0.0	0.6	7.2
6	**152**	n.d.	0.0	6.9	10.4	0.0
7	40	1.1	0.3	0.3	2.8	3.3
8	57.9	2.2	0.0	0.0	2.8	3.6
9	28.9	0.6	0.0	0.0	0.4	6.9
10	0.0	0.3	n.d.	0.0	0.6	0.0
11	12.8	1.1	0.0	0.0	0.3	**420.1**
12	13.1	0.8	0.0	0.0	1.8	0.0
13	2.4	0.3	0.0	1.7	0.3	4.4
14	1.7	0.0	0.0	0.2	0.2	0.0
15	1.3	0.0	n.d.	0.0	0.5	1.5

Tab. 3-4: Kinetic parameters for the decarboxylation of 2-keto acids by various 2-keto acid decarboxylases. All data were deduced using the coupled decarboxylase assay as described in publication IV and VII with the indicated buffers. Kinetic parameters were calculated by non-linear regression using the Michaelis-Menten equation in Origin 7G SR4 (OriginLab Coop., Northampton) except for ScPDC and AbPhPDC, where the Hill equation was used (for further information see publication IV). h = Hill coefficient, K_i= inhibition constant

enzyme	substrate	V_{max} [U/mg]	K_M [mM]	shape of the v/[S]-curve
ApPDC[a] (publication IV)	(pyruvate) 1	110 ± 1.9	2.8 ± 0.2	Michaelis-Menten kinetic
ZpPDC[a] (publication IV)	(pyruvate) 1	116 ± 2.0	2.5 ± 0.2	Michaelis-Menten kinetic
ZmPDC[b] (Siegert et al., 2005)	(pyruvate) 1	120-150	1.1 ± 0.1	Michaelis-Menten kinetic
ScPDC[a] (publication IV)	(pyruvate) 1	112.0 ± 35.6	$S_{0.5}$: 21.6 ± 7.4	sigmoidal (h = 1.6 ± 0.2) & substrate surplus inhibition (K_i = 35.1 ± 16.5)
LlKdcA[a] (publication V)	(3-methyl-2-oxobutanoate) 6	181.6 ± 1.7	5.02 ± 0.2	Michaelis-Menten kinetic
	(pyruvate) 1	3.7 ± 0.3	29.8 ± 6.4	Michaelis-Menten kinetic
	(phenylpyruvate) 12	15.7 ± 0.2	0.127 ± 0.007	Michaelis-Menten kinetic
	(indole-3-pyruvate) 15	1.55 ± 0.6	0.234 ± 0.024	Michaelis-Menten kinetic
AbPhPDC[c] (publication III)	(phenylpyruvate) 12	5.56 ± 0.13	1.08 ± 0.09	Michaelis-Menten kinetic (h = 1)
	(indole-3-pyruvate) 15	0.07 ± 0.002	0.13 ± 0.01	sigmoidal (h = 1.85)
PpBFD[d] (publication VII) (Iding et al., 2000)	(benzoylformate) 11	400 ± 7	0.37 ± 0.03	Michaelis-Menten kinetic

[a] reaction conditions as described in Tab. 3-1 (lowest row);
[b] 50 mM Mes/KOH-buffer pH 6.5, 2.5 mM $MgSO_4$, 0.1 mM ThDP (Siegert et al., 2005);
[c] 10 mM Mes-buffer pH 6.5, 2.5 mM $MgSO_4$, 0.1 mM ThDP;
[d] 50 mM potassium phosphate buffer pH 6.5, 2.5 mM $MgSO_4$, 0.1 mM ThDP.

3.2.3 Investigation of the carboligase activity

The new enzymes cloned in this work (*Ll*KdcA, *Ap*PDC, *Zp*PDC, *Rp*BAL) were subjected to a thorough investigation of their carboligase potential relative to the already well characterised enzymes in the toolbox (*Zm*PDC, *Sc*PDC, *Pp*BFD, *Pf*BAL). Some model reactions are described in order to compare the catalysts abilities. It should be noted that the reaction conditions might also influence the enzymes. Carboligase activity for *Sc*PhPDC and *Ab*PhPDC was elucidated especially for phenylpyruvate and acetaldehyde as substrates as described in chapter 3.3.2.

The self-ligation of aliphatic aldehydes is catalysed with low stereoselectivity
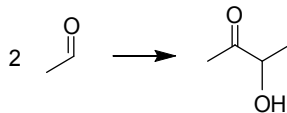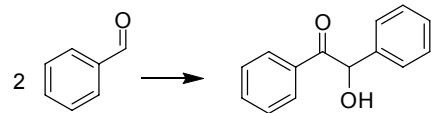
The formation of acetoin either starting from acetaldehyde or pyruvate has been described for *Sc*PDC (Chen & Jordan, 1984), some further yeast strains (Kurniadi *et al.*, 2003; Neuser *et al.*, 2000) and *Zm*PDC (Bringer-Meyer & Sahm, 1988; Crout *et al.*, 1991). The potential of *Pp*BFD and *Pf*BAL to catalyse the carboligation of two linear aliphatic aldehydes is described in **publication II**. The newly investigated enzymes were analysed towards the self-ligation of acetaldehyde to acetoin (Tab. 3-5). Whereas *Ap*PDC shows the comparatively highest activity, the stereocontrol is best with *Rp*BAL (*ee* 60% (*S*), which is inverse to *Pf*BAL (*ee* 40% (*R*)).

Nevertheless the enantioselectivity of this reaction is in all cases very low and most probably a consequence of less stabilisation of small aldehydes in the active site, which is further discussed in chapter 3.4.6.

The self-ligation of aromatic aldehydes is catalysed by the BALs

*Zp*PDC and *Ll*KdcA catalyse the benzoin formation from benzaldehyde in almost negligible amounts. With *Ap*PDC benzoin formation was observed, demonstrating that benzaldehyde can bind to the C2-atom of ThDP as a donor aldehyde, although the activity is very low. The benzoin ligase activity of *Rp*BAL is $4 \cdot 10^{-3}$ U/mg and thus much lower than the activity of the very potent *Pf*BAL (336 U/mg) (Janzen *et al.*, 2006). However, *Rp*BAL is the first catalyst showing (*S*)-selective formation of benzoin. Although the enantioselectivity with 15% (*S*) is very low, a further characterisation of this new enzyme is still of interest. According to the TTC-assay (not shown) *Rp*BAL is, like *Pf*BAL, able to catalyse the self-ligation of 3,5-dimethoxybenzaldehyde, a reaction which is not observed with all tested 2-keto acid decarboxylases.

Tab. 3-5: Self-ligation of (**A**) acetaldehyde (40 mM) towards acetoin and (**B**) benzaldehyde (20 mM) towards benzoin catalysed by various enzymes. Reaction conditions: 50 mM potassium phosphate buffer, pH 6.5/7.5, 2.5 mM MgSO$_4$, 0.1 mM ThDP, 20 vol% DMSO, 0.3 mg/mL enzyme, 30°C. Specific activities (spec. activity) were calculated after GC- (**A**) respectively HPLC-analysis (**B**). Assignment of the absolute configuration of acetoin (**A**) was done by chiral GC and circular dichroism (CD) (publication II) in accordance to *Sc*PDC (Baykal *et al.*, 2006; Crout *et al.*, 1991) and of benzoin (**B**) by chiral HPLC in accordance to *Pf*BAL (Janzen *et al.*, 2006).

enzyme	**A** 	**B**
***Ap*PDC**	spec. activity[a]: 58.9 U/mg *ee* 31% (*S*)	spec. activity[d]: 2.4 10^{-3} U/mg *ee* n.d.
***Zp*PDC**	spec. activity: 2 U/mg *ee* 58% (*S*)	not detectable, conversion too low
***Ll*KdcA**	spec. activity[a]: 4.4 U/mg *ee* 47% (*R*)	not detectable, conversion too low[e]
***Rp*BAL**	spec. activity: 0.03 U/mg *ee* 60% (*S*)	spec. activity: 4 10^{-3} U/mg *ee* ~ 15% (*S*)
***Pp*BFD**	spec. activity[b]: 0.7 U/mg *ee* 34-36% (*R*)	spec. activity[c]: 6.4 U/mg *ee* 99% (*R*)

[a] detected in the linear range after 1 h; [b] detected in the linear range after 19 h;
[c] detected in the linear range after 3 h; [d] diverse reaction conditions (see publication IV);
[e] with higher amounts of *Ll*KdcA small amounts of (*R*)-benzoin (*ee* > 98%) could be gained (publication V).

### *Mixed carboligations of benzaldehyde and acetaldehyde*

Determination of the product range obtained by the mixed carboligation of benzaldehyde and acetaldehyde was used as a preliminary test reaction for all enzymes to obtain valuable information about their chemo- and enantioselectivity.

For all investigated PDCs phenylacetylcarbinol (PAC) (Tab. 3-6, **17**) was identified as the only product, with small amounts of acetoin (**16**) in the case of *Zp*PDC. This result is in accordance to the other already intensively studied pyruvate decarboxylases *Sc*PDC and *Zm*PDC (chapter 1.5.2) and the results of whole cell transformation experiments for *Zp*PDC (Rosche *et al.*, 2003). Consequently all of them exclusively bind acetaldehyde as the donor aldehyde and benzaldehyde as the acceptor, if the acetoin formation is suppressed by an appropriate choice of the reaction conditions (e.g. excess of benzaldehyde). Concerning stereoselectivity of the carboligation (*R*)-PAC is formed enantioselectively by *Zm*PDC (*ee* 98%) (Siegert *et al.*, 2005) whereas *Sc*PDC (*ee* 90%) (Rosche *et al.*, 2001), *Zp*PDC and *Ap*PDC show lower stereoselectivity (Tab. 3-6). A possible structural explanation for this behaviour will be discussed in chapter 3.4.6 (***publication IV***).

By contrast *Pp*BFD (chapters 1.5.1) and *Rp*BAL accept benzaldehyde as a donor and acetaldehyde as an acceptor yielding predominantly 2-hydroxypropiophenone (HPP, **18**) with excess of the (*S*)-enantiomer.

*Ll*KdcA shows almost no chemoselectivity in carboligation reaction with acetaldehyde and benzaldehyde. As demonstrated in Tab. 3-6 both aldehydes may act either as donor or acceptors yielding almost equal amounts of **17** and **18**, which is probably a consequence of the sterical properties in the active site as discussed in chapter 3.4.5.

Tab. 3-6: Carboligation of benzaldehyde with acetaldehyde catalysed by various enzymes. Reaction conditions: 50 mM potassium phosphate buffer, pH 6.5/7.0, 2.5 mM MgSO₄, 0.1 mM ThDP, 20% (v/v) DMSO, 30 °C. Relative product distribution is given in mol% (NMR); *ee*-values were determined by HPLC-analysis (data generated by Lydia Walter, University of Freiburg). n.d. = not determined, - = no product formation

enzyme	substrates		possible products			
	benzaldehyde	acetaldehyde	**16**	**17**	**18**	**19**
***Ap*PDC**[a] (publication IV)	n.d.	n.d.	n.d.	30% *ee* 91% (*R*)	-	n.d.
***Zp*PDC**[b]	80%	-	8% *ee* n.d.	12% *ee* 89% (*R*)	-	-
***Ll*KdcA**[c] (publication V)	n.d.	n.d.	traces	41% *ee* 92% (*R*)	59% *ee* 93% (*R*)	traces *ee* 98% (*R*)
***Rp*BAL**[b]	47%	-	4% *ee* n.d.	5% *ee* 31% (*R*)	40% *ee* 78.5% (*S*)	4% *ee* 15% (*S*)
***Pp*BFD**[b] (publication VII) (Iding *et al.* 2000)	12%	3%	-	-	81% *ee* 92% (*S*)	4% *ee* 99% (*R*)

[a] 30 mM benzaldehyde, 150 mM acetaldehyde, 2 d reaction time; just mixed product (**17**+**18**) purified and yield detected by NMR-analysis; [b] 18 mM benzaldehyde, 18 mM acetaldehyde, 3 d reaction time;
[c] 30 mM benzaldehyde, 40 mM acetaldehyde, 3 d reaction time; just mixed ligation product (**17**+**18**) purified and analysed, distribution of **17** and **18** detected by NMR-analysis.

Mixed carboligation of larger aliphatic and substituted aromatic aldehydes
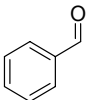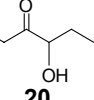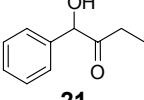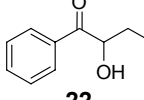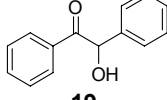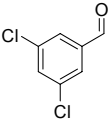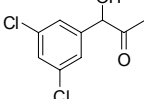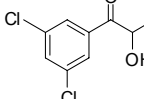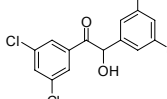
In this thesis *Sc*PDC, *Ap*PDC and *Ll*KdcA were investigated according to their carboligase ability in mixed carboligase reactions with substituted benzaldehydes and varying aliphatic aldehydes in order to gain versatile PAC-analogue products (see also chapter 3.3.3). The (*R*)-PAC-analogue of benzaldehyde and propanal is selectively catalysed by *Ll*KdcA and the PDCs (Tab. 3-7 A, **21**). Again *Ap*PDC is less stereoselective.

Chemo- and stereoselectivity are strongly influenced by the size of the aromatic aldehyde as demonstrated in mixed carboligations of 3,5-dichlorobenzaldehyde and acetaldehyde

(Tab. 3-7). The PAC-derivative (**23**) is accessible by *Ap*PDC, however only with very weak activity and without stereoselectivity, demonstrating that 3,5-dichlorobenzaldehyde is a weakly accepted donor and acceptor, probably due to steric hindrance (chapter 3.4.5). By contrast, *Sc*PDC catalyses the synthesis of the HPP-derivative (**24**) besides traces of the PAC-analogue (**23**). Low conversions, chemo- and enantioselectivities suggest that 3,5-dichlorobenzaldehyde can not fit optimally into the active site of the PDCs (chapter 3.4.5).

Although *Ll*KdcA catalyses the carboligation of 3,5-dichlorobenzaldehyde and acetaldehyde, the resulting product is, beside acetoin, not the expected PAC-analogue but the HPP-derivative (**24**) (*publication V*). The possibility to influence the chemoselectivity of the carboligation reaction by an appropriate combination of the donor- and acceptor aldehyde is unique among the currently known ThDP-dependent enzymes with carboligase activity and shows the big advantage of this enzyme to access HPP- as well as PAC-analogue 2-hydroxy ketones. A structural explanation of these properties of *Ll*KdcA is given in ***publication VI***.

Tab. 3-7: Carboligation of benzaldehyde with propanal (**A**) and 3,5-dichlorobenzaldehyde with acetaldehyde (**B**) catalysed by *Ap*PDC, *Sc*PDC and *Ll*KdcA. Reaction conditions: 50 mM potassium phosphate buffer, pH 6.8 (*Ll*KdcA)/7.0 (PDCs), 2.5 mM MgSO$_4$, 0.1 mM ThDP, 20% (v/v) DMSO, 30 °C. Relative product distribution is given in mol% (NMR); *ee*-values were determined by HPLC-analysis (data generated by Lydia Walter, University of Freiburg). n.d. = not determined

	substrates		possible products			
A	benzaldehyde	propanal	**20**	**21**	**22**	**19**
***Ap*PDC[a]** (publication IV)	preparative scale, product yield: 45% thereof:		16% *ee* n.d.	81% *ee* 82% (*R*)	3% *ee* n.d.	–
***Sc*PDC[a]**	preparative scale, product yield: 30% thereof:		–	91% *ee* 96.5% (*R*)	9% *ee* n.d.	–
***Ll*KdcA[b]** (publication V)	3%	17%	29%. *ee* n.d.	51% *ee* 98.5% (*R*)	–	–
B	3,5-dichlorobenzaldehyde	acetaldehyde	**16**	**23**	**24**	**25**
***Ap*PDC[c]**	75%	4%	16% *ee* n.d.	5% *ee* 4% (*R*)	–	–
***Sc*PDC[c]**	74%	–	–	7.5% *ee* 68% (*R*)	18.5% *ee* n.d.	–
***Ll*KdcA[d]** (publication V)	3%	17%	29% *ee* n.d.	–	51% *ee* 96.5% (*R*)	–

[a] 20 mM benzaldehyde, 80 mM propanal, reaction time 72 h; [b] 18 mM benzaldehyde, 18 mM propanal, reaction time 23 h; [c] 20 mM 3,5-dichlorobenzaldehyde, 40 mM acetaldehyde, reaction time 48 h;
[d] 20 mM 3,5-dichlorobenzaldehyde, 30 mM acetaldehyde, reaction time 72 h.

3.2.4 The enlarged enzyme toolbox

Due to their carboligase ability *Ap*PDC, *Zp*PDC, *Ll*KdcA and *Rp*BAL were added to the enzyme toolbox. *Ab*PhPDC and *Sc*PhPDC also showed ligase ability which will be described in chapter 3.3.2. Together with the already well characterised *Zm*PDC, *Pf*BAL, *Pp*BFD and *Sc*PDC (chapter 1.5) the platform now encompasses ten wild type enzymes (Tab. 3-8). Additionally two BFDs of *Pseudomonas aeruginosa* and *Bradyrhizobium japonicum* (Wendorff, 2006) as well as three BFDs, two originating from a chromosomal library of *Pseudomonas putida* ATCC 12633 and the third from an environmental DNA library (Henning *et al.*, 2006) were added. 66 active variants from earlier studies (chapter 1.5) with partially different substrate ranges as well as varied chemo- and enantioselectivities further broaden the range of the enzyme platform. As will be described in chapter 3.3.4, four active variants of *Pp*BFD could be added to the toolbox in the course of this thesis.

Tab. 3-8: In the course of this thesis the enzyme toolbox was extended to 15 wild type enzymes and 70 active variants.

enzyme (abbreviation)	source	variants
2-keto acid decarboxylases		

pyruvate decarboxylase
    from *Saccharomyces cerevisiae* (*Sc*PDC)
    from *Zymomonas mobilis* (*Zm*PDC)
    from *Acetobacter pasteurianus* (*Ap*PDC)
    from *Zymobacter palmae* (*Zp*PDC)
	
	S. König
	K. Mesch/M. Pohl
	this work
	this work
		
		-
		23
		-
		-

benzoylformate decarboxylase
    from *Pseudomonas putida* (*Pp*BFD)
    from *Pseudomonas putida* (*Pp*BFD) gene II
    from *Pseudomonas putida* (*Pp*BFD) gene III
    from *Pseudomonas aeruginosa* (*Pa*BFD)
    from *Bradyrhizobium japonicum* (*Bj*BFD)
    from metagenome
	
	H. Iding/M. Pohl/this work*
	H. Henning
	H. Henning
	M. Wendorff
	M. Wendorff
	H. Henning
		
		32 + 4*
		-
		-
		-
		-
		-

phenylpyruvate decarboxylase
    from *Saccharomyces cerevisiae* (*Sc*PhPDC)
    from *Azospirillum brasilense* (*Ab*PhPDC)
	
	M. Kneen
	S. Spaepen
		
		-
		-

branched-chain 2-ketoacid decarboxylase from *Lactococcus lactis* (*Ll*KdcA)	 this work	 -
benzaldehyde lyases		
from *Pseudomonas fluorescens* (*Pf*BAL) from *Rhodopseudomonas palustris* (*Rp*BAL)	E. Janzen / M. Pohl this work	11 -
total amount of wild type enzymes: 15		**variants: 70**

* 4 new variants of *Pp*BFD are described in chapter 3.3.4 (publication VII)

In this thesis the toolbox of ThDP-dependent enzymes with carboligase activity was enhanced now encompassing 15 wild type enzymes and 70 variants. Among the new enzymes the branched-chain 2-keto acid decarboxylase from *Lactococcus lactis* shows outstanding abilities due to its exceptional variable chemoselectivity depending on the sizes of the substrates. Furthermore the pyruvate decarboxylase from *Acetobacter pasteurianus* and the benzaldehyde lyase from *Rhodopseudomonas palustris* proved to be interesting catalysts as they catalyse the partially (*S*)-selective carboligation of certain aldehydes, which is rare among the predominantly (*R*)-selective enzymes.

3.3 Accessing new 2-hydroxy ketones

The enlarged enzyme toolbox was used to investigate the biocatalytical access of new diversely substituted and enantio-complementary 2-hydroxy ketones.

3.3.1 Self-ligation of branched-chain aliphatic aldehydes

Branched-chain aliphatic 2-hydroxy ketones such as 4-hydroxy-(2,5-dimethyl)-hexan-3-one (DMH) and 5-hydroxy-(2,7-dimethyl)-octan-4-on (DMO) (Fig. 3-4) are particularly interesting in order to access the corresponding chiral diols or amino alcohols which represent interesting building blocks for bioorganic chemistry (*publication II*). However, such branched-chain aliphatic 2-hydroxy ketones were not stereoselectively accessible by organic chemistry. To identify appropriate biocatalysts the enzyme toolbox (Tab. 3-8) was screened with respect to the self-ligation ability of isobutyric aldehyde, isovaleric aldehyde and pivaldehyde using the TTC- or NTC-assay, respectively (chapter 3.1.1).
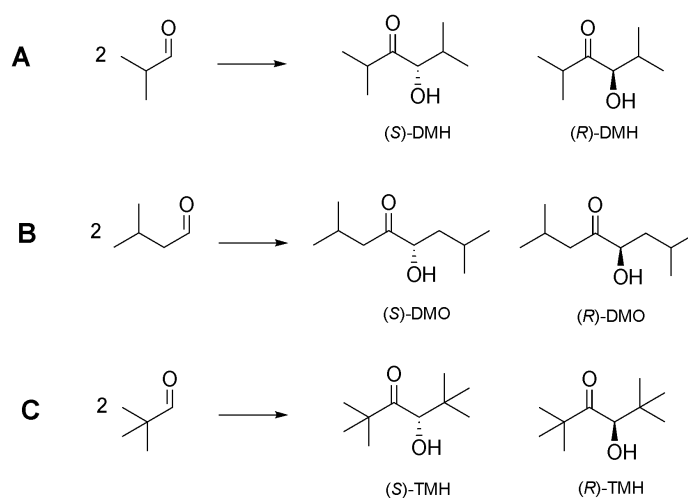

Fig. 3-4: Self-ligation of isobutyric aldehyde (**A**), isovaleric aldehyde (**B**) and pivaldehyde (**C**). The respective branched-chain 2-hydroxy ketones 4-hydroxy-(2,5-dimethyl)-hexan-3-one (DMH) (**A**), 5-hydroxy-(2,7-dimethyl)-octan-4-one (DMO) (**B**) and 4-hydroxy (2,2',5,5'-tetramethyl)-hexan-3-one (TMH) (**C**) can occur in the (*R*)- or (*S*)-form or a mixture of both.

No enzyme was found to catalyse the self-ligation of isobutyric aldehyde und pivaldehyde

Especially *Ll*KdcA was expected to catalyse the ligation of isobutyric aldehyde to DMH (Fig. 3-4 A) since the corresponding 2-keto acid is the natural substrate of this biocatalyst (Tab. 3-3). However, carboligase activity with isobutyric aldehyde was not observed, which is due to an inhibition caused by this aldehyde (for details see **publication V**). As none of the enzymes in the toolbox catalyses the self-ligation of isobutyric aldehyde or pivaldehyde it may be speculated that the branch in Cα-position to the carbonyl functions is limiting the carboligase ability of two of these substrates. However, in mixed reactions *Pf*BAL is able to catalyse the ligation of isobutyric aldehyde (acceptor) with benzaldehyde (donor) in small amounts (**publication I**).

Various enzymes catalyse the self-ligation of isovaleric aldehyde

In contrast various enzymes were identified catalysing the self-ligation of isovaleric aldehyde to DMO (Fig. 3-4 B), among them are *Ll*KdcA, *Pf*BAL, *Pf*BALH49A and *Pf*BALH281A. Weak signals in the TTC-assay were also obtained with *Pp*BFD, *Zm*PDC and their two variants *Pp*BFDH281A/F464I and *Zm*PDCW392M (Fig. 3-5). Biotransformations with all variants were subsequently analysed by GC with regard to stereoselectivity (*ee*) and relative activities. Analytical data of the enzymatic products and the chemical synthesised racemic material (**publication II**) fitted very well. As demonstrated in Fig. 3-5, the (*R*)-enantiomers are produced in excess, with *Pf*BAL giving the highest space time yield and the highest *ee* (89% (*R*)-DMO, GC).
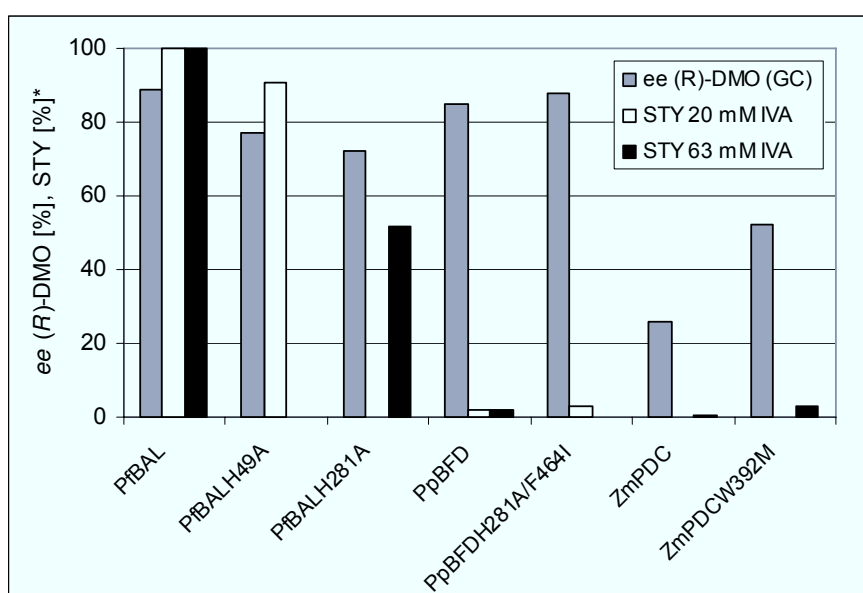
Fig. 3-5: Enantiomeric excess (*ee*) and relative space time yield (STY) (as a measure for activity) of the carboligation of isovaleric aldehyde (IVA) (20 mM and 63 mM, respectively) to 5-hydroxy-(2,7 dimethyl)-octan-4-on (DMO) by various wild type enzymes and variants determined by GC analysis. The highest STY of 13.15 g L^{-1} d^{-1} was obtained with *Pf*BAL with 63 mM substrate. Reaction conditions: 0.3 mg/mL enzyme in 50 mM potassium phosphate buffer, pH 8, 2.5 mM MgSO$_4$, 0.1 mM ThDP, 20% (v/v) DMSO, 30°C, 100 rpm.

Branched-chain aliphatic 2-hydroxy ketones are accessible with high ee

The absolute configuration of the DMO obtained by *Pf*BAL catalysis was determined by the Mosher-ester methodology (Sullivan *et al.*, 1973).[*] According to Sullivan the predominantly obtained low field shifted diastereomer is most likely derived from the (*R*)-enantiomer (GCMS *ee* 88%, ^{1}H-NMR *ee* 89%). These data correlate well with the CD spectra obtained with other chiral aliphatic 2-hydroxy ketones, all showing a negative Cotton effect at 278 nm for the (*R*)-enantiomer and a positive one for the (*S*)-enantiomer (***publication II***, Baykal *et al.*, 2006).

Although the stereoselectivity of *Ll*KdcA is much lower compared to *Pf*BAL (30-47% (*S*)-DMO) (data not shown in Fig. 3-6), *Ll*KdcA is the only enzyme which catalyses the formation of the (*S*)-product in excess (for structure modelling see ***publication VI***).

Despite the low stereoselectivities obtained with non-branched aliphatic 2-hydroxy ketones (e.g. *ees* of 32-60% for acetoin (Tab. 3-5), ***publication II***) the enzyme-catalysed ligation of the branched-chain aliphatic isovaleric aldehyde yields in predominantly higher enantiomeric excesses (> 85% (*R*)-DMO). This is probably a consequence of a better stabilisation of the sterically more demanding aldehydes in the active centre (chapter 3.4).

3.3.2 Carboligation of instable aldehydes

CH-acidic aldehydes such as phenylacetaldehyde and indole-3-acetaldehyde are prone to enolisation and are thus difficult substrates for enzymatic as well as chemical transformations (Schütz *et al.*, 2003). Moreover indole-3-acetaldehyde is very instable and decomposes rapidly, which makes its direct application in biotransformations impossible (Koga *et al.*, 1992). Phenylacetaldehyde is stable and commercially available. However, in aqueous buffer the aldol reaction occurs spontaneously. As suggested by Prof. Dr. Michael Müller (University Freiburg) this problem can easily be overcome by *in situ* production of these aldehydes by enzymatic decarboxylation of the corresponding 2-keto acids. Referring to the reaction mechanism (chapter 1.4.4, Fig. 1-8), the decarboxylation of a 2-keto acid results in a reaction intermediate (activated aldehyde) with the corresponding aldehyde bound to the cofactor ThDP, which can further react with the acceptor aldehyde.

LlKdcA catalyses the mixed carboligation of indol-3-acetaldehyde and acetaldehyde

Among the enzymes of the toolbox *Ll*KdcA and both phenylpyruvate decarboxylases from *Saccharomyces cerevisiae* and *Azospirillum brasilense* (Tab. 3-4) are able to decarboxylate

---

[*] Preparation of the Mosher-ester and subsequent NMR-analysis was performed during a research stay at the Institute of Pharmaceutical Sciences, Albert-Ludwigs-University Freiburg under skilful guidance of Lydia Walter.
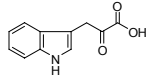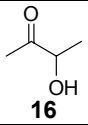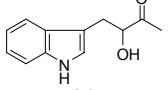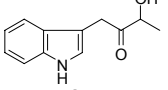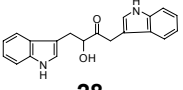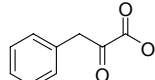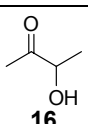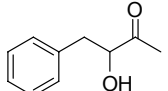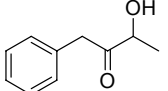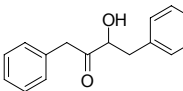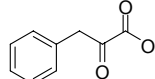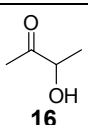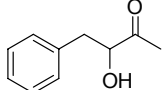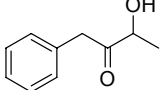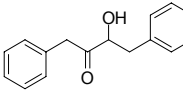
phenylpyruvate with good yields. *Ll*KdcA and *Ab*PhPDC were further found to decarboxylate indole-3-pyruvate with low activity. Carboligase data presented in Tab. 3-9 show *Ll*KdcA to form the HPP-analogue (**27**) in a ligation using acetaldehyde and indole-3-pyruvate as substrates with high chemoselectivity (***publication V***). Therewith *Ll*KdcA was the first ThDP-dependent enzyme tested accepting substrates as large as indole-3-acetaldehyde as donor aldehydes for carboligase reactions (***publication VI***).

PhPDCs and LlKdcA catalyse the carboligation of phenylpyruvate and acetaldehyde

Product formation in mixed carboligations of phenylpyruvate and acetaldehyde was demonstrated with *Ab*PhPDC, *Sc*PhPDC and *Ll*KdcA using the TTC-assay (plate not shown) and additionally characterised in detail using instrumental analysis for *Ll*KdcA and *Sc*PhPDC (Tab. 3-9 B). Both enzymes produce a mixture of the HPP- (**30**) and PAC-analogue (**29**) products. While with *Ll*KdcA the yield is much higher, the chemoselectivity of *Sc*PhPDC towards **30** is significantly better (**30:29** = 16:1) than with *Ll*KdcA (5:1). The general reason for the excess of **30** might be the fact that phenylpyruvate is first decarboxylated and the resulting activated phenylacetaldehyde (chapter 1.4.4) is subsequently coupled to acetaldehyde yielding the HPP-analogue product.

An interesting change of chemoselectivity towards the PAC-analogue (**29**) was observed with *Ll*KdcA, when both substrates, acetaldehyde and phenylacetaldehyde, were applied in form of the corresponding 2-keto acids (pyruvate and phenylpyruvate) (Tab. 3-9 C). This can be explained as follows: due to the different K_M-values for both substrates (0.13 mM for phenylpyruvate; 30 mM for pyruvate, Tab. 3-4) *Ll*KdcA starts with the decarboxylation of phenylpyruvate and the resulting phenylacetaldehyde is released. If phenylpyruvate is almost consumed, decarboxylation of pyruvate to acetaldehyde begins, with ThDP-bound acetaldehyde as an intermediate, which can now react with phenylacetaldehyde as the acceptor aldehyde yielding the PAC-analogue (**29**). This is a very nice example for the variation of chemoselectivity by reaction engineering and was just possible to discover by the synergic cooperation between the group of Prof. Dr. Michael Müller (Albert-Ludwigs-University Freiburg) and our group. Due to a lack of standards the enantioselectivity of the catalysts could not yet be determined (Lydia Walter, personal communication).

Tab. 3-9: C-C coupling of CH-acidic aldehydes with *Sc*PhPDC and *Ll*KdcA. Substrates are either applied as aldehydes or as the corresponding 2-keto acids which are in situ decarboxylated in the same active site. Reaction conditions: 50 mM potassium phosphate buffer, pH 6.8 (*Ll*KdcA) / 7.0 (*Sc*PhPDC), 2.5 mM $MgSO_4$, 0.1 mM ThDP, 20% (v/v) DMSO (*Ll*KdcA) / 5% (v/v) methyl tertiary-butyl ether (MTBE) (*Sc*PhPDC), 30 °C. Products were assigned by NMR-analysis by Lydia Walter (University Freiburg). The distribution is given after product purification of the mixed product from a preparative scale (therefore acetoin (**16**) and benzoin-analogues (**28**, **31**) were not detected).

	substrates		possible products			
A	indole-3-pyruvate	acetaldehyde	**16**	**26**	**27**	**28**
LlKdcA[a] (publication V)	preparative scale, product yield: 23%		–	–	100%	–
B	phenylpyruvate	acetaldehyde	**16**	**29**	**30**	**31**
ScPhPDC[b]	preparative scale, product yield: 14% + 6.3 mg aldol product		–	0.4 mg	6.8 mg	–
LlKdcA[c] (publication V)	preparative scale, product yield: 61%		–	20%	80%	–
C	phenylpyruvate	pyruvate	**16**	**29**	**30**	**31**
LlKdcA[d]	preparative scale, product yield: 32%		–	89%	11%	–

[a] 20 mM indole-3-pyruvate, 40 mM acetaldehyde, reaction time 104 h, distribution given in mol% (NMR);
[b] 20 mM phenylpyruvate, 20 mM acetaldehyde, reaction time 5 d, distribution given in weight [mg];
[c] 18 mM phenylpyruvate, 40 mM acetaldehyde, reaction time 26 h, distribution given in mol% (NMR);
[d] 18 mM phenylpyruvate, 18 mM pyruvate, reaction time 26 h, distribution given in mol% (NMR).

3.3.3 Carboligation of aliphatic donor- and aromatic acceptor aldehydes

Previously the mixed carboligation of diversely substituted aliphatic and aromatic aldehydes has been intensively studied, especially with *Pp*BFD and *Pf*BAL as biocatalysts. In both cases aromatic donor aldehydes are ligated with aliphatic acceptor aldehydes yielding HPP-analogues (chapter 1.5.1/1.5.5). One aim of the thesis was the enantioselective carboligation to substituted PAC-analogues in order to finally gain mandelic acid derivates after an additional enzymatic or chemical step (Fig. 3-6) which was devised and investigated by the group of Prof. Dr. Michael Müller, our cooperation partner at the Albert-Ludwigs-University Freiburg in an industrial project with Degussa AG (now Evonik, Essen). PAC-analogues have been obtained with PDCs and *Ll*KdcA as described above (chapter 3.2.3, ***publications IV*** and ***V***). Two possible pathways were followed (Fig. 3-6):
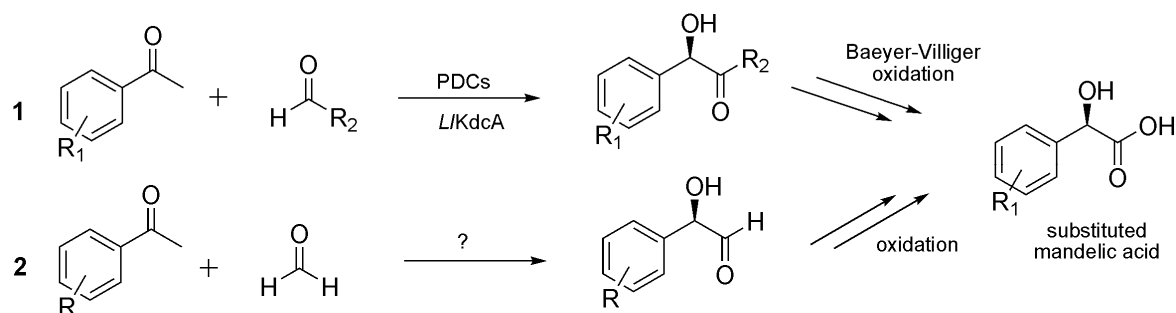
Fig. 3-6: Possible chemo-enzymatic approaches for the synthesis of chiral (*R*)-mandelic acids derivatives.

(**1**) Chemoselective carboligation of substituted benzaldehydes and aliphatic aldehydes to the (*R*)-PAC-analogue using various PDCs and *Ll*KdcA followed by a Baeyer-Villiger oxidation
(**2**) Search for an appropriate enzyme in the enzyme toolbox catalysing the ligation of substituted benzaldehydes as the acceptor with formaldehyde as the donor to the respective (*R*)-PAC-analogue followed by an additional oxidation step.

In this thesis we focused particularly on 3,5-substituted mandelic acid derivatives, which requires the selective acceptance of 3,5-substituted benzaldehyde derivatives as acceptor aldehydes by the respective enzyme.

Path 1: My part of the project was an enzyme toolbox-screen for biocatalysts catalysing the ligation of 3,5-dichlorobenzaldehyde (3,5-DChlBA) with various aliphatic aldehydes (acetaldehyde, propanal, isobutyric aldehyde, dimethoxyacetaldehyde), with isobutyric aldehyde and dimethoxyacetaldehyde being the most interesting donor aldehydes, since substitutions in Cα-position facilitate the subsequent Baeyer-Villiger oxidation. However, no enzyme was found to catalyse the ligation of 3,5-DChlBA with these branched-chain aliphatic aldehydes to the desired PAC-derivative. With *Pf*BAL high activities were observed for both aldehydes, but the product formed was again exclusively the 2-HPP-analogue. More promising results were obtained with acetaldehyde and propanal as donor substrates. The TTC-assay revealed almost all PDCs as well as *Ll*KdcA to catalyse the carboligation with 3,5-DChlBA. Detailed studies[*] using *Ap*PDC, *Ll*KdcA and *Zp*PDC demonstrated that the strong signal observed with *Ll*KdcA in the TTC-assay again originated from the HPP- and not from the PAC-derivative (chapter 3.2.3, Tab. 3-7, **24**). Although *Ap*PDC shows the desired PAC-analogue formation, the yield is very low. Due to the fact that no appropriate catalyst could be found for the desired reaction and the additional Baeyer-Villiger oxidation was more challenging than expected the work was finally focussed on the second path.

---

[*] Detailed instrumental analytics were conducted by Dr. Thomas Stillger and Lydia Walter, Albert-Ludwigs-University Freiburg.

Path 2: The carboligation of 3,5-DChlBA with formaldehyde seemed to be the most promising way to obtain the desired chiral mandelic acid derivative. *Pf*BAL accepts formaldehyde just as an acceptor (Demir *et al.*, 2004), yielding again the 2-HPP derivative but not the desired PAC-analogue. Screening of the enzyme toolbox for this reaction using the TTC-assay was performed with glyoxylate, the corresponding 2-keto acid of formaldehyde, which is easier to handle and not as volatile as formaldehyde. Best results were obtained with *Ll*KdcA and the variant *Zm*PDCI472A (Siegert *et al.*, 2005) (Fig 3-7).
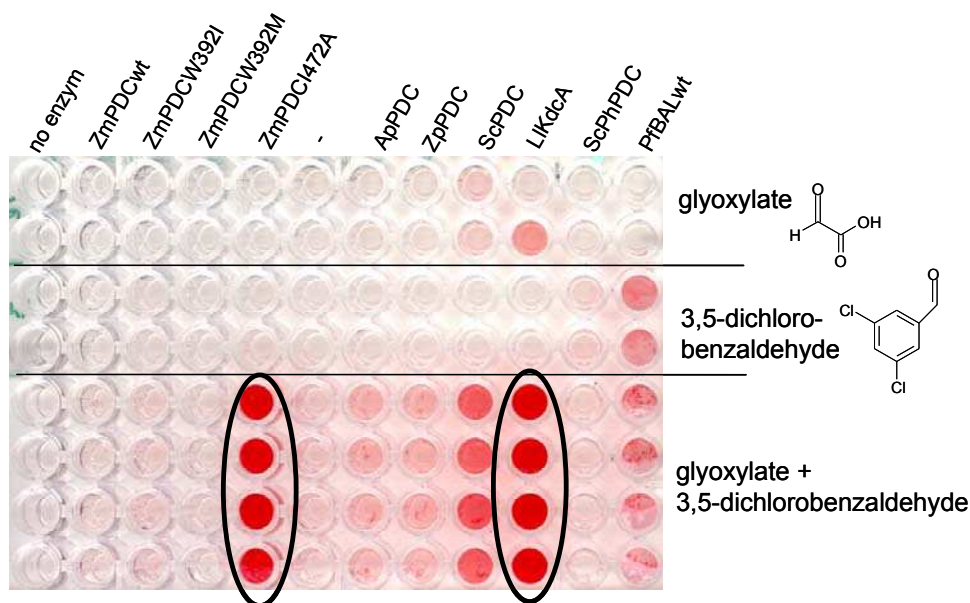
Fig. 3-7: TTC-assay monitoring the carboligation of formaldehyde (obtained by *in situ* decarboxylation of glyoxylate) and 3,5-dichlorobenzaldehyde. Four wild type pyruvate decarboxylases (PDCs), *Ll*KdcA, *Sc*PhPDC, *Pf*BAL and three variants of the *Zm*PDC are shown. As a negative control in the first column no catalyst was inserted. Reaction conditions: 50 µL (0.30 mg/mL enzyme) in 50 mM potassium phosphate buffer pH 7, 2.5 mM MgSO$_4$, 0.1 mM ThDP and 50 µL substrate solution (36 mM of each substrate, 40% (v/v) DMSO), incubation at 30°C for 18 h, 100 rpm. Scan taken after 1 min.

NMR-studies conducted by Dr. Ekaterina Gauchenova, Albert-Ludwigs-University Freiburg strongly hint to the formation of carboligation products, but the interpretation of the results was hindered since the products tend to form aggregates or polymers. Further investigations are currently performed by the Müller group (Dr. Ekaterina Gauchenova, personal communication).

Another new interesting pathway for the follow-up chemistry of substituted PAC-analogue products in general could be the transamination of the gained 2-hydroxy ketones by ω-transaminases towards achiral and chiral amines. Recently a screen of 16 different ω-transaminases was presented by the University of London, revealing the enzyme of the soil bacterium *Saccharopolyspora erythraea* to be able to convert 2-hydroxy ketones (Kaulmann *et al.*, 2007).

3.3.4 Accessing (*S*)-hydroxy ketones

Before this thesis was started the enzyme toolbox contained almost exclusively (*R*)-specific enzymes for the carboligation of benzoin derivatives and mixed aliphatic/aromatic 2-hydroxy ketones, with *Pp*BFD being the only exception coupling benzaldehyde and acetaldehyde to (*S*)-2-HPP with an *ee* of 92% (chapter 1.4.5) (Iding *et al.*, 2000). However, this (*S*)-selective carboligation does only work with acetaldehyde as the acceptor and fails with larger acceptor aldehydes like propanal (Fig. 3-8). The restriction to (*R*)-products limited the application range of the toolbox significantly. Therefore the molecular reasons for the strict (*R*)-selectivity of ThDP-dependent lyases were investigated in more detail in order to elucidate the underlying mechanistic principles allowing the introduction of (*S*)-selectivity by protein design. A molecular explanation for the exceptional behaviour of *Pp*BFD, based on the crystal structure, was recently suggested (Fig. 1-11, chapter 1.4.5). Here an *S*-pocket was identified that exactly fits the size of the small acetaldehyde side chain, when approaching the ThDP-bound aromatic donor aldehyde prior to the formation of the new C-C-bond (Knoll *et al.*, 2006). The hypothesis was affirmed by previous studies concerning the expansion of the substrate range of *Pp*BFD to accept *ortho*-substituted benzaldehydes as donors, where a variant *Pp*BFDM365L/L461S was identified catalysing the formation of various *ortho*-substituted (*S*)-2-HPP-analogues (Lingen *et al.*, 2003).

Shaping the *S*-pocket by site-directed mutagenesis of leucine 461 toward alanine and glycine in *Pp*BFD resulted in variants which open access to various (*S*)-hydroxy ketones with high enantioselectivity (***publication VII***, Fig. 3-8).
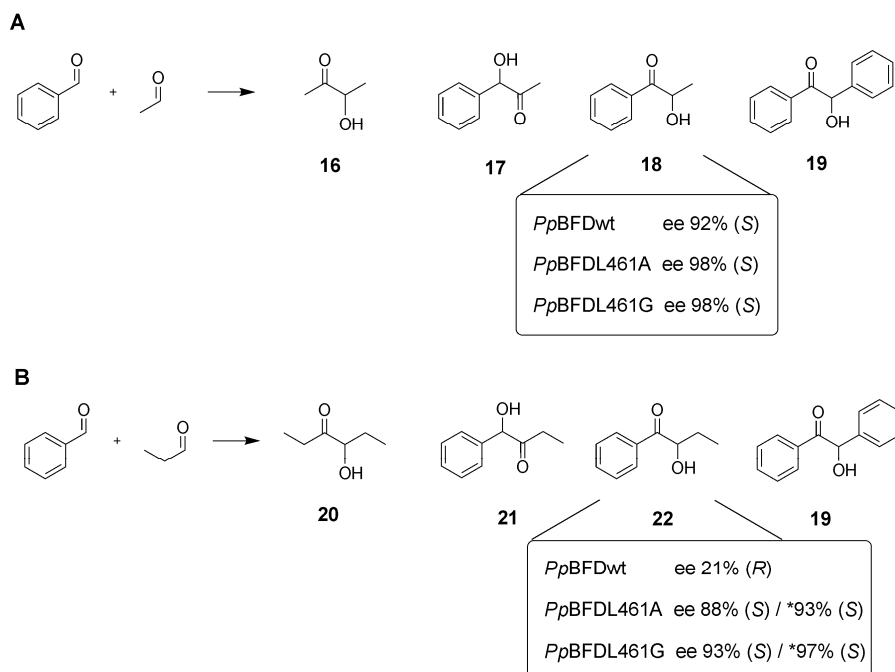
Fig. 3-8: C-C coupling of benzaldehyde with acetaldehyde (**A**) and propanal (**B**) catalysed by *Pp*BFDwt, *Pp*BFDL461A and *Pp*BFDL461G. Reaction conditions: 18 mM aromatic/aliphatic aldehyde, 50 mM potassium phosphate buffer, pH 7.0 (*pH 7.9), 2.5 mM MgSO$_4$, 0.1 mM ThDP, 20% (v/v) DMSO, 0.3 mg/mL purified enzyme, 30°C.

The biochemical data (***publication VII***) suggest that there is an optimal size for the acceptor aldehyde yielding high enantio- and chemoselectivity. If the acceptor is too large for the *S*-pocket, it is not optimally stabilised and a parallel arrangement relative to the donor aldehyde is an alternative (Fig. 1-11). Compared to benzaldehyde the stereo control in mixed carboligations with propanal is even better with 3,5-dimethoxybenzaldehyde and 3-cyanobenzaldehyde (Fig. 3-9). This shows that not only the perfect stabilisation of the acceptor aldehyde itself but also the interplay of donor and acceptor aldehydes fitting into the active site influences the chemo- and enantioselectivity of the biocatalysts. Further structural proofs for the *S*-pocket will be given in chapter 3.4.6.
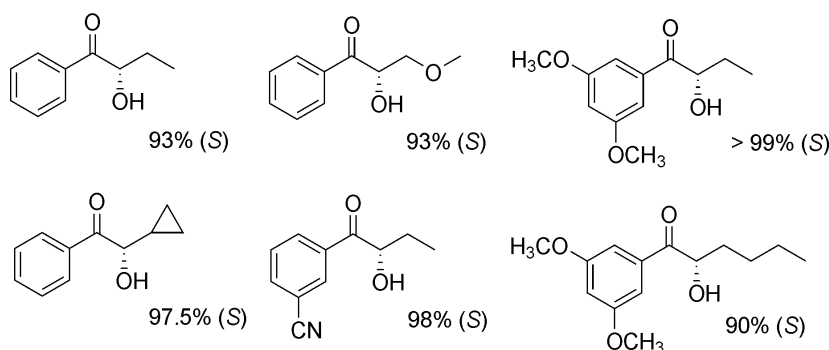
Fig. 3-9: Biocatalytically accessible 2-hydroxy ketones with high excess of the (*S*)-enantiomer catalysed by *Pp*BFDL461A (instrumental analysis performed by Lydia Walter, Albert-Ludwigs-University Freiburg, ***publication VII***). Reaction conditions: 18 mM aromatic/aliphatic aldehyde, 50 mM potassium phosphate buffer, pH 7.0, 2.5 mM MgSO$_4$, 0.1 mM ThDP, 20% (v/v) DMSO, 0.3 mg/mL purified enzyme, 30°C.

By accessing diversely substituted and enantio-complementary 2-hydroxy ketones the toolbox of versatile precursors for further bioorganic syntheses could be enlarged immensely. Among the new products are branched-chain aliphatic 2-hydroxy ketones and several substituted phenylacetylcarbinol-analogue products. By combining decarboxylase activity and carboligase activity the ligation of instable aldehydes like indole-3-acetaldehyde is possible using the corresponding 2-keto acid. Most interestingly the shaping of the *S*-pocket of the benzoylformate decarboxylase from *Pseudomonas putida* by site-directed mutagenesis now opens access to several (*S*)-2-hydroxy ketones.

*Ll*KdcA (*Sc*PhPDC)

*Ll*KdcA (?)   *Pp*BFDL461A/G   (*R*) = *Pf*BAL
(*S*) = *Ll*KdcA
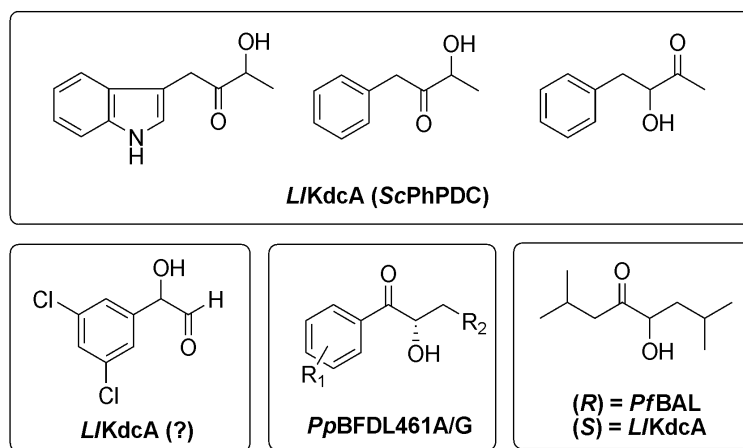
3.4 Investigation of structure-function relationships

3.4.1 Crystal structure analysis of *Ap*PDC, *Ll*KdcA and *Pp*BFDL461A

Crystals of *Ap*PDC and *Pp*BFDL461A could be grown and data on the well defracting crystals were collected and utilised for model building (Tab. 3-10).[*] The structures were solved to a resolution of 2.2 Å for *Pp*BFDL461A and 2.75 Å for *Ap*PDC (**publication IV and VII**). Additionally Catrine Berthold could finish the project by solving the structure of *Ll*KdcA (**publication VI**) to a resolution of 1.6 Å.

---

[*] Crystal growing, data collection (at MAX-lab, Lund, Sweden), processing, structure solution and crystallographic refinement were performed during a ten weeks research stay at the Institute of Structural biology, Karolinska Institutet in Stockholm, Sweden, under helpful guidance of Catrine Berthold.

Tab. 3-10: Most important collection and model statistics of the crystal structures of *Ap*PDC (publication IV), *Ll*KdcA (publication VI) and *Pp*BFDL461A (publication VII). Values in parentheses are given for the highest resolution interval.

	*Ap*PDC	*Ll*KdcA	*Pp*BFDL461A
data collection statistics			
space group	c2	$P2_12_12_1$	$P2_12_12_1$
nmol/asym unit	8	2	4
resolution [Å]	2.75 (2.9-2.75)	1.60 (1.69-1.60)	2.2 (2.32-2.2)
mean (I/σ(I))	10.8 (2.0)	11.3 (2.2)	8.8 (2.2)
completeness [%]	97.2 (98.5)	92.4 (63.8)	97.0 (99.2)
Wilson B-factor [Å^2]	69.7	18	27.8
structure refinement and final model statistics			
R-factor [%]	22.7	16.3	18.0
R-free [%]	24.5	20.0	22.8
Ramachandran plot			
most favoured [%]	91.2	91.2	92.0
additionally allowed [%]	8.6	8.6	8.0
generously allowed [%]	0.2	0.2	0
disallowed [%]	0	0	0

Concerning the secondary structure *Ap*PDC and *Ll*KdcA show the general fold of ThDP-dependent enzymes consisting of three domains with α/β-topology (chapter 1.4.3). The quaternary structure of *Ap*PDC displays the typical formation of a dimer of dimers, which was expected from the determination of the native molecular mass (***publication IV***) and is typical for ThDP-dependent decarboxylases. In contrast, *Ll*KdcA is exceptional among decarboxylases as it is active as a dimer, which was deduced from size-exclusion chromatography and supported by the crystal structure (***publication VI***).

3.4.2 Deducing general principles for chemo- and enantioselectivity

Based on the now available large structural and biochemical database structure-function studies are possible. The newly obtained 3D structures of *Ap*PDC and *Ll*KdcA were superimposed with the structures of *Sc*PDC, *Zm*PDC, *Pp*BFD and *Pf*BAL (Tab. 3-11) and the active sites as well as the channels connecting the active sites with the protein surface have been compared.[*] An overview of the assignment of amino acids lining the active site at similar positions in the various enzymes is given in Tab. 3-11.

---

[*] Comparisons were done during a research stay at the Institute of Technical Biochemistry, University of Stuttgart in cooperation with Michael Knoll.

Tab. 3-11: Residues lining the substrate channel and the active site of various ThDP-dependent enzymes. Corresponding residues are found in similar structural positions, if not otherwise indicated. For the positions of the respective areas see the schematic presentation (Fig. 3-10). Residues in parenthesis are situated at a similar position but are only partially involved in lining the specific site. pdb = RCBS protein data bank code (www.rcbs.org/pdb)

*Pf*BAL (pdb: 2ag0)	*Pp*BFD (pdb: 1mcz, 1bfd)	*Ap*PDC (pdb: 2vbi)	*Zm*PDC (pdb: 1zpd)	*Ll*KdcA (pdb: 2vbf)	*Sc*PDC (pdb: 1pyd)	available variants (references)
residues lining the *S*-pocket (area A1 in Fig. 3-10)						
L25	(N23)	V24	V24	V23	L25	
H26	P24	A25	A25	P24	P26	
G27	G25	G26	G26	G25	G27	*Pp*BFDG25A (Siegert, 2000)
A28	S26	D27	D27	D26	D28	*Pp*BFDS26A (Kneen *et al.*, 2005; Polovnikova *et al.*, 2003); *Pf*BALA28S (Janzen, 2002); *Zm*PDCD27E,N (Chang *et al.*, 1999; Wu *et al.*, 2000)
F484	F464	I472	I476	I465	I480	*Pp*BFDF464I (Siegert *et al.*, 2005)
(W487)	(Y458)	Y466	Y470	Y459	Y474	
-	L461	E469	E473	E462	E477	*Pp*BFDM365L/L461S (Lingen *et al.*, 2003); *Zm*PDCE473I,D,N,T,A (very low expression (Mesch, 1997))
entrance to the *S*-pocket (area A2 in Fig. 3-10)						
A480	A460	I468	I472	V461	I476	*Pp*BFDA460I, *Zm*PDCI472A (Siegert *et al.*, 2005); *Pp*BALA480I (Janzen, 2002)

Tab. 3-11: continued

*Pf*BAL	*Pp*BFD	*Ap*PDC	*Zm*PDC	*Ll*KdcA	*Sc*PDC	available variants (references)
further important residues lining the donor binding site (area B in Fig. 3-10)						
H415 (C414, L398) Y397	(T380) C398 F397	W388	W392	F381	A392	*Zm*PDCW392A (Bruhn *et al.*, 1995); *Zm*PDCW392M,I (Mesch, 1997)
residue stabilising the V-conformation of ThDP						
M421	L403	I411	I415	I404	I415	
proton relay system*						
H29 Q113	H70					*Pp*BFDH70 (Kneen *et al.*, 2005; Polovnikova *et al.*, 2003; Siegert, 2000); *Pf*BALH29A (Kneen *et al.*, 2005)
		H113 H114	H113 H114	H112 H113	H114 H115	*Zm*PDCH113/H114 (Bruhn, 1995; Huang *et al.*, 2001); *Sc*PDCH114F/H115F (Liu *et al.*, 2001)
α-helix covering the entrance to the substrate channel (area C in Fig. 3-10)						
aa 550-555 α-helix, 1.5 turns	no helix	aa 538-555 α-helix, 4 turns	aa 544-566 α-helix, >4 turns	aa 532-547 α-helix, 4 turns	aa 564-556 α-helix, 2-3 turns	*Zm*PDC C-terminal deletion variants (Chang *et al.*, 2000)

* Due to a backbone displacement the proton relay system in *Pp*BFD and *Pf*BAL is located at different sequence positions compared to PDCs and *Ll*KdcA.

As a result of the information gained by the superimposition, schematic models of the different steric properties in the substrate channel and the active sites of the enzymes could be developed. This allows the visualisation of hotspots regarding catalytic activity and selectivity. As obvious in Fig. 3-10 the catalysts differ in the size of the donor binding site and the substrate channel as well as in the existence and form of an *S*-pocket.
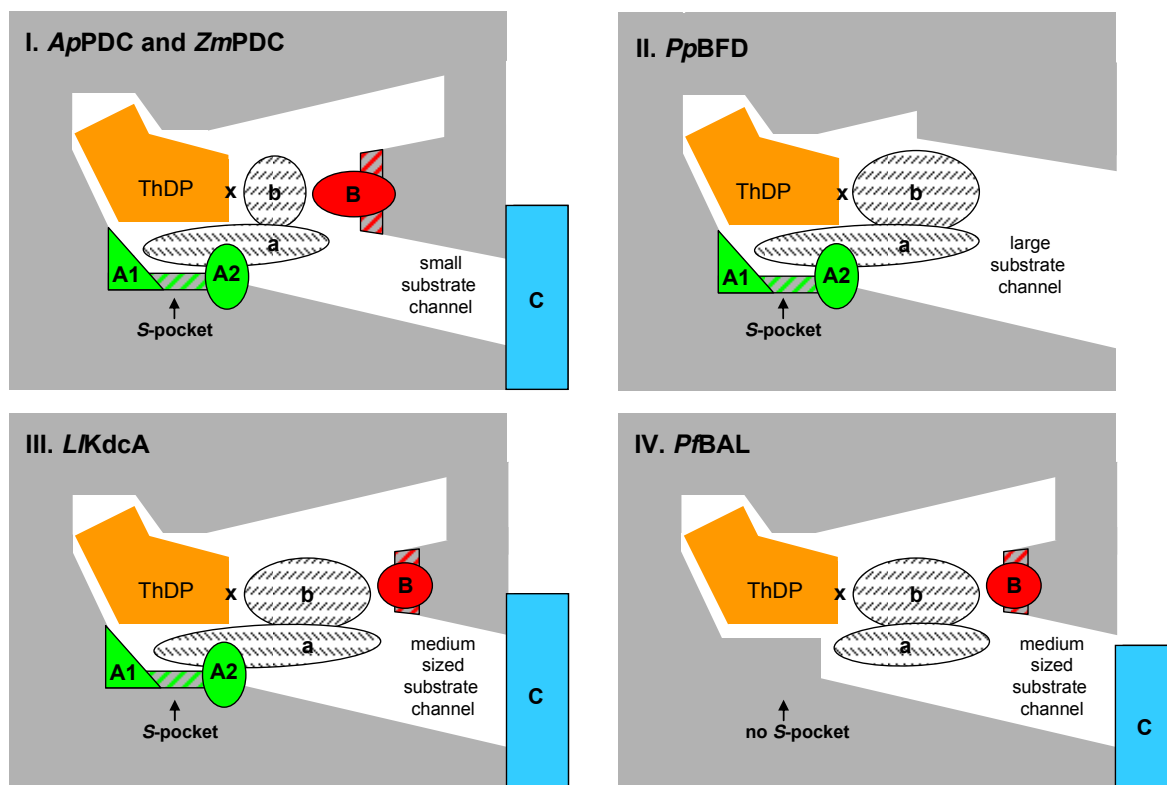
Fig. 3-10: Schematic presentation of the substrate channel and the active site of *Ap*PDC, *Zm*PDC (**I**), *Pp*BFD (**II**), *Ll*KdcA (**III**) and *Pf*BAL (**IV**). ThDP is bound in a V-conformation in the active site. x = C2-atom of the thiazolium ring where binding of a 2-keto acid (decarboxylation) or a donor aldehyde (carboligation) occurs; area **A** = residues lining the *S*-pocket; **A1** = most prominent residues defining the *S*-pocket's size, **A2** = residues defining the entrance to the *S*-pocket, **a** = acceptor binding site; **B** = bulky residue opposite to the donor binding site, **b** = binding site for 2-keto acid and donor aldehyde; **C** = C-terminal α-helix covering the entrance of the substrate channel.

In the following the different structural regions are comparatively discussed with respect to biochemical data in order to rationalise the structural base for the chemo- and enantio-selectivity of the various enzymes.

3.4.3 Substrate channel

In *Pp*BFD the channel entrance is clearly visible on the surface of the 3D structure (Knoll *et al.*, 2006). The situation is different in all other examined enzymes as a C-terminal α-helix covers this entrance (area C, Fig. 3-10). The helix is assumed to be flexible in solution in order to allow the substrates to enter the substrate channel. Furthermore it has some impact on the activity as was investigated for *Sc*PDC (König, 1998). This is

supported by results obtained with *Zm*PDC variants with a C-terminal truncation (Chang *et al.*, 2000). Whereas the deletion of the last seven C-terminal amino acids has no effect on the decarboxylase activity, the next few, R561 and S560, are critical not only for the decarboxylase activity resulting in a drastically decrease of k_{cat} for the decarboxylation of the main substrate pyruvate but also for the cofactor binding.

Beside the uncovered entrance *Pp*BFD also possesses a very wide and straight substrate channel. In contrast, *Ap*PDC, *Zm*PDC and *Sc*PDC show a less wide and curved substrate channel (Fig. 3-10). The size of the channel of *Pf*BAL and *Ll*KdcA is between that of the PDCs and *Pp*BFD and additionally slightly curved upwards (not educible in Fig. 3-10 III and IV).

3.4.4 Proton relay system

For *Pf*BAL and *Pp*BFD catalytically important residues have been identified, operating as proton acceptors and donors in the several protonation and deprotonation steps during the catalytic cycle (chapter 1.4.4). In *Pf*BAL these residues are H29 (Kneen *et al.*, 2005; Knoll *et al.*, 2006) and Q113 while H70 is the proton relay in *Pp*BFD (Kneen *et al.*, 2005; Polovnikova *et al.*, 2003; Siegert, 2000). In *Zm*PDC H113 and H114 are localised at a similar position (Tab. 3-11) and might have a proton acceptor/donor function as mutations at this position validate (Polovnikova *et al.*, 2003), yielding completely inactive variants (Bruhn, 1995; Huang *et al.*, 2001). *Ap*PDC, also containing histidine residues at positions 113 and 114 is probably comparable to *Zm*PDC. *Ll*KdcA has H112/H113 and *Sc*PDC H114/H115 (Liu *et al.*, 2001) as well as E477 at a similar position.

3.4.5 Donor binding site

Differences in the substrate range for both, the decarboxylase- and carboligase ability of the enzymes, are directly explainable by different sizes of the donor binding sites (area b in Fig. 3-10). The preference of PDCs for short aliphatic 2-keto acids as well as short aliphatic donor aldehydes is mainly a result of the residue located opposite to the C2-atom (area B in Fig. 3-10), which is occupied by a bulky tryptophane in the bacterial *Zm*PDC and *Ap*PDC (Tab. 3-11). However, this binding site in *Ap*PDC is larger than in *Zm*PDC allowing the binding of at least small amounts of benzaldehyde in this position (Fig. 3-11) which goes in line with the observed decarboxylation of benzoylformate (Tab. 3-3, **11**). Furthermore the enzyme shows a very week benzoin forming activity (Tab. 3-5 B). Although small amounts of aromatic donors can bind, smaller donor aldehydes are the preferred substrates in mixed ligase reactions yielding predominantly PAC-analogues.
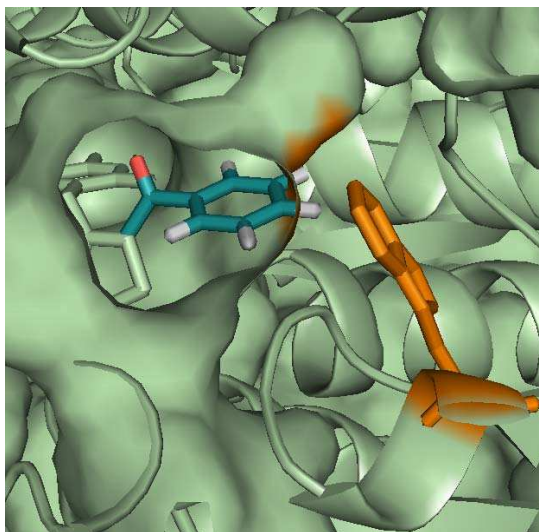
Fig. 3-11: Donor binding site of *Ap*PDC with benzaldehyde (blue) modelled to the C2-atom of the thiazolium ring of ThDP. In the crystal structure the hydrogen atoms of the aromatic ring interfere slightly with the aromatic site chain of the amino acid W388 (yellow) (picture generated by Michael Knoll, University of Stuttgart).

In *Pp*BFD as well as *Pf*BAL and *Ll*KdcA (Fig. 3-10, II-IV) this part of the active site is large enough to bind large aromatic substrates. However, the size of the binding site is not the only criterion to explain substrate transformation. Optimal stabilisation of a molecule in a binding site is also important to allow the rapid formation of an enzyme-substrate complex. This may explain why *Pp*BFD shows only very little decarboxylase activity towards aliphatic 2-keto acids (***publication VII***, supplementary material) and also little carboligase activity of aliphatic donor aldehydes although there is enough space available in the substrate binding site. In *Ll*KdcA the shape of the substrate binding pocket allows an optimal fit of the indole moiety, which makes *Ll*KdcA the only biocatalyst known so far which accepts indole-3-acetaldehyde in carboligation reactions (***publication VI***).

The various constraints in the different enzymes result in an almost selective formation of one predominant ligation product if the two aldehydes to be ligated are sufficiently different. Thus, in the case of an aliphatic and an aromatic aldehyde, PDCs prefer the binding of aliphatic aldehydes as donors, whereas *Pp*BFD prefers aromatic aldehydes. Especially in the case of *Ll*KdcA the phenomenon of optimal fitting is obvious. With this catalyst in mixed ligase reactions of benzaldehyde and acetaldehyde both aldehydes can act as donors or acceptors (chapter 3.2.3). Larger aliphatic aldehydes than acetaldehyde are preferably bound as acyl donors, since sterical hindrance with F382 and F542 makes their binding in the acceptor binding site less favourable (for details see ***publication VI***). Therefore the ligation of larger aliphatic aldehydes with benzaldehyde leads exclusively to the production of PAC-analogues (chapter 3.2.3). On the other hand, binding of a substituted benzene ring in the acceptor position would result in several unfavourable interactions with surrounding hydrophobic residues, which also prevents these compounds

to act as acceptors in the carboligation reaction, whereas they fit into the donor binding site. Consequently the carboligation of substituted aromatic aldehydes with acetaldehyde yields exclusively the HPP-analogues (chapter 3.2.3).

3.4.6 Acceptor binding site

After binding of the donor aldehyde to ThDP the acceptor aldehyde has to bind to the active site. This can only occur in a strictly defined area of the active site (area a in Fig. 3-10). There are predominantly two possibilities for the acceptor aldehyde to approach the ThDP-bound donor: in a parallel mode, with both side chains pointing into the substrate channel or in an antiparallel mode, with the acceptor side chain pointing into the so-called *S*-pocket (Fig. 1-11, chapter 3.3.4). This step might be not only decisive for the stereoselectivity of the ligation step but also for the chemoselectivity of the reaction, as principally both aldehydes present in the reaction mixture can also react as acceptors.

The S-pocket approach

As described in chapter 3.3.4 (*S*)-selectivity in *Pp*BFD can be explained by an existing *S*-pocket which is accessible by acetaldehyde in the wild type enzyme and was shaped by rational protein design to accept larger acceptor aldehydes. In this chapter the structural confirmation of the enlarged *S*-pocket in the variant *Pp*BFDL461A is described. Additional evidence will be provided that the *S*-pocket is not unique in *Pp*BFD as similar structural features are also present in other ThDP-dependent enzymes which opens access to a broad range of (*S*)-hydroxy ketones.

New (S)-hydroxy ketones with a PpBFD variant

In order to prove that the site-directed mutagenesis did not alter the 3D structure of *Pp*BFDL461A apart from the exchanged amino acid, the crystal structure of the variant was solved (Tab. 3-10) and compared to the structure of wild type *Pp*BFD (Fig. 3-12). Despite the desired increase of the *S*-pocket, no significant structural modifications have been detected (***publication VII***). In contrast to the wild type enzyme the *S*-pocket in the variant *Pp*BFDL461A offers optimal space for the ethyl group of propanal (Fig. 3-12 B) providing higher stereoselectivity (Fig. 3-8). The data show that enantioselectivity towards the (*S*)-product is predominantly a consequence of the optimal stabilisation of the acceptor aldehyde side chain in the *S*-pocket and is additionally influenced by the size of the donor aldehyde, as larger donor aldehydes, such as substituted benzaldehydes result in increased selectivity (Fig. 3-9, ***publication VII***).
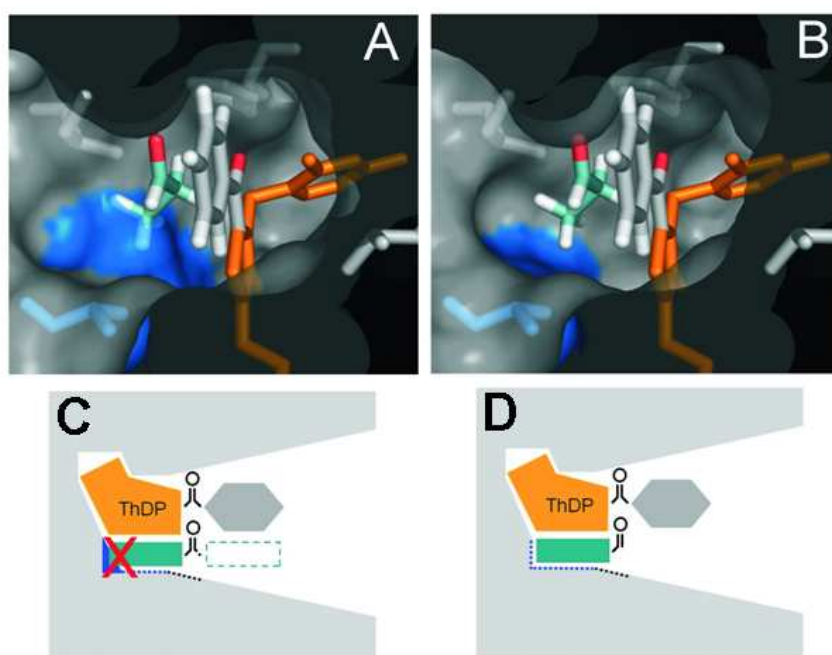
Fig. 3-12: Comparison of the active sites of *Pp*BFDwt (**A**) and *Pp*BFDL461A (**B**) with benzaldehyde and propanal modelled inside (from ***publication VII***). The side chain of the amino acid residue in position 461 is marked in blue. Benzaldehyde (grey) is bound to the C2-atom of the thiazolium ring (orange) and is coplanarly arranged due to steric and electronic demands. Propanal (light green) is located in the *S*-pocket. Models show a perfect stabilisation of the acceptor aldehyde in the *S*-pocket of the variant (**B, D**) while L461 causes sterical hindrance with propanal in *Pp*BFDwt (**A, C**). Consequently, in BFDwt propanal approaches predominantly parallel to benzaldehyde (**C**) yielding mainly the (*R*)-enantiomer while in *Pp*BFDL461A the perfect fit allows an antiparallel arrangement (**D**) leading almost exclusively to the (*S*)-product (***publication VII***).

S-pockets are widespread among ThDP-dependent enzymes but not always accessible

A superimposition of the crystal structures of *Pp*BFD, *Pf*BAL, *Zm*PDC, *Ap*PDC, *Sc*PDC and *Ll*KdcA revealed *S*-pockets in all structures except in *Pf*BAL (Fig. 3-10 IV). The *S*-pockets size increases in the series *Pp*BFD < *Ll*KdcA < *Zm*PDC / *Sc*PDC < *Ap*PDC. In all these enzymes except for *Pp*BFD a glutamate residue occupies the position analogous to *Pp*BFDL461 (area A1 in Fig 3-10, Tab. 3-11). The backbone of the three PDCs is almost identical in this area and the slightly different sizes of the pockets are just a result of different orientations of the side chains (***publication IV***).

From this structural comparison the question arose why only *Pp*BFD is able to catalyse the formation of (*S*)-2-HPP from benzaldehyde and acetaldehyde (chapter 1.5), whereas all other enzymes mentioned are (*R*)-selective in mixed reactions. One explanation is the chemoselectivity of the PDCs. Since small aliphatic aldehydes such as acetaldehyde are exclusively the donor aldehydes and benzaldehyde is the acceptor aldehyde, bulky aromatic aldehydes are usually too large for these *S*-pockets (see also next chapter). A further important factor for the strict (*R*)-selectivities are the almost completely blocked entrances of the *S*-pockets in *Ll*KdcA and in the PDCs by branched bulky residues such as isoleucine and valine (Fig. 3-13, Tab. 3-11), preventing the access of even small acceptor

aldehydes. Consequently (*S*)-2-hydroxy ketones could be formed by improving the access to these pockets. This has successfully been shown for the variant *Zm*PDCI472A which catalyses the formation of (*S*)-HPP (*ee* 70%) while exclusively (*R*)-PAC (*ee* > 98%) is formed with the wild type enzyme using benzaldehyde and acetaldehyde as substrates (Siegert *et al.*, 2005). Therefore it is most likely that *S*-selective variants of *Ap*PDC and *Sc*PDC are possible to be created by opening the access to the *S*-pockets via site-directed mutagenesis. *Ap*PDC, with the largest *S*-pocket, has the highest potential to accept large acceptor aldehydes.

enzyme	involved amino acid
S-pocket	
*Pp*BFD	L461
*Ap*PDC	E469
*Zm*PDC	E473
*Sc*PDC	E477
*Ll*KdcA	E462
entrance to S-pocket	
*Pp*BFD	A460
*Ap*PDC	I468
*Zm*PDC	I472
*Sc*PDC	I476
*Ll*KdcA	V461
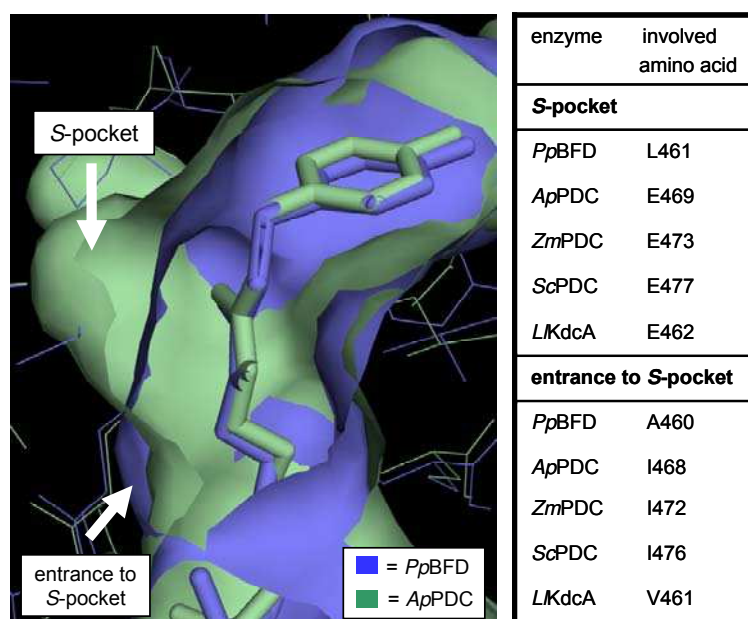
Fig. 3-13: Superimposition of wild type *Pp*BFD and *Ap*PDC. In contrast to *Pp*BFD, *Ap*PDC shows an enlarged *S*-pocket but the entrance is blocked by residue I468 (left). The amino acids mainly bordering the *S*-pocket as well as those defining the entrance to the pocket are given for *Pp*BFD, *Ap*PDC and additional for *Zm*PDC, *Sc*PDC and *Ll*KdcA (right) (**publication VII**).

The large *S*-pocket in *Ap*PDC is most probably also the reason for the comparative low enantioselectivity of the enzyme, yielding e.g. (*R*)-PAC (Tab. 3-6, **17**) with an enantiomeric excess of just 91%, whereas *Zm*PDC catalyses this reaction with 98% enantioselectivity (chapter 3.2.3). Modelling studies demonstrate that in the case of *Ap*PDC the *S*-pocket is large enough to bind even benzaldehyde (Fig. 3-14). Although the parallel arrangement of the donor and acceptor aldehyde, with the aromatic ring directed towards the substrate channel is preferred (Fig. 3-14 **B**), yielding (*R*)-PAC, benzaldehyde fits also into the *S*-pocket (Fig. 3-14 **D**). This antiparallel arrangement yields small amounts of (*S*)-PAC.
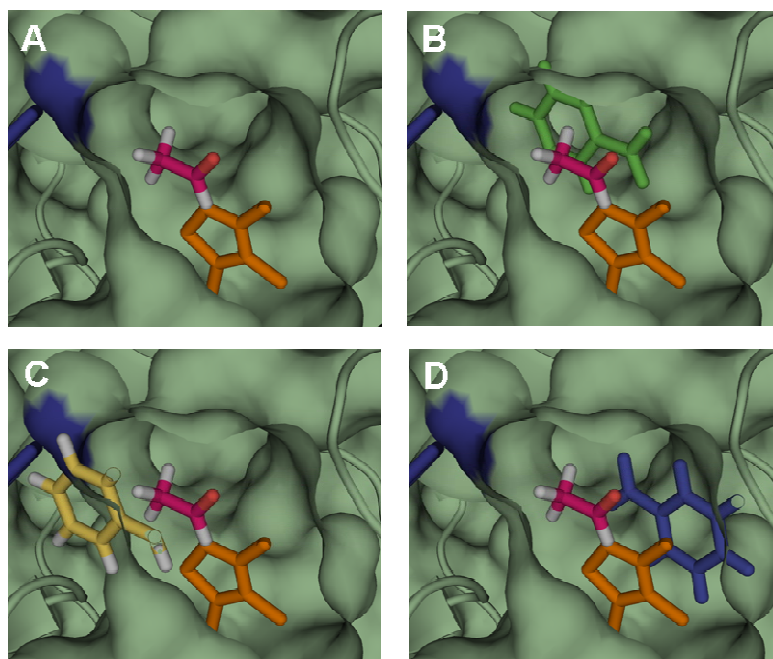
Fig. 3-14: Active site of *Ap*PDC with benzaldehyde and acetaldehyde modelled inside. The side chain of W388 is marked in blue. Acetaldehyde (pink) is bound in an almost coplanar arrangement to the C2-atom of the thiazolium ring (orange). (**A**) Without acceptor. (**B**) Main pathway: benzaldehyde (green) approaches towards ThDP-bound acetaldehyde in a parallel mode from the substrate channel, yielding (*R*)-PAC. (**C**) Due to sterical hindrance predominantly with W388 (blue) there is no possibility for benzaldehyde (yellow) to approach from the opposite site, which would theoretically also lead to (*S*)-PAC formation. (**D**) Explains the formation of low amounts of (*S*)-PAC, as the *S*-pocket in *Ap*PDC is large enough to bind benzaldehyde. However, this pathway is less favoured than **B** (picture generated by Michael Knoll, University of Stuttgart).

The biochemical data can very well be explained based on the structure. Furthermore the modelling studies allowed a reliable prediction of the chemoselectivity as well as enantioselectivity for all ThDP-dependent enzymes described in this thesis. However, prognoses are limited to mixed carboligation reactions of two substrates of significantly different size, e.g. an aromatic and an aliphatic aldehyde. Only in these cases preferences for the binding of the small and the large substrate can be assigned with sufficient reliability.

Carboligation of two similar aldehydes

In contrast, no enzyme of the current toolbox seems to be able to dissect sufficiently between e.g. two similar aromatic aldehydes with a different substitution pattern or two aliphatic aldehydes as their preferred binding as donor or acceptor is primarily a function of their steric demands. In such reactions usually a low chemoselectivity is observed.

A special case is the ligation of two small aliphatic aldehydes. The smaller the aldehydes, the more space they have in the active site and the less steric restrictions are present. This is especially obvious for the formation of acetoin from acetaldehyde. As demonstrated in Tab. 3-5 A, all tested enzymes show low enantioselectivity for this reaction. The PDCs produce (*S*)-acetoin in excess, probably making use of their *S*-pockets. *Pf*BAL which does

not have any *S*-pocket produces (*R*)-acetoin with an *ee* of only 40%. This example shows that there must be an alternative route to obtain aliphatic (*S*)-2-hydroxy ketones, probably according to the arrangement shown in Fig 3.14 C. This route is most likely just possible in the case of a large donor binding site in combination with little space demanding aliphatic aldehydes.

The conformation of the acceptor aldehyde bound in the active site could be demonstrated for *Ll*KdcA by implementing 2[(*R*)-1-hydroxyethyl]deazaThDP[*] (Leeper *et al.*, 2005), an inactive enantiomerically pure ThDP-analogue mimicking the covalent reaction inter-mediate 2-[(*R*)-1-hydroxyethyl]ThDP, which results from the binding of acetaldehyde to ThDP (**publication VI**). It was the first time that a donor aldehyde, in this case acetaldehyde, could be co-crystallised as an intermediate bound to the C2-atom of the thiazolium ring. While a coplanar arrangement of an aromatic donor aldehyde bound to ThDP is most likely, the prediction of the hydroxyethyl-moiety of an aliphatic donor aldehyde is less specific due to less constraints. Nevertheless the crystal structure shows an arrangement of the hydroxyethyl-moiety which is very similar to the arrangement predicted for aromatic ThDP-bound aldehydes, with the methyl group being directed into the substrate channel and the α–hydroxyl group forming a hydrogen bond to the 4'-amino group of the pyrimidine ring of ThDP (**publication VI**). As alternative conformations of ThDP-bound aliphatic aldehydes can be assumed depending on the size of the active centre, more structural data, preferably of co-crystallised donor aldehydes, are necessary to understand the conformation of ThDP-bound aliphatic aldehydes better. It can be assumed that the orientation of the aliphatic donor aldehyde will most probably be also influenced by the acceptor aldehyde.

The new crystal structures of *Ll*KdcA, *Ap*PDC and the variant BFDL461A increase the structure information about the sensitive parts of the active centre: the binding sites for the donor and acceptor aldehyde during carboligation. Based on these data a mapping of the active sites of most ThDP-dependent enzymes used for carboligation in this thesis was possible, which explains the observed chemo- and enantioselectivity of the carboligase reactions as a result of optimal stabilisation of the substrates in the active site. The results furthermore pave the way to expand the shaping strategy of *S*-pockets to a broad range of other 2-keto acid decarboxylases. This is a powerful tool to enlarge the toolbox of enzymatically accessible 2-hydroxy ketones by (*S*)-enantio-mers therewith providing a valuable platform for chemo-enzymatic synthesis.

---

[*] This inhibitor was kindly provided by Prof. Finian Leeper from the University of Cambridge.

4 CONCLUSIONS AND FUTURE PERSPECTIVES

With the extension of the enzyme toolbox the 2-hydroxy ketone platform was strongly expanded. Thus, apart from diversely substituted aromatic, aliphatic and branched-chain (*R*)-2-hydroxy ketones also various (*S*)-products are now available. The results provide a profound understanding of the stereocontrol in ThDP-dependent enzymes, thereby supporting and expanding a theoretic model which was initially only based on *Pp*BFD.

During the course of this thesis the database of available crystal structures was expanded by three further structures of *Ap*PDC, *Ll*KdcA and *Pp*BFDL461A. Together with the already known crystal structures and the biochemical data the relevant factors for the enantioselectivity in different enzymes are consistently explainable by the orientation of the two substrates in the active centre prior to formation of a new C-C bond. Based on the proved *S*-pockets, which have been identified in several 2-keto acid decarboxylases the combination of model studies and rational protein design is a promising tool to open access to a broad range of (*S*)-2-hydroxy ketones. Especially the opening of the entrances to the large pockets in the PDCs has a big potential.

Several affords were made in order to visualise the binding of donor and acceptor aldehyde in the active site, however trials to soak crystals with aldehydes, especially benzaldehyde, were not successful so far. Still crystals with both, the donor and acceptor aldehyde arranged in the active site, would give the final proof about speculations of the enzymes chemo- and enantioselectivity and are therefore a worth target of further trials. As a first step some success could be already achieved with a transition state analogue of an aliphatic aldehyde bound to a inhibitory ThDP-analogue.

To expand the potential of the toolbox further, next strategies should focus on the implementation of further enzymatic steps to modify the 2-hydroxy ketone building blocks by e.g. reduction of the keto group by dehydrogenases, oxidases or reductive amination using ω-transaminases, thereby offering an onion skin-like chiral building block toolbox.

5 REFERENCES

Adam, W., Lazarus, M., Saha-Möller, C. R. & Schreier, P. (1999). Biocatalytic synthesis of optically alpha-oxyfunctionalized carbonyl compounds. Acc. Chem. Res., 32, 837-845.

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.

Ansorge-Schumacher, M. B., Greiner, L., Schroeper, F., Mirtschin, S. & Hischer, T. (2006). Operational concept for the improved synthesis of (R)-3,3'-furoin and related hydrophobic compounds with benzaldehyde lyase. Biotechnol. J., 1, 564-568.

Arjunan, P., Nemeria, N., Brunskill, A., Chandrasekhar, K., Sax, M., Yan, M., Jordan, F., Guest, J. R. & Furey, W. (2002). Structure of the pyruvate dehydrogenase multienzyme complex E1 component from *Escherichia coli* at 1.85 Å resolution. Biochemistry, 41, 5213-5221.

Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D. & Jordan, F. (1996). Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* at 2.3 Å resolution. J. Mol. Biol., 256, 590-600.

Arnold, F. H. (1998). When blind is better: Protein design by evolution. Nat. Biotechnol., 16, 617-618.

Asakawa, T., Wada, H. & Yamano, T. (1968). Enzymatic conversion of phenylpyruvate to phenylacetate. Biochim. Biophys. Acta, 170, 375-391.

Azerad, R. (1995). Application of biocatalysts in organic synthesis. Bull. Soc. Chim. Fr., 132, 17-51.

Barrowman, M. M. & Fewson, C. A. (1985). Phenylglyoxylate decarboxylase and phenylpyruvate decarboxylase from *Acinetobacter calcoaceticus*. Curr. Microbiol., 12, 235-239.

Barrowman, M. M., Harnett, W., Scott, A. J., Fewson, C. A. & Kusel, J. R. (1986). Immunological comparison of microbial TPP-dependent non-oxidative alpha-keto acid decarboxylases. FEMS Microbiol. Lett., 34, 57-60.

Baykal, A. T., Chakraborty, S., Dodoo, A. & Jordan, F. (2006). Synthesis with good enantiomeric excess of both enantiomers of α-ketols and acetolactates by two thiamin diphosphate dependent decarboxylases. Bioorg. Chem., 34, 380-392.

Bera, A. K., Polovnikova, L. S., Roestamadji, J., Widlanski, T. S., Kenyon, G. L., McLeish, M. J. & Hasson, M. S. (2007). Mechanism-based inactivation of benzoylformate decarboxylase, a thiamin diphosphate-dependent enzyme. J. Am. Chem. Soc., 129, 4120-4121.

Bettendorff, L. & Wins, P. (2004). New perspectives on the cellular role of thiamine triphosphate and thiamine triphosphatase. Marcel Dekker Inc., New York/Basel.

BLAST. 2007. Basic Local Alignment Search Tool. www.ncbi.nlm.nih.gov/blast/.

BMBF (2007). Weiße Biotechnologie - Chancen für neue Produkte und umweltschonende Prozesse. Bundesministerium für Bildung und Forschung (www.biotechnologie.de), Bonn/Berlin.

Bornemann, S., Crout, D. H. G., Dalton, H., Hutchinson, D. W., Dean, G., Thomson, N. & Turner, M. M. (1993). Stereochemistry of the formation of lactaldehyde and acetoin produced by the pyruvate decarboxylase of yeast (*Saccharomyces sp.*) and *Zymomonas*

mobilis: different Boltzmann distributions between bound forms of the electrophile, acetaldehyde, in the two enzymatic reactions. J. Chem. Soc. Perkin Trans. I, 309-311.

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. I, 1, 425 - 430.

Bornscheuer, U. T. & Kazlauskas, R. J. (2004). Untreue Enzyme in der Biokatalyse: mit alten Enzymen zu neuen Bindungen und Synthesewegen. Angew. Chem., 116, 6156-6165.

Breslow, R. (1957). Rapid deuterium exchange in thiazolium salts. J. Am. Chem. Soc., 79, 1762.

Breuer, M., Ditrich, K., Habicher, T., Hauer, B., Keßeler, M., Stürmer, R. & Zelinski, T. (2004). Industrial methods for the production of optically active intermediates. Angew. Chem., Int. Ed., 43, 788-824.

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.

Bringer-Meyer, S., Schimz, K. L. & Sahm, H. (1986). Pyruvate decarboxylase from *Zymomonas mobilis*. Isolation and partial characterization. Arch. Microbiol., 146, 105-110.

Brosi, H. (2006). Klonierung und Charakterisierung neuer Thiamindiphosphat-abhängiger Enyzme. Diploma thesis. RWTH Aachen, Aachen.

Bruhn, H. (1995). Verbesserung der Acyloinkondensationsfähigkeit der Pyruvatdecarboxylase aus *Zymomonas mobilis*. Doctoral thesis. Heinrich-Heine University, Düsseldorf.

Bruhn, H., Pohl, M., Grötzinger, J. & Kula, M. R. (1995). The replacement of Trp392 by alanine influences the decarboxylase/carboligase activity and stability of pyruvate decarboxylase from *Zymomonas mobilis*. Eur. J. Biochem., 234, 650-655.

Buchner, E. (1897). Alkoholische Gährung ohne Hefezellen. Ber. Dt. Chem. Ges., 30, 117-124.

Campbell, N. A. (1997). Biologie. Spektrum Akademischer Verlag, Heidelberg/Berlin/Oxford.

Candy, J. M. & Duggleby, R. G. (1994). Investigation of the cofactor-binding site of *Zymomonas mobilis* pyruvate decarboxylase by site-directed mutagenesis. Biochem. J., 300 (Pt 1), 7-13.

Candy, J. M. & Duggleby, R. G. (1998). Structure and properties of pyruvate decarboxylase and site-directed mutagenesis of the *Zymomonas mobilis* enzyme. Biochim. Biophys. Acta, 1385, 323-338.

Candy, J. M., Koga, J., Nixon, P. F. & Duggleby, R. G. (1996). The role of residues glutamate-50 and phenylalanine-496 in *Zymomonas mobilis* pyruvate decarboxylase. Biochem. J., 315 (Pt 3), 745-751.

Chang, A. K., Nixon, P. F. & Duggleby, R. G. (1999). Aspartate-27 and glutamate-473 are involved in catalysis by *Zymomonas mobilis* pyruvate decarboxylase. Biochem. J., 339, 255-260.

Chang, A. K., Nixon, P. F. & Duggleby, R. G. (2000). Effects of deletions at the carboxyl terminus of *Zymomonas mobilis* pyruvate decarboxylase on the kinetic properties and substrate specificity. Biochemistry, 39, 9430-9437.

Chen, G. C. & Jordan, F. (1984). Brewers' yeast pyruvate decarboxylase produces acetoin from acetaldehyde: a novel tool to study the mechanism of steps subsequent to carbon dioxide loss. Biochemistry, 23, 3576-3582.

CLIB2021 (2007). CLIB 2021: Cluster industrielle Biotechnologie. Clusterwettbewerb der Bioindustrie gefördert durch das Bundesministerium für Bildung und Forschung. http://www.bmbf.de/pub/clib2021_nrw.pdf.

Costacurta, A., Keijers, V. & Vanderleyden, J. (1994). Molecular cloning and sequence analysis of an *Azospirillum brasilense* indole-3-pyruvate decarboxylase gene. Mol. Gen. Genet. (today Mol. Genet. Genomics), 243, 463-472.

Costacurta, A. & Vanderleyden, J. (1995). Synthesis of phytohormones by plant-associated bacteria. Crit. Rev. Microbiol., 21, 1-18.

Crout, D. H. G., Dalton, H., Hutchinson, D. W. & Miyagoshi, M. (1991). Studies on pyruvate decarboxylase: acyloin formation from aliphatic, aromatic and heterocyclic aldehydes. J. Chem. Soc., Perkin Trans. I, 1329-1334.

Curci, R., D'Accolti, L., Dinoi, A., Fusco, C. & Rosa, A. (1996). Selective oxidation of o-isopropylidene derivatives of diols to 2-hydroxy ketones employing dioxiranes. Tetrahedron Lett., 37, 115-118.

Dalby, P. A. (2003). Optimising enzyme function by directed evolution. Curr. Opin. Struct. Biol., 13, 1-6.

Dale, B. E. (2003). Greening the chemical industry: research and development priorities for biobased industrial products. J. Chem. Technol. Biotechnol., 78, 1093-1103.

Dawes, E. A., Ribbons, D. W. & Large, P. J. (1966). The route of ethanol formation in *Zymomonas mobilis*. Biochem. J., 98, 795-803.

Demir, A. S., Ayan, P., Igdir, A. C. & Guygu, A. (2004). Enzyme catalyzed hydroxymethylation of aromatic aldehydes with formaldehyde. Synthesis of hydroxyacetophenones and (*S*)-benzoins. Tetrahedron, 60, 6509-6512.

Demir, A. S., Dünnwald, T., Iding, H., Pohl, M. & Müller, M. (1999). Asymmetric benzoin reaction catalyzed by benzoylformate decarboxylase. Tetrahedron: Asymmetry, 10, 4769-4774.

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., Şeşenoglu, 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., Şeşenoglu, 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.

Dietrich, A. & König, S. (1997). Substrate activation behaviour of pyruvate decarboxylase from *Pisum sativum* cv. Miko. FEBS Lett., 400, 42-44.

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.

Domínguez de María, P., Pohl, M., Gocke, D., Gröger, H., Trauthwein, H., Walter, L. & Müller, M. (2007). Asymmetric synthesis of aliphatic 2-hydroxy ketones by enzymatic carboligation of aldehydes. Eur. J. Org. Chem., 2940-2944.

Domínguez de María, P., Stillger, T., Pohl, M., Wallert, S., Drauz, K., Gröger, H., Trauthwein, H. & Liese, A. (2006). Preparative enantioselective synthesis of benzoins and (*R*)-2-hydroxy-1-phenylpropane using benzaldehyde lyase. J. Mol. Catal. B: Enzym., 38, 43-47.

DSHS Köln. 2007. Institut für Biochemie der Deutschen Sporthochschule Köln IOC akkreditierts Labor für Dopinganalytik. www.dshs-koeln\biochemie.

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.

Dünkelmann, P., Pohl, M. & Müller, M. (2004). Enantiomerically pure 2-hydroxy carbonyl compounds through enzymatic C-C bond formation. Chim. Oggi/Chemistry Today supplement Chiral Catalysis, 22, 24-28.

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., 2000, 2161-2170.

Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B. & Jordan, F. (1993). Catalytic centers in the thiamin diphosphate dependent enzyme pyruvate decarboxylase at 2.4-Å resolution. Biochemistry, 32, 6165-6170.

Enders, D. & Kallfass, U. (2002). An efficient nucleophilic carbene catalyst for the asymmetric benzoin condensation. Angew. Chem., Int. Ed., 41, 1743-1745.

Entner, N. & Doudoroff, M. (1952). Glucose and gluconic acid oxidation of *Pseudomonas saccharophila*. J. Biol. Chem., 196, 853-862.

Faber, K. (2000). Biotransformations in organic chemistry, 4th ed. Springer-Verlag, Berlin/Heidelberg.

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, Q. F., Han, Z., Grover, P., Kessler, D., Senanayake, C. H. & Wald, S. A. (2000). Rapid access to enantiopure bupropion and its major metabolite by stereospecific nucleophilic substitution on an α-ketotriflate. Tetrahedron: Asymmetry, 11, 3659-3663.

Ferrer, M., Martínez-Abarca, F. & Golyshin, P. N. (2005). Mining genomes and 'metagenomes' for novel catalysts. Curr. Opin. Biotechnol., 16, 588-593.

Fessner, W.-D. & Walter, C. (1997). Enzymatic C-C bond formation in asymmetric synthesis. Topics Curr. Chem., 184, 97-194

Festel, G., Knöll, J., Götz, H. & Zinke, H. (2004). Der Einfluss der Biotechnologie auf die Produktionsverfahren in der Chemieindustrie. Chem. Ing. Tech., 76, 307-312.

Fischer, E. (1894). Einfluß der Conformation auf die Wirkung der Enzyme. Chem. Ber., 27, 2985-2993.

Forlani, G. (1999). Purification and properties of a pyruvate decarboxylase from *Zea mays* cultured cells. Phytochemistry, 50, 1305-1310.

Frank, R. A., Titman, C. M., Pratap, J. V., Luisi, B. F. & Perham, R. N. (2004). A molecular switch and proton wire synchronize the active sites in thiamine enzymes. Science, 306, 872-876.

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, 829-905.

Frey, P. & Hegeman, A. D. (2006). Enzymatic Reactions Mechnisms. Oxford University Press Inc., Oxford.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res., 31, 3784-3788.

Goetz, G., Iwan, P., Hauer, B., Breuer, M. & Pohl, M. (2001). Continuous production of (*R*)-phenylacetylcarbinol in an enzyme-membrane reactor using a potent mutant of pyruvate decarboxylase from *Zymomonas mobilis*. Biotechnol. Bioeng., 74, 317-325.

González, B., Merino, A., Almeida, M. & Vicuña, R. (1986). Comparative growth of natural bacterial isolates on various lignin-related compounds. Appl. Environ. Microbiol., 52, 1428-1432.

González, B. & Vicuña, R. (1989). Benzaldehyde lyase, a novel thiamine PPi-requiring enzyme, from *Pseudomonas fluorescens* biovar I. J. Bacteriol., 171, 2401-2405.

Graf, T. (2005). Charakterisierung der Pyruvatdecarboxylase aus *Acetobacter pasteurianus*. Diploma thesis (FH). Fachhochschule Aachen, Jülich.

Guo, Z., Goswami, A., Mirfakhrae, K. D. & Patel, N. (1999). Asymmetric acyloin condensation catalyzed by phenylpyruvate decarboxylase. Tetrahedron: Asymmetry, 10, 4667-4675.

Guo, Z., Goswami, A., Nanduri, V. B. & Patel, R. N. (2001). Asymmetric acyloin condensation catalysed by phenylpyruvate decarboxylase. Part 2: Substrate specificity and purification of the enzyme. Tetrahedron: Asymmetry, 12, 571-577.

Hasson, M. S., Muscate, A., Henehan, G. T., Guidinger, P. F., Petsko, G. A., Ringe, D. & Kenyon, G. L. (1995). Purification and crystallization of benzoylformate decarboxylase. Protein Sci., 4, 955-959.

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 Å 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*. III. Isolation and properties of constitutive mutants. J. Bacteriol., 91, 1161-1167.

Held, M., Schmid, A., van Beilen, J. B. & Witholt, B. (2000). Biocatalysis. Biological systems for the production of chemicals. Pure Appl. Chem., 72, 1337-1343.

Henning, H., Leggewie, C., Pohl, M., Müller, M., Eggert, T. & Jaeger, K. E. (2006). Identification of novel benzoylformate decarboxylases by growth selection. Appl. Environ. Microbiol., 72, 7510-7517.

Hildebrand, F., Kühl, S., Pohl, M., Vasic-Racki, D., Müller, M., Wandrey, C. & Lütz, S. (2007). The production of (*R*)-2-hydroxy-1-phenyl-propan-1-one derivatives by benzaldehyde lyase from *Pseudomonas fluorescens* in a continuously operated membrane reactor. Biotechnol. Bioeng., 96, 835-843.

Hischer, T., Gocke, D., Fernández, M., Hoyos, P., Alcántara, A. R., Sinisterra, J. V., Hartmeier, W. & Ansorge-Schumacher, M. B. (2005). Stereoselective synthesis of novel benzoins catalysed by benzaldehyde lyase in a gel-stabilised two-phase system. Tetrahedron, 61, 7378-7383.

Hodgson, J. (1994). The changing bulk biocatalyst market. Bio/Technology, 12, 789-790.

Hoffritz, J. (2007). Faible für Vampire; Vom Pionier zum Resteverwerter: Die Biotech-Branche lebt von alten Patenten der Pharmakonzerne - mit wechselndem Erfolg. Die Zeit/Wirtschaft, 09.08.2007.

Holloway, P. & Subden, R. E. (1993). The isolation and nucleotide sequence of the pyruvate decarboxylase gene from *Kluyveromyces marxianus*. Curr. Genet., 24, 274-277.

Holzer, H., Schultz, G., Villar-Palai, C. & Jüntgen-Schell, J. (1956). Isolierung der Hefecarboxylase und Untersuchung über die Aktivität des Enzyms in lebenden Zellen. Biochem. Zeitschr., 327, 331-344.

Horn, S. J., Aasen, I. M. & Østgaard, K. (2000). Production of ethanol from mannitol by *Zymobacter palmae*. J. Ind. Microbiol. Biotechnol., 24, 51-57.

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 Mol. Biol., 31, 761-770.

Huang, C. Y., Chang, A. K., Nixon, P. F. & Duggleby, R. G. (2001). Site-directed mutagenesis of the ionizable groups in the active site of *Zymomonas mobilis* pyruvate decarboxylase: effect on activity and pH dependence. Eur. J. Biochem., 268, 3558-3565.

Hübner, G., Weidhase, R. & Schellenberger, A. (1978). The mechanism of substrate activation of pyruvate decarboxylase. A first approach. Eur. J. Biochem., 92, 175-181.

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.

Iwan, P., Goetz, G., Schmitz, S., Hauer, B., Breuer, M. & Pohl, M. (2001). Studies on the continuous production of (*R*)-(-)-phenylacetylcarbinol in an enzyme-membrane reactor. J. Mol. Catal. B: Enzym., 11, 387-396.

Janzen, E. (2002). Die Benzaldehydlyase aus *Pseudomonas fluorescens*. Doctoral thesis. Heinrich-Heine University, Düsseldorf.

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.

Jordan, F. (2003). Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. Nat. Prod. Rep., 20, 184-201.

Jordan, F., Liu, M., Sergienko, E., Zhang, Z., Brunskill, A., Arjunan, P. & Furey, W. (2004). Yeast pyruvate decarboxylase: new features of the structure and mechanism. Marcel Dekker Inc., New York / Basel.

Kakeya, H., Morishita, M., Koshino, H., Morita, T., Kobayashi, K. & Osada, H. (1999). Cytoxazone: a novel cytokine modulator containing a 2-oxazolidinone ring produced by *Streptomyces* sp. J. Org. Chem., 64, 1052-1053.

Kaulmann, U., Smithies, K., Smith, M. E. B., Hailes, H. C. & Ward, J. M. (2007). ω-Transaminases for the biotransformation of ketones, aldehydes and ketodiols. Poster presentation (P45). Presented at the Biotrans 2007, Oviedo, Spain.

Kazlauskas, R. J. (2005). Enhancing catalytic promiscuity for biocatalysis. Curr. Opin. Chem. Biol., 9, 195-201.

Kazlauskas, R. J. (2000). Molecular modeling and biocatalysis: explanations, predictions, limitations, and opportunities. Curr. Opin. Chem. Biol., 4, 81-88.

Kern, D., Kern, G., Neef, H., Tittmann, K., Killenberg-Jabs, M., Wikner, C., Schneider, G. & Hübner, G. (1997). How thiamine diphosphate is activated in enzymes. Science, 275, 67-70.

Kihumbu, D., Stillger, T., Hummel, W. & Liese, A. (2002). Enzymatic synthesis of all stereoisomers of 1-phenylpropane-1,2-diol. Tetrahedron: Asymmetry, 13, 1069-1072.

Kirk, O., Borchert, T. V. & Fuglsang, C. C. (2002). Industrial enzyme applications. Curr. Opin. Biotechnol., 13, 345-351.

Kluger, R. (1987). Thiamine diphosphate: a mechanistic update on enzymic and nonenzymic catalysis of decarboxylation. Chem. Rev., 87, 863-876.

Kneen, M. M., Pogozheva, I. D., Kenyon, G. L. & McLeish, M. J. (2005). Exploring the active site of benzaldehyde lyase by modeling and mutagenesis. Biochim. Biophys. Acta, 1753, 263-271.

Knight, R. L. & Leeper, F. J. (1998). Comparison of chiral thiazolium and thiazolium salts as asymmetric catalysts for the benzoin condensation. J. Chem. Soc., Perkin Trans. I, 12, 1891-1893.

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.

Koga, J. (1995). Structure and function of indolepyruvate decarboxylase, a key enzyme in indole-3-acetic acid biosynthesis. Biochim. Biophys. Acta, 1249, 1-13.

Koga, J., Adachi, T. & Hidaka, H. (1992). Purification and characterization of indolepyruvate decarboxylase. A novel enzyme for indole-3-acetic acid biosynthesis in *Enterobacter cloacae*. J. Biol. Chem., 267, 15823-15828.

König, S. (1998). Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. Biochim. Biophys. Acta, 1385, 271-286.

Kühne, W. (1876). Über das Verhalten verschiedener organisirter und sog. ungeformter Fermente.-Über das Trypsin (Enzym des Pankreas). FEBS Lett., 62, 3-7.

Kurlemann, N. & Liese, A. (2004). Immobilization of benzaldehyde lyase and its application as a heterogeneous catalyst in the continuous synthesis of a chiral 2-hydroxy ketone. Tetrahedron: Asymmetry, 15, 2955-2958.

Kurniadi, T., Bel Rhlid, R., Fay, L. B., Juillerat, M. A. & Berger, R. G. (2003). Chemoenzymatic synthesis of aroma active 5,6-dihydro- and tetrahydropyrazines from aliphatic acyloins produced by baker's yeast. J. Agric. Food Chem., 51, 3103-3107.

Kutter, S., Weiss, M., Wille, G., Golbik, R., Spinka, M. & König, S. (2008). Covalently bound substrate at the regulatory site triggers allosteric enzyme activation. Nat. Proc., precedings.nature.com.

Lee, K. J., Tribe, D. E. & Rogers, P. L. (1979). Ethanol production of *Zymomonas mobilis* in continuous culture at high glucose concentrations. Biotechnol. Lett., 1, 421-432.

Lee, L. G. & Whitesides, G. M. (1986). Preparation of optical active 1,2-diols and α-hydroxy ketones using glycerol dehydrogenase as catalyst: Limits to enzyme-catalyzed synthesis due to noncompetitive and mixed inhibition by product. J. Org. Chem., 51, 25-36.

Leeper, F. J., Hawksley, D., Mann, S., Perez Melero, C. & Wood, M. D. (2005). Studies on thiamine diphosphate-dependent enzymes. Biochem. Soc. Trans., 33, 772-775.

Leuenberger, H. G. W., Matzinger, P. K. & Wirz, B. (1999). Synthesis and Metabolism of drugs by means of Enzyme-Catalysed Reactions. Chimia, 53, 536-540.

Lindqvist, Y., Schneider, G., Ermler, U. & Sundström, M. (1992). Three dimensional structure of transketolase, a thiamine diphosphate dependent enzyme at 2.5 Å 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.

Lingen, B., Kolter-Jung, D., Dünkelmann, P., Feldmann, R., Grötzinger, J., Pohl, M. & Müller, M. (2003). Alteration of the substrate specificity of benzoylformate decarboxylase from *Pseudomonas putida* by directed evolution. ChemBioChem, 4, 721-726.

Lingen, B., Pohl, M., Liese, A., Demir, A. S. & Müller, M. (2004). Enantioselective synthesis of hydroxyketones via benzaldehyde lyase and benzoylformate decarboxylase catalyzed C-C bond formation, 113-130. *In:* F. Jordan and M. S. Patel (ed.), Thiamine: Catalytic mechanisms in normal and disease states. Marcel Dekker Inc., New York, Basel.

Liu, M., Sergienko, E. A., Guo, F., Wang, J., Tittmann, K., Hübner, G., Furey, W. & Jordan, F. (2001). Catalytic acid-base groups in yeast pyruvate decarboxylase. 1. Site-directed mutagenesis and steady-state kinetic studies on the enzyme with the D28A, H114F, H115F, and E477Q substitutions. Biochemistry, 40, 7355-7368.

Lockington, R. A., Borlace, G. N. & Kelly, J. M. (1997). Pyruvate decarboxylase and anaerobic survival in *Aspergillus nidulans*. Gene, 191, 61-67.

Lorenz, P. & Eck, J. (2005). Metagenomics and industrial applications. Nat. Rev. Microbiol., 3, 510-516.

Lowe, S. E. & Zeikus, J. G. (1992). Purification and characterization of pyruvate decarboxylase from *Sarcina ventriculi*. J. Gen. Microbiol., 138, 803-807.

Lu, G., Dobritzsch, D., Baumann, S., Schneider, G. & König, S. (2000). The structural basis of substrate activation in yeast pyruvate decarboxylase. A crystallographic and kinetic study. Eur. J. Biochem., 267, 861-868.

Lu, G., Dobritzsch, D., König, S. & Schneider, G. (1997). Novel tetramer assembly of pyruvate decarboxylase from brewer's yeast observed in a new crystal form. FEBS Lett., 403, 249-253.

Margolin, A. L. (1993). Enzymes in the synthesis of chiral drugs. Enzyme Microb. Technol., 15, 266-280.

McSweeney, P. L. H. & Sousa, M. J. (2000). Biochemical pathways in the production of flavour compounds in cheeses during ripening. Lait, 80, 293-324.

Mesch, K. (1997). Rationales Protein-Design an der Pyruvatdecarboxylase aus *Zymomonas mobilis*. Doctoral thesis. Heinrich-Heine University, Düsseldorf.

Moller, K., Langkjaer, R. B., Nielsen, J., Piskur, J. & Olsson, L. (2004). Pyruvate decarboxylases from the petite-negative yeast *Saccharomyces kluyveri*. Mol. Genet. Genomics, 270, 558-568.

Morazzoni, P. & Bombardelli, E. (1995). *Silybum marianum* (*Carduus marianus*). Fitoterapia, 66, 3-42.

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.

Mücke, U., König, S. & Hübner, G. (1995). Purification and characterisation of pyruvate decarboxylase from pea seeds (*Pisum sativum* cv. Miko). Biol. Chem. Hoppe-Seyler, 376, 111-117.

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, 1, 95-103.

Nakamura, K., Kondo, S.-I., Kawai, Y., Hida, K., Kitano, K. & Ohno, A. (1996). Enantio- and regioselective reduction of α-diketones by baker's yeast. Tetrahedron: Asymmetry, 7, 409-412.

Neuberg, C. & Hirsch, J. (1921). Über ein Kohlenstoffketten knüpfendes Ferment (Carboligase). Biochem. Zeitschr., 115, 282.

Neuberg, C. & Karczag, L. (1911). Über zuckerfreie Hefegärungen. IV. Carboxylase, ein neues Enzym der Hefe. Biochem. Zeitschr., 36, 68.

Neuser, F., Zorn, H. & Berger, R. G. (2000). Generation of odorous acyloins by yeast pyruvate decarboxylases and their occurrence in sherry and soy sauce. J. Agric. Food Chem., 48, 6191-6195.

Neuser, F., Zorn, H., Richter, U. & Berger, R. G. (2000). Purification, characterisation and cDNA sequencing of pyruvate decarboxylase from *Zygosaccharomyces bisporus*. Biol. Chem. Hoppe Seyler, 381, 349-353.

Nguyen, C. L. (2006). Charakterisierung der verzweigtkettigen 2-Ketosäuredecaboxylase aus *Lactococcus lactis* subspecies *cremoris* B1157. Diploma thesis. Heinrich-Heine-University, Düsseldorf.

Okamoto, T., Taguchi, K., Nakamura, H., Ikenaga, H., Kuraishi, H. & Yamasato, K. (1993). *Zymobacter palmae* gen. nov., sp. nov., a new ethanol-fermentating peritrichious bacterium isolated from palm sap. Arch. Microbiol., 160, 333-337.

Okon, Y. & Labandera-Gonzalez, C. A. (1994). Agronomic applications of *Azospirillum*: an evaluation of 20 years worldwide field inoculation. Soil Biol. Biochem., 26, 1591-1601.

Oku, H. & Kaneda, T. (1988). Biosynthesis of branched-chain fatty acids in *Bacillis subtilis*. J. Biol. Chem., 263, 18386-18396.

Otten, L. G. & Quax, W. J. (2004). Directed evolution: selecting today's biocatalysts. Biomol. Eng., 22, 1-9.

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.

Panke, S., Held, M. & Wubbolts, M. (2004). Trends and innovations in industrial biocatalysis for the production of fine chemicals. Curr. Opin. Biotechnol., 15, 272-279.

Pasteur, L. (1848). Researches on the molecular asymmetry of natural organic products. English translation of French original, published by Alembic Club Reprints (1905), 14, 1-46.

Patel, R. N. (2001). Enzymatic synthesis of chiral Intermediates for drug development. Adv. Synth. Catal., 343, 527-546.

Patten, C. L. & Glick, B. R. (1996). Bacterial biosynthesis of indole-3-acetic acid. Can. J. Microbiol., 42, 207-220.

Pohl, M. (2000). Optimierung von Biokatalysatoren und Reaktionsbedingungen für den technischen Einsatz. Habilitation. Heinrich-Heine-University, Düsseldorf.

Pohl, M. (1997). Protein design on pyruvate decarboxylase (PDC) by site-directed mutagenesis. Application to mechanistical investigations, and tailoring PDC for the use in organic synthesis, pp.15-43, Advances in Biochemical Engineering/Biotechnology, vol. 58. Springer Verlag, Berlin/Heidelberg.

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.

Pohl, M., Lingen, B. & Müller, M. (2002). Thiamin-diphosphate-dependent enzymes: new aspects of asymmetric C-C bond formation. Chem. Eur. J., 8, 5288-5295.

Pohl, M., Mesch, K., Rodenbrock, A. & Kula, M. R. (1995). Stability investigations on the pyruvate decarboxylase from *Zymomonas mobilis*. Biotechnol. Appl. Biochem., 22, 95-105.

Pohl, M., Siegert, P., Mesch, K., Bruhn, H. & Grötzinger, J. (1998). Active site mutants of pyruvate decarboxylase from *Zymomonas mobilis* - a site-directed mutagenesis study of L112, I472, I476, E473, and N482. Eur. J. Biochem., 257, 538-546.

Pohl, M., Sprenger, G. A. & Müller, M. (2004). A new perspective on thiamine catalysis. Curr. Opin. Biotechnol., 15, 335-342.

Polovnikova, E. S., McLeish, M. J., Sergienko, E. A., Burgner, J. T., Anderson, N. L., Bera, A. K., Jordan, F., Kenyon, G. L. & Hasson, M. S. (2003). Structural and kinetic analysis of catalysis by a thiamin diphosphate-dependent enzyme, benzoylformate decarboxylase. Biochemistry, 42, 1820-1830.

Raj, K. C., Ingram, L. O. & Maupin-Furlow, J. A. (2001). Pyruvate decarboxylase: a key enzyme for the oxidative metabolism of lactic acid by *Acetobacter pasteurianus*. Arch. Microbiol., 176, 443-451.

Raj, K. C., Talarico, L. A., Ingram, L. O. & Maupin-Furlow, J. A. (2002). Cloning and characterization of the *Zymobacter palmae* pyruvate decarboxylase gene (pdc) and comparison to bacterial homologues. Appl. Environ. Microbiol., 68, 2869-2876.

Reetz, M. T. (2006). Directed evolution of enantioselective enzymes as catalysts for organic synthesis. Adv. Catal., 49, 1-69.

Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. (2005). Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. Angew. Chem., Int. Ed., 44, 4192-4196.

Rivoal, J., Ricard, B. & Pradet, A. (1990). Purification and partial characterization of pyruvate decarboxylase from *Oryza sativa* L. Eur. J. Biochem., 194, 791-797.

Rosche, B., Breuer, M., Hauer, B. & Rogers, P. L. (2003). Increased pyruvate efficiency in enzymatic production of (*R*)-phenylacetylcarbinol. Biotechnol. Lett., 25, 847-851.

Rosche, B., Breuer, M., Hauer, B. & Rogers, P. L. (2005). Role of pyruvate in enhancing pyruvate decarboxylase stability towards benzaldehyde. J. Biotechnol., 115, 91-99.

Rosche, B., Breuer, M., Hauer, B. & Rogers, P. L. (2003). Screening of yeasts for cell-free production of (*R*)-phenylacetylcarbinol. Biotechnol. Lett., 25, 841-845.

Rosche, B., Sandford, V., Breuer, M., Hauer, B. & Rogers, P. (2001). Biotransformation of benzaldehyde into (*R*)-phenylacetylcarbinol by filamentous fungi or their extracts. Appl. Microbiol. Biotechnol., 57, 309-315.

Rouhi, A. M. (2003). Chirality at work - Drug developers can learn much from recent successful and failed chiral switches. CENEAR, 81, 56-61.

Saehuan, C., Rojanarata, T., Wiyakrutta, S., McLeish, M. J. & Meevootisom, V. (2007). Isolation and characterization of a benzoylformate decarboxylase and a NAD+/NADP+-dependent benzaldehyde dehydrogenase involved in D-phenylglycine metabolism in *Pseudomonas stutzeri* ST-201. Biochim. Biophys. Acta, 1770, 1585-1592.

Sanchez-Gonzalez, M. & Rosazza, J. P. N. (2003). Mixed aromatic acyloin condensations with recombinant benzaldehyde lyase: Synthesis of alpha-hydroxydihydrochalcones and related alpha-hydroxy ketones. Adv. Synth. Catal., 345, 819-824.

Sauermost, R. (1994). Herder-Lexikon der Biologie. Spektrum Akad. Verl., Heidelberg/Berlin/ Oxford.

Schlee, D. & Kleber, H.-P. (1991). Wörterbücher der Biologie – Biotechnologie. Gustav Fischer Verlag, Jena.

Schmitt, H. D. & Zimmermann, F. K. (1982). Genetic analysis of the pyruvate decarboxylase reaction in yeast glycolysis. J. Bacteriol., 151, 1146-1152.

Schulze, B. & Wubbolts, M. G. (1999). Biocatalysis for industrial production of fine chemicals. Curr. Opin. Biotech., 10, 609-615.

Schütz, A., Golbik, R., Tittmann, K., Svergun, D. I., Koch, M. H., Hübner, G. & König, S. (2003). Studies on structure-function relationships of indolepyruvate decarboxylase from *Enterobacter cloacae*, a key enzyme of the indole acetic acid pathway. Eur. J. Biochem., 270, 2322-2331.

Schütz, A., Sandalova, T., Ricagno, S., Hübner, G., König, S. & Schneider, G. (2003). Crystal structure of thiamin diphosphate-dependent indolepyruvate decarboxylase from *Enterobacter cloacae*, an enzyme involved in the biosynthesis of the plant hormone indole-3-acetic acid. Eur. J. Biochem., 270, 2312-2321.

Shukla, V. B. & Kulkarni, P. R. (2000). L-Phenylacetylcarbinol (L-PAC): biosynthesis and industrial applications. World J. Microbiol. Biotechnol., 16, 499-506.

Siegert, P. (2000). Vergleichende Charakterisierung der Decarboxylase- und Carboligasereaktion der Benzoylformiatdecarboxylase aus *Pseudomonas putida* und der Pyruvatdecarboxylase aus *Zymomonas mobilis* mittels gerichteter Mutagenese. Doctoral thesis. Heinrich-Heine University, Düsseldorf.

Siegert, P., McLeish, M. J., Baumann, M., Iding, H., Kneen, M. M., Kenyon, G. L. & Pohl, M. (2005). Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. Protein Eng., Des. Sel., 18, 345-357.

Sijbesma, F. & Schepens, H. (2003). White Biotechnology: Gateway to a More Sustainable Future. EuropaBio. Lyon, 10.04.2003.

Skory, C. D. (2003). Induction of *Rhizopus oryzae* pyruvate decarboxylase genes. Curr. Microbiol., 47, 59-64.

Smit, B. A., Engels, W. J. M., Bruinsma, J., van Hylckama Vlieg, J. E. T., Wouters, J. T. M. & Smit, G. (2004). Development of a high throughput screening method to test flavour-forming capabilities of anaerobic micro-organisms. J. Appl. Microbiol., 97, 306-313.

Smit, B. A., Vlieg, J., Engels, W. J. M., Meijer, L., Wouters, J. T. M. & Smit, G. (2005a). Identification, cloning, and characterization of a *Lactococcus lactis* branched-chain alpha-keto acid decarboxylase involved in flavor formation. Appl. Environ. Microbiol., 71, 303-311.

Smit, G., Smit, B. A. & Engels, W. J. (2005b). Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. FEMS Microbiol. Rev., 29, 591-610.

Soetaert, W. & Vandamme, E. (2006). The impact of industrial biotechnology. Biotechnol. J., 1, 756-769.

Somers, E., Ptacek, D., Gysegom, P., Srinivasan, M. & Vanderleyden, J. (2005). *Azospirillum brasilense* produces the auxin-like phenylacetic acid by using the key enzyme for indole-3-acetic acid biosynthesis. Appl. Environ. Microbiol., 71, 1803-1810.

Stanier, R. Y. & Ornston, L. N. (1973). The beta-ketoadipate pathway. Adv. Microbiol. Physiol., 9, 89-151.

Stermitz, F. S., Lorenz, P., Tawara, J. N., Zenewicz, L. A. & Lewis, K. (2000). Synergy in a medicinal plant: antimicrobial action of berberine potentiated by 5'-methoxyhydnocarpin, a multidrug pump inhibitor. Proc. Natl. Acad. Sci. USA, 97, 1433-1437.

Stillger, T. (2004). Enantioselektive C-C Knüpfung mit Enzymen - Charakterisierung und reaktionstechnische Bearbeitung der Benzaldehydlyase aus *Pseudomonas fluorescens* Biovar I. Doctoral thesis. Rheinische Friedrich-Wilhelms-University Bonn.

Stillger, T., Pohl, M., Wandrey, C. & Liese, A. (2006). Reaction engineering of benzaldehyde lyase catalyzing enantioselective C-C bond formation. Org. Process Res. Dev., 10, 1172-1177.

Straathof, A. J. J., Panke, S. & Schmid, A. (2002). The production of fine chemicals by biotransformations. Curr. Opin. Biotechnol., 13, 548-556.

Strauss, U. T., Felfer, U. & Faber, K. (1999). Biocatalytic transformation of racemates into chiral building blocks in 100% chemical yield and 100% enantiomeric excess. Tetrahedron: Asymmetry, 10, 107-117.

Streit, W. R., Daniel, R. & Jaeger, K.-E. (2004). Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. Curr. Opin. Biotechnol., 15, 285-290.

Sukumaran, J. & Hanefeld, U. (2005). Enantioselective C-C bond synthesis catalysed by enzymes. Chem. Soc. Rev., 34, 530-542.

Sullivan, R. G., Dale, J. A. & Mosher, H. S. (1973). Correlation of configuration and 19 F chemical shifts of α-methoxy-α-trifluoromethylphenylacetate derivatives. J. Org. Chem., 98, 2143-2147.

Sumner, J. B. (1926). The isolation and crystallization of the enzyme urease. J. Biol. Chem., 69, 435-441.

SusChem Solutions (2007). European technology platform for sustainable chemistry. http://www.suschem.org/.

Talarico, L. A., Ingram, L. O. & Maupin-Furlow, J. A. (2001). Production of the gram-positive *Sarcina ventriculi* pyruvate decarboxylase in *Escherichia coli*. Microbiology, 147, 2425-2435.

Thayer, A. M. (2006). Enzymes at work. Chem. Eng. News, 84, 15-25.

Tittmann, K., Mesch, K., Pohl, M. & Hübner, G. (1998). Activation of thiamine diphosphate in pyruvate decarboxylase from *Zymomonas mobilis*. FEBS Lett., 441, 404-406.

Toscano, M. D., Woycechowsky, K. J. & Hilvert, D. (2007). Minimale Umgestaltung aktiver Enzymtaschen - wie man alten Enzymen neue Kunststücke beibringt. Angew. Chem., 119, 3274-3300.

Toukoniitty, E., Mäki-Arvela, P., Kuzma, M., Villela, A., Neyestanaki, A. K., Salmi, T., Sjöholm, R., Leino, R., Laine, E. & Murzin, D. Y. (2001). Enantioselective hydrogenation of 1-phenyl-1,2-propanedione. J. Catal., 204, 281-291.

Trost, B. M. (2004). Asymmetric catalysis: an enabling science. Proc. Natl. Acad. Sci. USA, 101, 5348-5355.

Tsou, A. Y., Ransom, S. C., Gerlt, J. A., Buechter, D. D., Babbitt, P. C. & Kenyon, G. L. (1990). Mandelate pathway of *Pseudomonas putida*: sequence relationships involving mandelate racemase, (*S*)-mandelate dehydrogenase, and benzoylformate decarboxylase and expression of benzoylformate decarboxylase in *Escherichia coli*. Biochemistry, 29, 9856-9862.

Turner, M. K. (1998). Perspectives in Biotransformations, 5-23. *In:* D. R. Kelly (ed.), Biotechnology: Biotransformations I vol. 8a. Wiley VCH, Weinheim.

Ukai, T., Tanaka, S. & Dokawa, S. (1943). New catalysts for acyloin condensation. J. Pharm. Soc. Jpn., 63, 296-300.

Versées, W., Spaepen, S., Vanderleyden, J. & Steyaert, J. (2007). The crystal structure of phenylpyruvate decarboxylase from *Azospirillum brasilense* at 1.5 Å resolution. Implications for its catalytic and regulatory mechanism. FEBS J., 274, 2363-2375.

Vuralhan, Z., Morais, M. A., Tai, S. L., Piper, M. D. & Pronk, J. T. (2003). Identification and characterization of phenylpyruvate decarboxylase genes in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol., 69, 4534-4541.

Wackett, L. P. (2004). Novel biocatalysis by database mining. Curr. Opin. Biotechnol., 15, 280-284.

Ward, O. P. & Singh, A. (2000). Enzymatic asymmetric synthesis by decarboxylases. Curr. Opin. Biotechnol., 11, 520-526.

Watson, J. D. & Crick, F. H. C. (1953). Molecular structure of nucleic acids: a structure for DNA. Nature, 171, 737-738.

Wendorff, M. (2006). Neue Benzoylformiatdecarboxylasen für die Biokatalyse. Doctoral thesis. Heinrich-Heine University, Düsseldorf.

Wilcocks, R. & Ward, O. P. (1992). Factors effecting 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. & Liebig, J. (1832). Untersuchungen über das Radikal der Benzoesäure. Ann. Pharm., 3, 249-282.

Wu, Y. G., Chang, A. K., Nixon, P. F., Li, W. & Duggleby, R. G. (2000). Mutagenesis at Asp27 of pyruvate decarboxylase from *Zymomonas mobilis*. Effect on its ability to form acetoin and acetolactate. Eur. J. Biochem., 267, 6493-6500.

Yep, A., Kenyon, G. L. & McLeish, M. J. (2006a). Determinants of substrate specificity in KdcA, a thiamin diphosphate-dependent decarboxylase. Bioorg. Chem., 34, 325-336.

Yep, A., Kneen, M. M., Kenyon, G. L. & McLeish, M. J. (2006b). Structural basis of substrate specificity in thiamin diphosphate dependent decarboxylases. FASEB J., 20:A476

Die vorliegende Dissertation habe ich vollständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Köln, den 28. November 2007