Stefanie Hanreich, BSc.

Identification and selection of active α-ketoglutarate dependent dioxygenases in engineered *Escherichia coli*

MASTER'S THESIS
to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Biotechnology

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dr.rer.nat. Robert Kourist

Institute of Molecular Biotechnology

Graz, September 2019

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master's thesis.

Date

Signature

Acknowledgements

I would like to express my deepest gratitude to my supervisor Univ.-Prof. Dr.rer.nat Robert Kourist for the opportunity of completing my master's thesis in his research group and providing me with this fascinating project. Moreover, I would like to express my appreciation for the invaluable advices and scientific input he provided me during this research project.

I would also like to express my heartfelt gratitude to Anna Schweiger, MSc. and Aline Telzerow, MSc. for their professional and patient guidance and warm encouragement they provided me throughout this thesis. Me and my master's thesis have greatly benefited from the invaluable suggestions and generous support. Moreover, I especially appreciated the time they spent proofreading this thesis.

My sincerest thanks are also expressed to everyone in the research group for their support and helpful advices. Special thanks to Dipl.-Biochem. Dr.rer.nat. Sandy Schmidt and Mgr. dr Bc. Ivana Drienovska for the scientific discussions and input. I am also especially thankful to Maria Schabhüttl and Daniel Mertschnigg for their assistance in the everyday lab routine.

I also want to acknowledge Prof. Dr. Dirk Tischler of the Ruhr-Universität Bochum for the supply of plasmids encoding different dioxygenases. Moreover, a special thanks to Prof. Dr. Bruno Bühler of the Helmholtz centre for environmental research providing the selection strain *E. coli* 3Δ*sucA*.

I cannot express how thankful I am for my family, especially for my mother supporting me financially and encouraging me during my study and my whole life. My deepest and warmest gratitude to our "master student self-supporting group" Andrea Nigl, Kristin Bauer, Magdalena Ingeborg Wessely and Michael Runda supporting and encouraging me during all the ups and downs of my research. I am extremely thankful for all the valuable advices and marvelous time we had.

Last but not the least important, I would like to thank Gabrijela Mijić and Joanna Laßnig, who have always been there for me and cheered me up during frustrating times.

Abstract

Hydroxylated amino acids are widely ubiquitous in nature and serve as valuable building blocks for various chemical products. Fe(II)/α-ketoglutarate dependent dioxygenases belong to the group of non-heme iron (II) dependent dioxygenases and catalyze a wide spectrum of reactions in a chemo-, regio- and stereoselective manner, whereby the hydroxylation of L-amino acids is the most common. In recent years, many of those hydroxylases were discovered and characterized occurring in several organisms ranging from prokaryotes to fungi. In this study, a dioxygenase toolbox was generated investigating the soluble expression and the substrate spectrum of overall eleven enzymes with the aim to identify enzymes for the hydroxylation of sterically hindered amino acids. Among the examined dioxygenases, ten were successfully expressed in *E. coli* BL21(DE3) varying in soluble expression levels. Enzymatic reactions performed with the L-leucine dioxygenase from *Anabaena variabilis* (AvLDO) demonstrated full conversion of L-leucine into (2*S*,4*S*)-5-hydroxyleucine within 24 hours. Throughout this thesis, the implementation of an engineered *E. coli* strain was examined for the establishment of a selection tool for active Fe(II)/α-ketoglutarate dependent dioxygenases. A deletion of the genes *sucA* (encoding α-KG dehydrogenase E1 subunit) and *aceA* (encoding isocitrate lyase) leads to blocked succinate production and the mutant strain should be unable to grow in minimal medium. Growth of the Δ*sucA*Δ*aceA*Δ*putA* (3Δ*sucA*) mutant could be restored in minimal medium with TCA re-cyclization via dioxygenase catalyzed L-amino acid hydroxylation, thus coupling hydroxylation of a desired substrate with cell growth. Nevertheless, growth of the mutant strain was also restored without the hydroxylating activity of a dioxygenase in some cases. Furthermore, a comparison of the growth behavior of *E. coli* BL21(DE3) and the mutant strain was studied. Due to the metabolic burden, the mutant strain showed a decelerated growth compared to *E. coli* BL21(DE3).

Zusammenfassung

Hydroxyaminosäuren sind in der Natur allgegenwärtig auffindbar und werden für verschiedenste Chemieerzeugnisse als Synthesebausteine verwendet. Fe(II)/α-Ketoglutarat-abhängige Dioxygenasen gehören zur Gruppe der nicht-Häm Fe(II)-abhängigen Dioxygenasen und katalysieren ein breites Spektrum an Reaktionen in einer chemo-, regio- und stereoselektiven Art und Weise, wobei die Hydroxylierung von L-Aminosäuren präferiert wird. Ihr Vorkommen reicht von Prokaryoten bis hin zu Pilzen und in den letzten Jahren wurden viele dieser Hydroxylasen entdeckt und charakterisiert. In dieser Arbeit wurde eine Dioxygenasen-Toolbox etabliert, wobei das Expressionslevel von löslichem Protein und das Substratspektrum von insgesamt elf Dioxygenasen untersucht wurde mit dem Ziel Enzyme zu identifizieren, welche in der Lage sind sterisch gehinderte Aminosäuren zu hydroxylieren. Zehn Enzyme wurden erfolgreich in *E. coli* BL21(DE3) exprimiert, jedoch mit unterschiedlichen Erträgen an löslichem Protein. Enzymreaktionen, welche mit gereinigtem Protein der L-Leucin Dioxygenase aus *Anabaena variabilis* (AvLDO) durchgeführt wurden, zeigten innerhalb von 24 Stunden eine erfolgreiche Umwandlung von L-Leucin zu (2*S*,4*S*)-5-Hydroxyleucin. Weiters wurde in dieser Arbeit die Anwendung eines modifizierten *E. coli* Stammes untersucht, der als Selektionstool für aktive α-Ketoglutarat-abhängige Dioxygenasen dienen soll. Die Synthese von Succinat wurde durch einen Knockout an den Genen *sucA* und *aceA* blockiert, wodurch das Wachstum des Mutanten unterbunden wird. Das Wachstum von *E. coli* 3Δ*sucA* konnte mittels der Aktivität der Dioxygenasen wiederhergestellt werden, wobei die Hydroxylierung eines Substrates mit der Wiederherstellung des Zellwachstumes verbunden ist. Wachstumsversuche zeigten jedoch, dass das Wachstum des Mutanten in manchen Fällen auch ohne die Aktivität einer Dioxygenase wiederhergestellt werden kann. Darüber hinaus, wurde auch das Wachstumsverhalten von *E. coli* BL21(DE3) mit dem des Mutanten verglichen und Untersuchungen ergaben, dass der Mutant durch die metabolische Belastung die ihm widerfährt ein verzögertes Wachstum aufweist.

List of abbreviations

(2*S*,3*R*,4*S*)-HIL	(2*S*,3*R*,4*S*)-4-hydroxyisoleucine
ACCO	1-amino-1-cyclopropanecarboxylic acid oxidase
AMKP	(2S,3R)-2-amino-3-methyl-4-ketopentanoate
amp	Ampicillin
AnPip4H	Cyclic amino acid hydroxylase from *A. nidulans* FGSC A4
ATP	Adenosine triphosphate
AvLDO	Leucine dioxygenase from *A. variabilis*
BCA	Bicinchoninic acid
bp	Base pairs
BtIDO	Isoleucine dioxygenase from *B. thuringiensis*
CAS	Clavaminate synthase
CFE	Cell free extract
cv	Column volume
DACS	Deacetylcephalosporin C synthase
DAOCS	Deacetoxycephalosporin C synthase
DAP	Diaminopimelate
ddH_2O	Double-distilled water
dH_2O	Distilled water
dNTPs	Nucleoside triphosphate
DO	Dioxygenase
DSBH	Double-stranded β-helix fold
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EV	Empty vector
FoPip4H	Cyclic aa hydroxylase from *F. oxysporum*
Fw primer	Forward primer
g	Relative centrifugal compared to gravity
GriE	Leucine dioxygenase from Streptomyces strain DSM 40835
IPNS	Isopenicillin N synthase
IPTG	Isopropanyl β-D-1 thiogalactopyranoside
kDa	Kilo Dalton
KDO1	Lysine dioxygenase
km	Kanamycin
KPi	Potassium phosphate buffer
LB	*Lysogeny* broth
LdoA	L-Leucine 5-hydroxylase from *Nostoc punctiforme* PCC 73102
MfIDO	Leucine dioxygenase from *M. flagellatus*
MMGDO	Putative dioxygenase from marine metagenome
MOPS	(3-(*N*-morpholino)propanesulfonic acid)
n.d	Not determined
nc	Negative control
NMR	Nuclear magnetic resonance
O/N	Overnight
OD_{600}	Optical density at 600 nm
ONC	Over-night culture
P4H	Proline-4-hydroxylase
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
R_f-value	Retention value

rpm	Rotation per minute
Rv primer	Reverse primer
SadA	N-succinyl amino acid hydroxylase from *B. ambifaria* AMMD
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SOC	Super optimal broth with catabolite repression
TAE	TRIS-acetate-EDTA
TB	Terrific broth
TCA	Tricarboxylic acid cycle
TLC	Thin layer chromatography
TmDO	Putative dioxygenase from *T. mobilis*
Tris	Tris(hydroxymethyl)aminomethane
U	Enzymatic units
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume
α-KG	α-ketoglutarate

Table of contents

1 Introduction

1.1 Oxidoreductases and their industrial application as biocatalysts

The application of biocatalysts at the industrial level and in research has nowadays gained increasing interest due to the great demand for green technologies. Biocatalysts serve as a great substitution to traditional chemical procedures as enzymes are energy-saving, operate under mild conditions reducing problems of unwanted side reactions and minimizing downstream processing thus decreasing waste production. Moreover, enzymes possess the ability to accept a wide spectrum of substrates and perform reactions in a chemo-, regio- and stereoselective manner. (Choi et al., 2015; Faber, 2018; Sheldon et al., 2007). One example of such an enzyme class are oxidoreductases (EC 1.-) which find great demand in organic synthesis such as chemical and pharmaceutical industries but also in food, environmental and medicinal fields due to their sustainable and ecofriendly properties (Gamenara et al., 2012; Xu, 2005).

Oxidoreductases are characteristic for catalyzing oxidation and reduction reactions via electron transfer between a reductant to an oxidant. Besides of hydroxylation reactions, Baeyer-Villiger oxidations or the oxidation of alcohols, acids and aldehydes, they are also capable of catalyzing the reduction of aldehydes or ketones, C=C and C=N double bonds (Figure 1). Depending on the required coenzyme and their structural configuration, oxidoreductases can be classified in oxygenases, oxidases, peroxidases and dehydrogenases (Gamenara et al., 2012; Wohlgemuth, 2010; Xu, 2005). For instance, reactions catalyzed by oxygenases are highly interesting as they activate C-H bonds of small molecules by incorporation of oxygen resulting in the formation of new functional groups (Dong et al., 2018; Gamenara et al., 2012; Hüttel, 2013). One of the most popular C-H activating enzymes are the cytochrome P450 monooxygenases followed by the non-heme Fe(II) dependent dioxygenases (Hüttel, 2013; Schofield and Zhang, 1999).

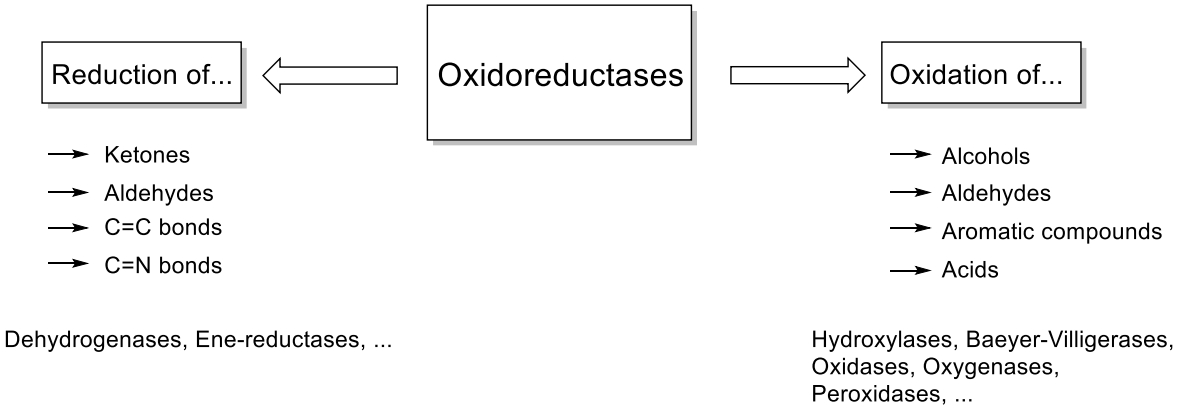

Figure 1: Examples of reduction and oxidation reactions catalyzed by various oxidoreductases.

Oxygenases are divided into two subclasses including monooxygenases (EC 1.13.-) and dioxygenases (EC 1.14.-). Reactions, which are catalyzed by monooxygenases only incorporate one oxygen atom from dioxygen into the product and the second one is reduced by the cofactors NADH or NADPH generating water. Dioxygenases on the other hand incorporate both oxygen atoms from molecular oxygen (Bugg, 2003; Dong et al., 2018; Gamenara et al., 2012). The utilization of oxygenases in chemical synthesis has also some drawbacks and is limited by a low stability of the enzymes, the reliance of cofactors, a narrow substrate scope and in rare cases a restricted regio- and stereoselectivity (Gamenara et al., 2012; Hüttel, 2013).

Dioxygenases are categorized into two main classes including a) non-heme iron (II) dependent dioxygenases such as α-ketoglutarate dependent dioxygenases, extradiol- and Rieske-type dioxygenases and b) iron (III)-dependent dioxygenases comprising intradiol dioxygenases and lipoxygenases. The first class of dioxygenases activates O_2 directly, whereas the second class is activating the substrate (Gamenara et al., 2012). The α-ketoglutarate-dependent dioxygenases occupy the biggest group of non-heme iron(II) oxygenases and are ubiquitous in a variety of organisms such as eukaryotes and prokaryotes (Hüttel, 2013; Schofield and Zhang, 1999).

1.2 Fe(II)/α-ketoglutarate dependent dioxygenases

Fe(II)/α-ketoglutarate dependent dioxygenases, belonging to the non-heme iron(II)-dependent oxygenases, are capable of catalyzing oxyfunctionalization reactions by incorporating one oxygen atom from molecular oxygen into the hydroxylated amino acid and the second one into succinate. These enzymes need Fe(II) as cofactor and α-ketoglutarate as cosubstrate when catalyzing important reactions such as hydroxylations, sulfoxidations, desaturations, epoxidations and oxidative ring closure and expansions reactions, whereas hydroxylation reactions are the most common. (Bugg, 2003; Herr and Hausinger, 2018; Hüttel, 2013; Islam et al., 2018). However, enzymes with structural similarities to Fe(II)/α-ketoglutarate dependent dioxygenases such as isopenicillin N synthase (IPNS) and 1-amino-1-cyclopropanecarboxylic acid oxidase (ACCO) use ascorbate as cosubstrate instead of α-ketoglutarate (Schofield and Zhang, 1999).

Due to their highly specific oxyfunctionalization, Fe(II)/α-ketoglutarate dependent dioxygenases play a key factor in a wide range of biological activities including the modification of proteins such as hydroxylating side chains of proline, aspartate, lysine and asparagine, metabolism of fatty acids, repair of nucleic acids like DNA and RNA, the biosynthesis of valuable antibiotics especially β-lactams and secondary metabolites such as flavonoids, alkaloids and gibberellins (Bollinger et al., 2005; Hüttel, 2013; Islam et al., 2018; Martinez and Hausinger, 2015). For instance, prolyl 4-hydroxylase, the first identified α-ketoglutarate dependent dioxygenase by Udenfriend (Hutton et al., 1967), is indispensable for collagen biosynthesis as they catalyze the hydroxylation of proline side chains in procollagen. The enzyme is a α2β2 tetramer, whereby the α-subunit is responsible for the hydroxylating activity of the protein and the β-subunit supports the solubility. About 30 % of the protein content of mammals consist of type one collagen, whereas 10 % comprise of *trans*-4-hydroxyproline (Bugg, 2003; Hausinger, 2004; Hüttel, 2013; Schofield and Zhang, 1999). Nucleic acids such as DNA and RNA are often alkylated by S-adenosylmethionine and other methylating chemicals. For instance, SN1 agents which comprise of methylnitrosourea and N-methyl-N'-nitro-N-nitrosoguanidine alkylating guanine, adenine, pyrimidines and the phosphate backbone, whereas compounds of SN2 include methylmethanesulfonate and dimethylsulfate interact with adenine and cytosine. These reactions lead to 1-methyladenine and 3-methylcytosine DNA lesions blocking DNA replication. In order

to eliminate these toxic and mutagenic lesions, DNA repairing enzymes such as AlkB, belonging to Fe(II)/α-ketoglutarate dependent dioxygenases, find utilization as the enzyme oxidizes the toxic methyl groups generating formaldehyde. Besides, AlkB seems to possess a resistance against SN2 alkylating compounds when expressed in human cells (Duncan et al., 2002; Hausinger, 2004).

Another important role of a Fe(II)/α-ketoglutarate dependent dioxygenase that should be showcased is the biosynthesis of β-lactams. For instance, clavaminate synthase (CAS) and the two isoenzymes CS1 and CS2 from *Streptomyces clavuligerus* catalyze a three-step oxidation reaction in the biosynthesis of clavulanic acid. In between those reactions proclavaminate amidinohydrolase performs a hydrolyzation generating proclavaminic acid followed by cyclization and desaturation by CAS (Hausinger, 2004; Townsend, 2002). Furthermore, the Fe(II)/α-ketoglutarate dependent dioxygenases deacetylcephalosporin C synthase (DACS) and deacetoxycephalosporin C synthase (DAOCS) are important key factors during cephamycin biosynthesis catalyzing ring expansions and hydroxylation reactions beginning from penicillin N. DACS converts penicillin N into deacetoxycephalosporin C by ring expansion on the thiazolidine ring. Interestingly, the activity of the enzyme is heavily provoked by ascorbic acid instead of ATP. Afterwards, the hydroxylation of deacetoxycephalosporin C catalyzed by DAOCS leads to the formation of deacetylcephalosporin C. Both reactions are highly stimulated by ascorbic acid, α-ketoglutarate, Fe(II) and dithiothreitol (DTT) (Hausinger, 2004; Islam et al., 2018; Martín et al., 1999).

1.3 Catalytic mechanism and three-dimensional structure of Fe(II)/α-ketoglutarate dependent dioxygenases

A characteristic feature of these dioxygenases is a double-stranded β-helix fold (DSBH) known as “jellyroll” which is composed of eight β-strands in a shape of two, four-stranded sheets. Additional α-helices, loops and β-strands strengthen the stability of the protein core. Within this DSBH, Fe(II) is bound by two histidine side chains and a carboxylate from a glutamate or an aspartate side chain (His1-X-Asp/Glu-Xn-His2 motif) with three additional water molecules as coordination sites (Figure 2 (1)). By binding of α-ketoglutarate to Fe(II) via its 1-carboxylate and 2-oxo groups and further stabilization of a C-5 carboxylate group, two water molecules are displaced (Figure 2 (2)) (Gamenara et al., 2012; Hausinger, 2004; Martinez and Hausinger, 2015; Qin et al., 2013). Next, the ‘prime’ substrate binds in the active

site of the enzyme resulting in a removal of the third water molecule and generation of a binding site for an O_2 molecule (Figure 2 (3)). Furthermore, this leads to the formation of an Fe(III)-superoxo intermediate (Figure 2 (4)). An oxygen atom of the Fe(III)-superoxo intermediate performs a nucleophilic attack on the C2 of α-ketoglutarate resulting in a peroxohemiketal bicyclic intermediate (Figure 2 (5)). Followed by that, α-ketoglutarate is decarboxylated generating CO_2 and a Fe (IV)-oxo species with succinate as ligand (Figure 2 (6)). Finally, substrate hydroxylation occurs in a two-step process in which a hydrogen atom is abstracted from the 'prime' substrate pursued by an oxygen rebound and release of the hydroxylated substrate, succinate and CO_2 concluding the catalytic cycle (Figure 2 (7, 8)) (Gamenara et al., 2012; Hüttel, 2013; Martinez and Hausinger, 2015; Wu et al., 2016).

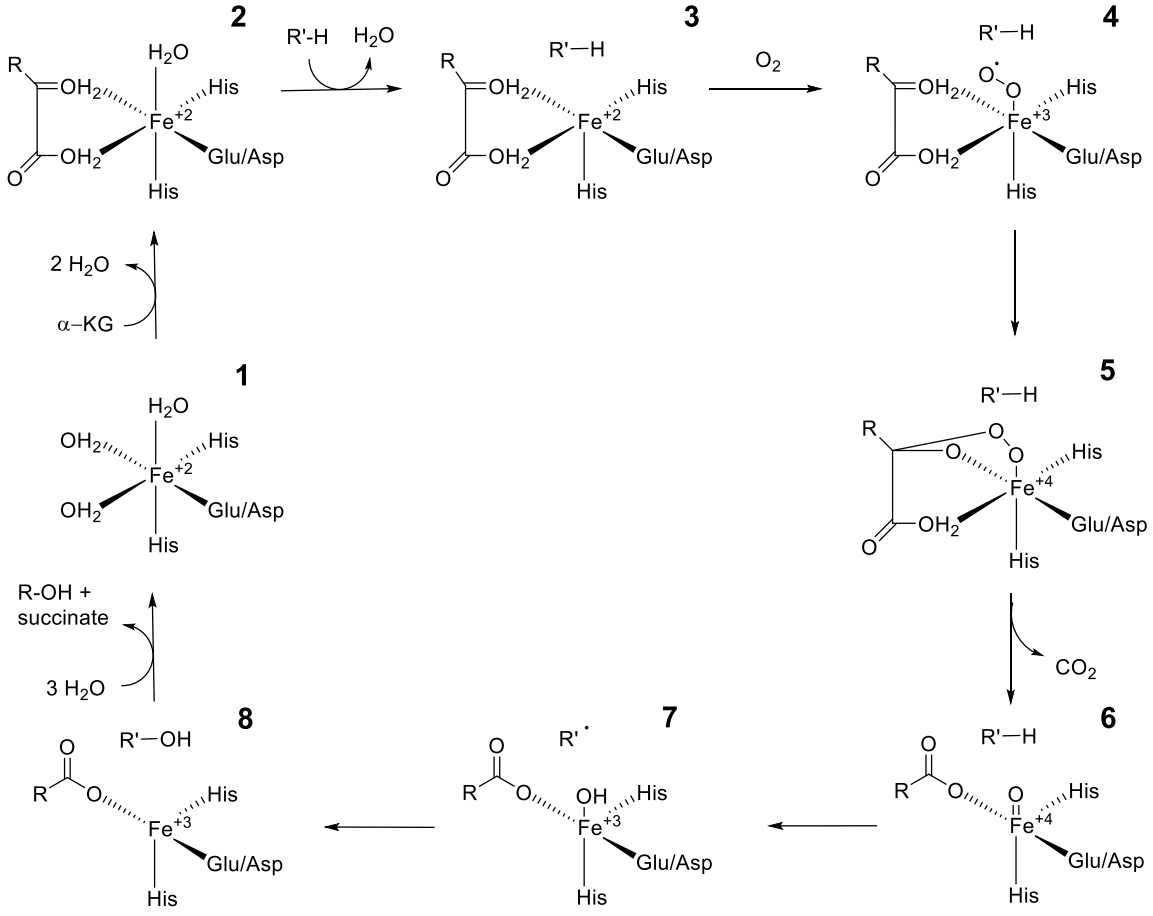

Figure 2: Hydroxylation mechanism of α-ketoglutarate dependent dioxygenases. Adapted from (Martinez and Hausinger, 2015).

Besides catalyzing hydroxylation reactions, Fe(II)/α-ketoglutarate dependent dioxygenases are also capable of performing other major oxidative reactions including halogenation, ring formation and desaturation reactions (Martinez and Hausinger, 2015). For instance, in case of halogenation reactions SyrB2 from

Pseudomonas syringae B301D, the first identified Fe(II)/α-ketoglutarate dependent halogenase, is able to chlorinate L-tyrosine which is bound to SyrB1. Most of these enzymes seem to halogenate substrates which are bound to a phosphopantetheine group in a carrier protein's thiolation domain. Moreover, their reaction mechanism is analogously to those of Fe(II)/α-ketoglutarate dependent hydroxylases (Martinez and Hausinger, 2015; Vaillancourt et al., 2005). Initial steps of the catalytic mechanism are identical but instead of the typical His1-X-Asp/Glu-Xn-His2 motif, here the Fe(II) is bound by two histidine side chains and one chloride ion. The generated ferryl intermediate (Figure 3) is now termed haloferryl species and is abstracting a hydrogen atom from the prime substrate resulting in a *cis*-halohydroxo-ferric species and substrate radical. The chloride atom is then transferred to the substrate radical leading to the formation of a halogenated product (Martinez and Hausinger, 2015; Pratter et al., 2014).

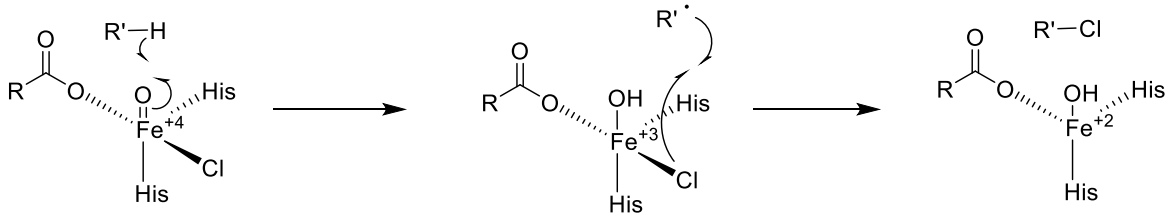

Figure 3: Halogenation reaction catalyzed by α-ketoglutarate dependent halogenases. Adapted from (Martinez and Hausinger, 2015).

Another major role among this enzyme class is oxidative ring formation and desaturation during the biosynthesis of β-lactams, showcased by CAS, DACS and DAOCS (Hausinger, 2004; Islam et al., 2018; Martinez and Hausinger, 2015). For instance, in case of CAS the catalytic mechanism is subdivided into three different oxidative pathways after the formation of the ferryl intermediate, depending on which hydrogen atom is abstracted from procalvaminic acid. An abstraction can occur on the C4 carbon at the proS position, the C3 hydroxide or also in a coupled manner (Martinez and Hausinger, 2015). Further Fe(II)/α-ketoglutarate dependent dioxygenase reactions to showcase are non-productive reactions which occur with oxygen. Oxygen is able to bind at the catalytic site of substrate free enzymes leading to inactive Fe(III)-containing enzymes. α-ketoglutarate bound enzymes show "uncoupled turnover" in which the cosubstrate collapses to succinate and CO_2 without the conversion of the prime substrate. This phenomenon is known in prolyl hydroxylases in which the metal is oxidized through this uncoupled reaction (Hausinger, 2004).

1.4 Dioxygenase toolbox

As previously mentioned, the most common reactions α-ketoglutarate dependent dioxygenases catalyze are hydroxylations of non-activated methylene groups followed by the oxidation of sulfur containing amino acids. The investigated hydroxylases have a wide substrate spectrum ranging from cyclic amino acids including pipecolic acid and proline to non-cyclic amino acids such as leucine, isoleucine or methionine. Moreover, the substrate tolerance towards α-methyl-L-leucine was investigated as the hydroxylation of α-methyl amino acids gives rise to several promising chemical building blocks (Figure 4). Six of these dioxygenases were received from the plasmid collection of the working group of Prof. Dr. Dirk Tischler (Bochum) and the remaining four are representative enzymes from literature (Table 2). During this thesis, particular high attention was dedicated to the following outlined dioxygenases the L-leucine dioxygenase from *Anabaena variabilis* (AvLDO), L-isoleucine dioxygenase from *Bacillus thuringiensis* (BtIDO) and L-leucine dioxygenase from *Streptomyces* strain DSM 40835 (GriE).

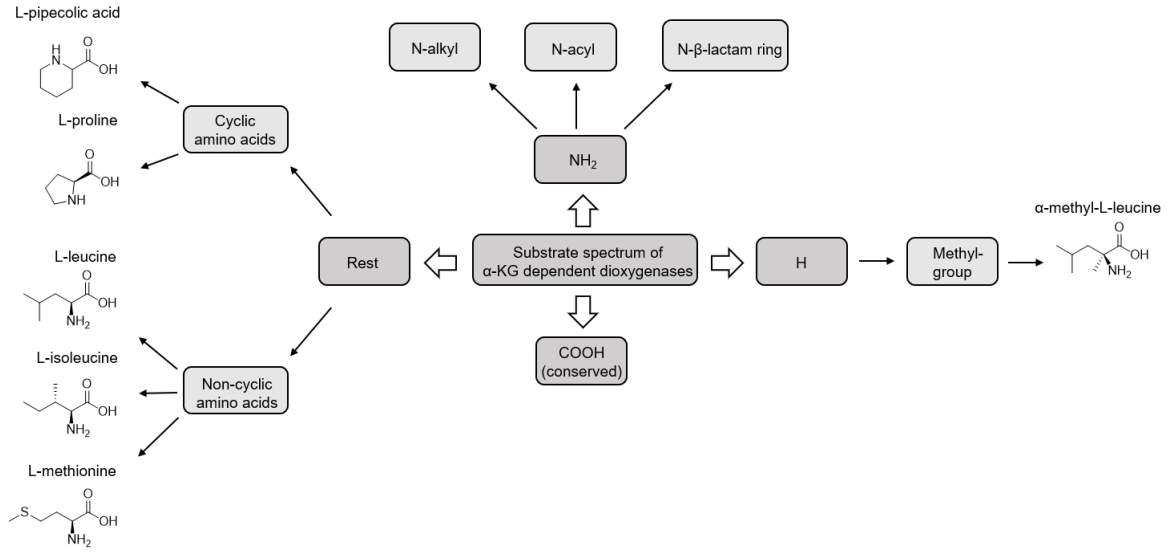

Figure 4: Substrate spectrum of α-ketoglutarate dependent dioxygenases ranging from various cyclic and non-cyclic amino acids to β-lactams.

1.5 L-leucine dioxygenase from *Anabaena variabilis* - AvLDO

The L-leucine dioxygenase from *Anabaena variabilis*, recently discovered and characterized by Correia Cordeiro *et al.* (Correia Cordeiro et al., 2018) belongs to one of three L-leucine dioxygenases from cyanobacteria that were investigated at

the time. By investigating the substrate scope of AvLDO, it was demonstrated that the enzyme hydroxylates L-leucine and L-norleucine in C5 position generating (2*S*,4*S*)-5-hydroxyleucine and 5-hydroxynorleucine. AvLDO is also capable of catalyzing the sulfoxidation of L-methionine and L-ethionine (Figure 5). However, while sulfoxidation of L-methionine catalyzed by AvLDO generates (*R*)-sulfoxides, BtIDO offers the opposite stereopreference and provides (*S*)-sulfoxides (Correia Cordeiro et al., 2018; Peters and Buller, 2019). Moreover, an important fact is that AvLDO is incapable of converting L-isoleucine, valine and norvaline although structure similarities to the accepted substrates are given. Correia Cordeiro *et al* (Correia Cordeiro et al., 2018) observed that AvLDO favors acidic conditions with a limited pH profile at pH 3.5–5.3. At a higher pH the protein undergoes denaturation. For enzymatic reactions which were performed with L-leucine, the highest activity was observed at 25 °C with a specific activity of 0.22 ± 0.024 U/mg. At lower temperatures the activity remained at 90 % (Correia Cordeiro et al., 2018). AvLDO possesses 82 % homology to the L-leucine 5-hydroxylase from *Nostoc punctiforme* (LdoA). Therefore, LdoA and AvLDO have many properties in common including the substrate scope and the desire for acidic conditions. Especially, the fact that they favor more acidic conditions should be highlighted, as most other Fe(II)/α-ketoglutarate dependent enzymes act under a neutral pH optimum (Correia Cordeiro et al., 2018; Hibi et al., 2013).

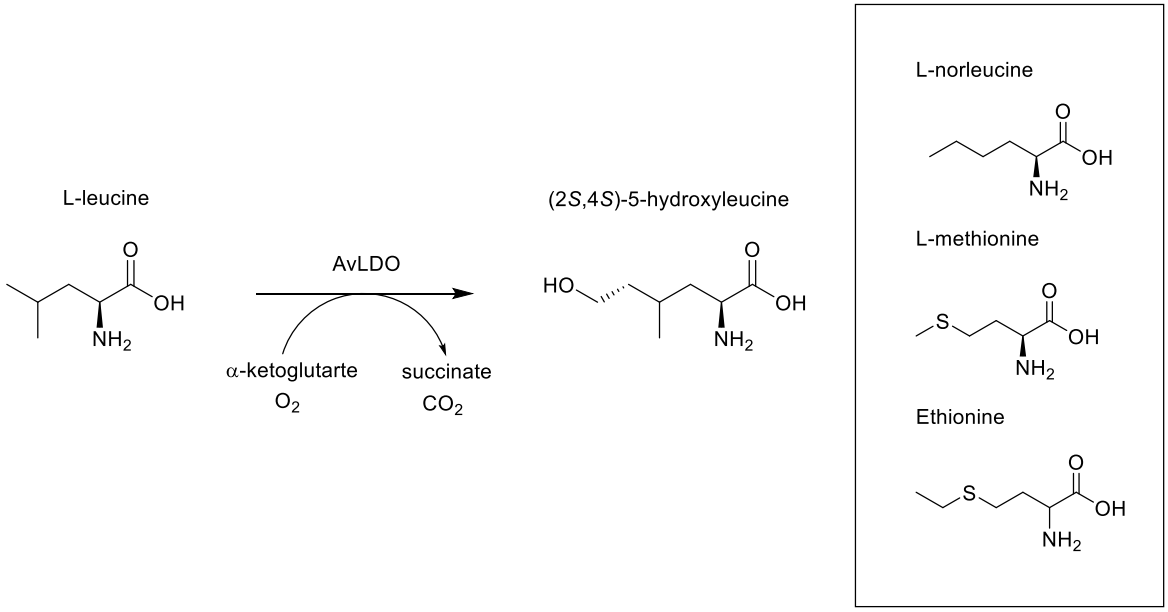

Figure 5: C5-hydroxylation of L-leucine catalyzed by AvLDO. Amino acids in the box are accepted by AvLDO as substrates. Adapted from (Correia Cordeiro et al., 2018).

1.6 L-isoleucine dioxygenase from *Bacillus thuringiensis* - BtIDO

The L-isoleucine dioxygenase from *Bacillus thuringiensis* initially characterized by Kodera *et al.* (Kodera et al., 2009), catalyzes three distinct oxidation reactions. First is the hydroxylation of several hydrophobic non-aromatic L-amino acids such as L-isoleucine leading to the formation of (2*S*,3*R*,4*S*)-4-hydroxyisoleucine (HIL) (Figure 6). (2*S*,3*R*,4*S*)-HIL, originally found in fenugreek seeds (*Trigonella foenum-graecum*), gains increasing interest due to its anti-diabetic effect and anti-obesity activity as its preventing hypoglycemia and insulinemia (Hibi et al., 2011; Kodera et al., 2009; Peters and Buller, 2019). Recombinant BtIDO was initially generated in *E. coli* by Hibi *et al.* (Hibi et al., 2011) in 2011 where they identified the substrate scope and substrate specificity of BtIDO. Enzymatic reactions performed with L-isoleucine revealed that the optimal reaction conditions are at pH 6.0 and 25 °C with a specific activity of 0.71 ± 0.29 U/mg. Overall, it was found that BtIDO is capable of catalyzing C3-, C4- or C5-hydroxylations such as the hydroxylation of L-norleucine into 4-hydroxynorleucine and 5-hydroxynorleucine (Hibi et al., 2011).

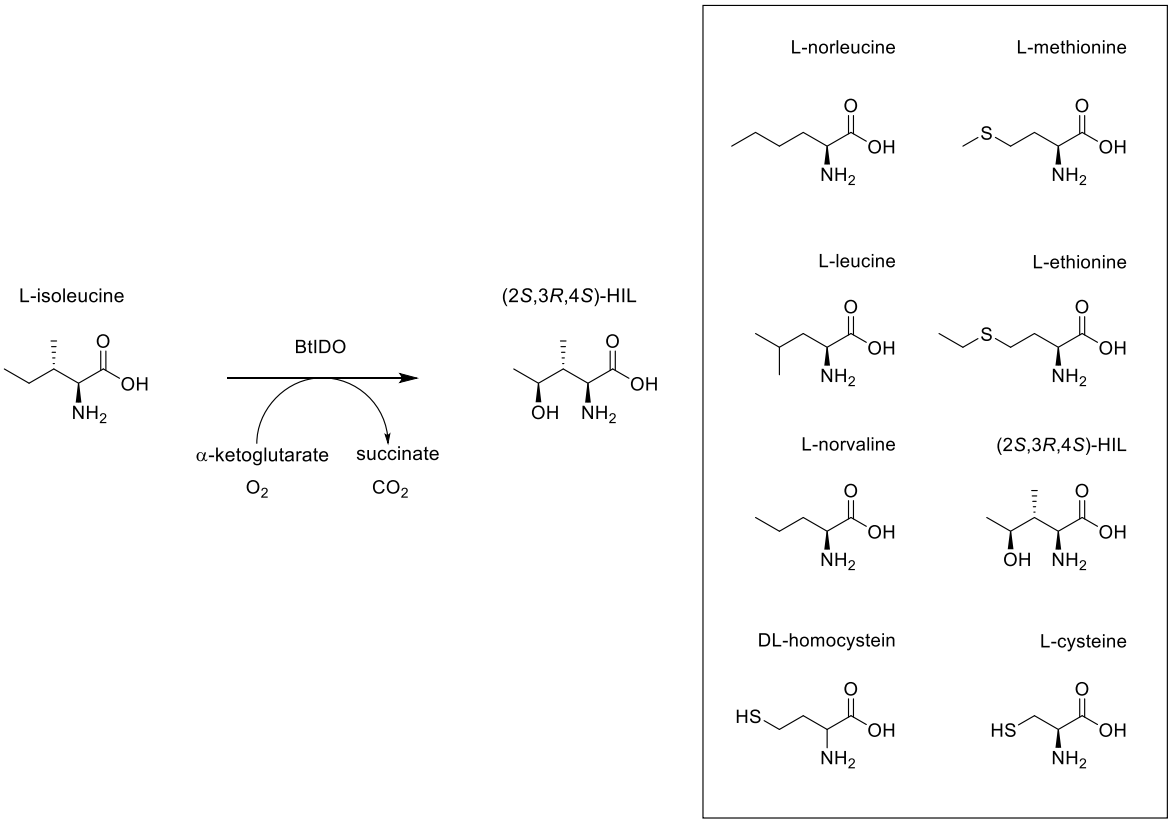

Figure 6: BtIDO catalyzes the oxyfunctionalization of several amino acids such as the hydroxylation of L-isoleucine generating (2*S*,3*R*,4*S*)-HIL. Adapted from (Hibi et al., 2011).

The second oxidation reaction BtIDO catalyzes is the dehydrogenation of (2*S*,3*R*,4*S*)-HIL into (2*S*,3*R*)-2-amino-3-methyl-4-ketopentanoate (AMKP), also

familiar as a vitamin B_{12} antimetabolite (Hibi et al., 2011; Peters and Buller, 2019). This pathway is highly important in *Bacillus* species as it shunts an incomplete tricarboxylic acid (TCA) cycle. For instance, *Bacillus thuringiensis* possesses an incomplete TCA cycle and is therefore unable to convert α-ketoglutarate via the oxidative branch. However, succinate production is ensured by the γ-aminobutyrate catabolic pathway converting α-ketoglutarate into L-glutamate via transamination reaction of (2*S*,3*R*,4*S*)-HIL and AMKP and further to succinate (Aronson et al., 1975; Hibi and Ogawa, 2014; Ogawa et al., 2011). Finally, yet importantly, the third oxidation reaction catalyzed by BtIDO is the sulfoxidation of L-amino acids containing a sulfur atom such as the enantioselective sulfoxidation of L-ethionine, L-methionine, *S*-ethyl-L-cysteine and *S*-methyl-L-cysteine. These sulfoxides find great demand in metabolic pathways of mammals, whereas (+)-methiin is an effective drug due to its antibiotic, antidiabetic and anti-inflammatory properties (Hibi and Ogawa, 2014; Hibi et al., 2011).

Furthermore, IDO belongs to the Pfam family PF10014 of conserved, yet uncharacterized, bacterial proteins. Smirnov *et al.* (Smirnov et al., 2012) found via *in silco* screening five α-ketoglutarate dependent dioxygenases homologous to IDO in their amino acid sequence. Substrate scope analysis revealed that all discovered dioxygenases are able to catalyze the sulfoxidation of L-methionine. In addition, the leucine dioxygenase from *Methylobacillus flagellatus*, short MfIDO, is even capable of hydroxylating L-leucine into (2*S*)-4-hydroxyleucine, but is inactive towards L-isoleucine (Hibi and Ogawa, 2014; Smirnov et al., 2012). Although, BtIDO and LdoA have low levels of sequence identity (19 %), both share the ability to catalyze the sulfoxidation of L-methionine and L-ethionine Moreover, both possess similar pH and temperature optima and kinetic parameters (Hibi et al., 2013).

1.7 L-leucine dioxygenase from *Streptomyces* strain DSM 40835 - GriE

The L-leucine dioxygenase from *Streptomyces* strain DSM 40835 catalyzes the selective hydroxylation of L-leucine at C5 position. Zwick *et al.* (Zwick and Renata, 2018a) investigated the substrate scope of GriE and found further amino acid derivates which are accepted as substrates to be hydroxylated at γ-position (Figure 7). Besides, the enzyme seems to be sensitive for substitutions in β-position (Peters and Buller, 2019; Zwick and Renata, 2018a). For instance, very appealing

diastereoselectivities (>99 %) were obtained when L-leucine, L-norleucine and DL-homonorleucine are hydroxylated at γ-position. No activity was observed for D-leucine, L-valine and L-isoleucine, whereas L-allo-isoleucine on the other hand was hydroxylated by GriE yielding (*S*)-5-hydroxy-l-*allo*-isoleucine (Zwick and Renata, 2018a).

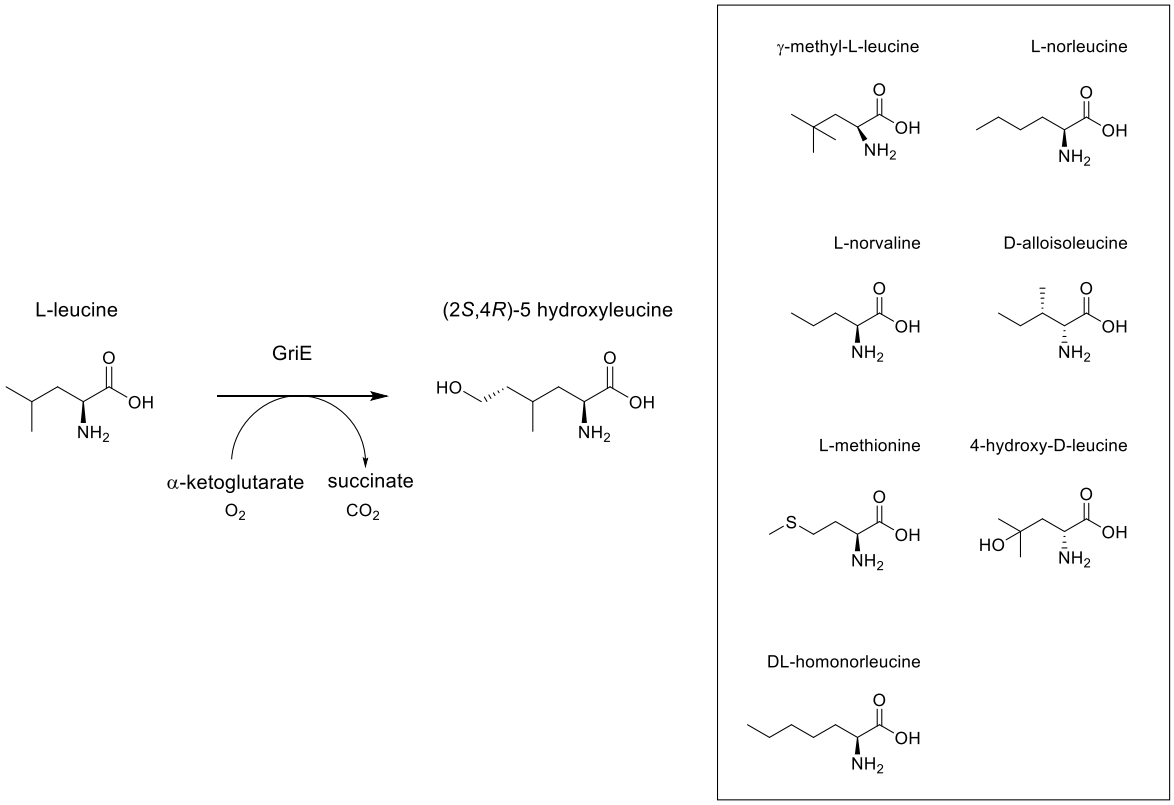

Figure 7: Reaction mechanisms and substrate spectrum of GriE. The hydroxylation of L-leucine into (2*S*,4*S*)-5-hydroxyleucine has been investigated in this thesis. Adapted from (Zwick and Renata, 2018a).

Moreover, GriE is a promising biocatalyst for the generation of valuable building blocks as the enzyme can be employed in a chemo enzymatic pathway for the synthesis of manzacidin C, cavinafungin B and the anti-tuberculosis griselimycins Important building blocks for the biosynthesis of those compounds are (2*S*,4*S*)-5-hydroxyleucine and (2*S*,4*R*)-4-methylproline which are both generated by the hydroxylation of L-leucine by GriE (Lukat et al., 2017; Peters and Buller, 2019; Zwick and Renata, 2018b, 2018a). Studies reported, that two other leucine 5-hydroxylases, LdoA5a and EcdK, are part of the biosynthesis of 4-methylproline. Nevertheless, due to the low substrate scope of LdoA and the instability of EcdK during reactions, GriE possess better properties as a biocatalyst (Zwick and Renata, 2018a). Furthermore, Lukat *et al.*(Lukat et al., 2017) determined the crystal structure

of GriE giving insights into the reaction mechanism α-ketoglutarate dependent dioxygenases are willing to perform.

1.8 Selection of Fe(II)/α-ketoglutarate dependent dioxygenases

Hydroxylated amino acids find appealing interest as valuable building blocks for various industries ranging from food industries to pharmaceutical industries. α-Ketoglutarate dependent dioxygenases require the cosubstrate α-ketoglutarate for catalyzing hydroxylation reactions which is obtained either from external addition or by the microbial metabolism from the TCA cycle. Amino acid hydroxylases assimilate α-ketoglutarate leading to the formation of the coproduct succinate which is then recycled by the carbon metabolism. Therefore, engineering of such metabolic pathways is highly fascinating for improving the biosynthesis of hydroxylated amino acids as well as providing insight into metabolic pathways (Peters and Buller, 2019; Theodosiou et al., 2017). For instance, in studies of Smirnov *et al.* (Smirnov et al., 2010) and Theodosiou *et al.* (Theodosiou et al., 2017) the TCA cycle of *E. coli* was engineered in order to enhance α-ketoglutarate levels and increase flux through the hydroxylating activity of P4H and IDO leading to an optimized production of (2*S*,3*R*,4*S*)-HIL and *trans*-4-hydroxy-L-proline (Peters and Buller, 2019; Smirnov et al., 2010; Theodosiou et al., 2017). Due to the deletion of the genes *sucA*, encoding for the E1 subunit of α-KG dehydrogenase, and *sucC*, encoding for succinyl-CoA synthetase β subunit together with *aceA*, which encodes for isocitrate lyase, succinate production via cellular TCA cycle was blocked (Figure 8). However, the enhanced carbon flux through the dioxygenases led to succinate generation and restoration of the TCA cycle (Theodosiou et al., 2017).

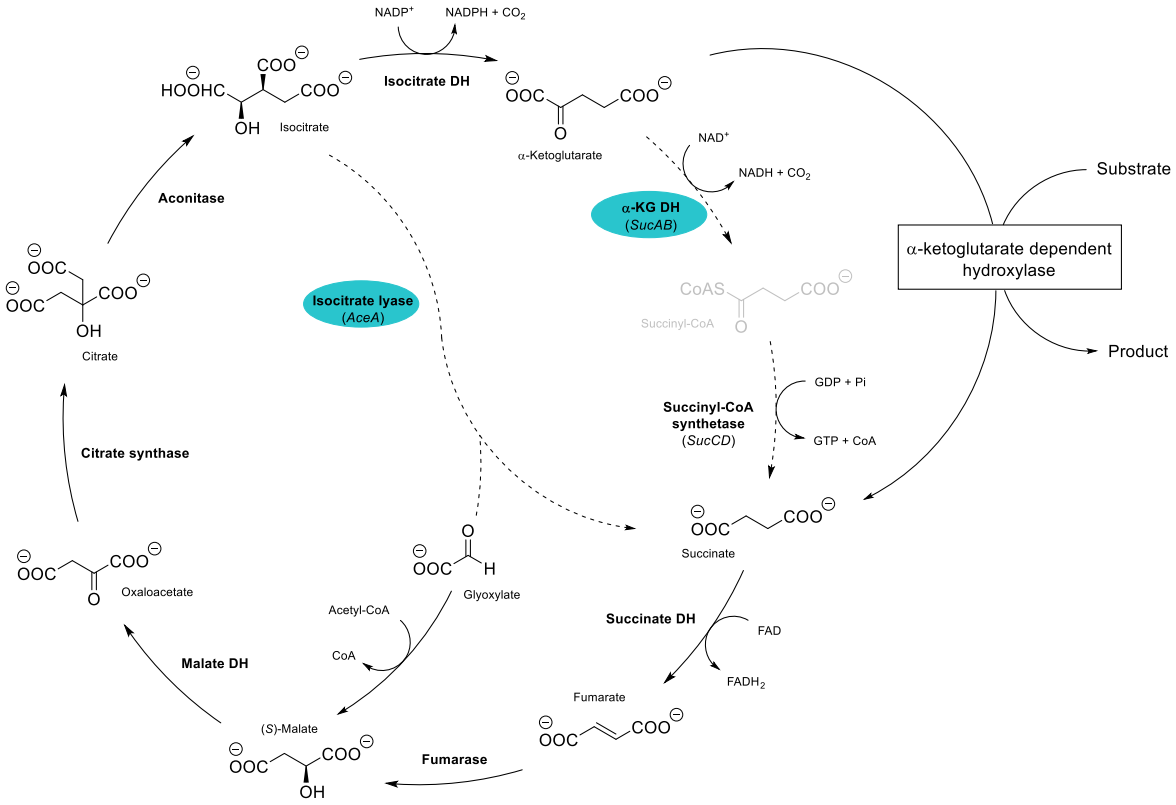

Figure 8: Proposed mechanism of the TCA re-cyclization in engineered *E. coli* 3Δ*sucA* cells via dioxygenase catalyzed L-amino acid hydroxylation. Adapted from (Theodosiou et al., 2017).

The reduced succinyl-CoA pool and the lack of the glyoxylate pathway provokes a heavy metabolic stress on the mutant strain leading to 27 % decreased cell size, lower growth rates and final biomass concentrations. The limitations in succinyl-CoA have impact on the physiology and biocatalytic efficiency, as succinyl-CoA serves as a precursor for the biosynthesis of lysine, methionine and diaminopimelate (DAP), where the latter is involved in the cell wall formation (Theodosiou et al., 2017). Zhang *et al.* (Zhang et al., 2018) reported that external addition of succinate has a positive effect on the growth of *E. coli* 3Δ*sucA* enhancing biomass production with increasing concentrations of succinate. Moreover, accumulation of (2*S*,3*R*,4*S*)-HIL and biomass can be improved by enhanced addition of IPTG as the increased activity of IDO leads to higher succinate production (Zhang et al., 2018). If other succinate pathways are blocked, *E. coli* 3Δ*sucA* is only able to grow in minimal medium when the activity of a dioxygenase shunts the broken TCA cycle, thereby coupling hydroxylation of a desired substrate and restore growth of *E. coli* (Figure 8). In case of *E. coli* 3Δ*sucC* growth can be restored in minimal medium without the hydroxylating activity of a dioxygenase but dependent on the fumarate reductase activity (Theodosiou et al., 2017). Substrate uptake restrictions can be avoided by overexpressing Na^+ dependent transporters such as *putA* in case of proline uptake

and BrnQ for L-isoleucine uptake (Ohnishi et al., 1988; Smirnov et al., 2010; Theodosiou et al., 2017). Li *et al.* (Li et al., 2006) reported that a *sucA* deletion results in an upregulation of pentose phosphate pathway enzymes as well as anaplerotic reactions including phosphoenolpyruvate (PEP) carboxykinase, PEP carboxylase and malate dehydrogenase. Several enzymes of glycolysis and TCA cycle enzymes on the other hand were downregulated. Due to the absence of the glyoxylate shunt, the only anaplerotic reaction for oxaloacetate production is PEP carboxylation (Li et al., 2006; Sauer and Eikmanns, 2005; Theodosiou et al., 2017). Although *E. coli* 3Δ*sucA* and *E. coli* 3Δ*sucC* produce acetate after glucose is depleted, both are incapable of assimilating it due to the broken glyoxylate shunt (Theodosiou et al., 2017).

These studies showcase the great potential of *E. coli* 3Δ*sucA* as a selection tool for active Fe(II)/α-ketoglutarate dependent hydroxylases. Besides, this selection method can be applied for screening of enzyme libraries, which were generated through random mutagenesis or metagenomics (Theodosiou et al., 2017; Zhang et al., 2018). For instance, Zhang *et al.* (Zhang et al., 2018) generated a (2*S*,3*R*,4*S*)-HIL library via random mutagenesis and obtained five mutants with enhanced IDO activities. Moreover, one variant showed even improved catalytic properties and thermal stability compared to the wild type. To overcome substrate uptake limitations as well as export of the hydroxylated product (2*S*,3*R*,4*S*)-HIL, the research group generated resting cells by freezing *E. coli* BL21(DE3) cells which overexpress the IDO variant at -80 °C. Biotransformations performed with these resting cells overexpressing the IDO variant resulted in synthesis of 151.9 mmol of (2*S*,3*R*,4*S*)-HIL /L (22.4 g/L) in 12 h (Zhang et al., 2018). Furthermore, *E. coli* 3Δ*sucA* can be applied for the screening of substrate scopes, whereas non-canonical amino acids would find tremendous interest for investigation. Any substrate can be employed for hydroxylation as long as the following conditions are respected. First, expression of the dioxygenase gene in an active form has to be ensured, moreover, toxicity of the substrate and the hydroxylated amino acid for cells has to be excluded, and an efficient transfer of the substrate and product through the cell membrane has to be guaranteed (Smirnov et al., 2010; Theodosiou et al., 2017).

1.9 Aims

The aim of this study was to (a) establish a non-heme Fe(II)/α-ketoglutarate dependent dioxygenase toolbox of overall eleven enzymes by investigating soluble expression levels in recombinant *E. coli* cells and conduct a TLC screening in order to determine the substrate acceptance of each dioxygenase.

(b) adapt soluble expression levels of AvLDO according to literature as the yields obtained in this study are insufficient for *in vitro* reactions. An optimization was attempted via co-expression of chaperones, adaptions in inducer concentrations and cultivation temperatures as well as performing expression in autoinduction medium.

(c) Create a selection tool for active Fe(II)/α-ketoglutarate dependent dioxygenases via the implementation of an engineered *E. coli* strain. Growth assays in liquid medium were executed with three dioxygenases including BtIDO, GriE and AvLDO. In addition, the growth behavior of *E. coli* BL21(DE3) was compared to the mutant strain.

2 Results

2.1 Expression of recombinant dioxygenases

2.1.1 Whole cells

The expression of recombinant dioxygenases was analyzed by SDS-PAGE as described in 5.5.5.1. The t_0 samples served as controls since no protein was overexpressed at this point and only native *E. coli* proteins are visible (Figure 9). Overnight samples of the putative dioxygenase from *Tistrella mobilis* (TmDO) and GriE showed bands with a high intensity at the expected height and clearly indicate overexpression of the desired protein. For AvLDO the highest amount of protein was reached after overnight expression. The t_4 sample on the other hand contained low amounts of protein. A high protein content was obtained in the t_4 fraction of the putative dioxygenase from marine metagenome (MMGDO) at the expected height. The overnight sample of MMGDO showed lower amounts of protein.

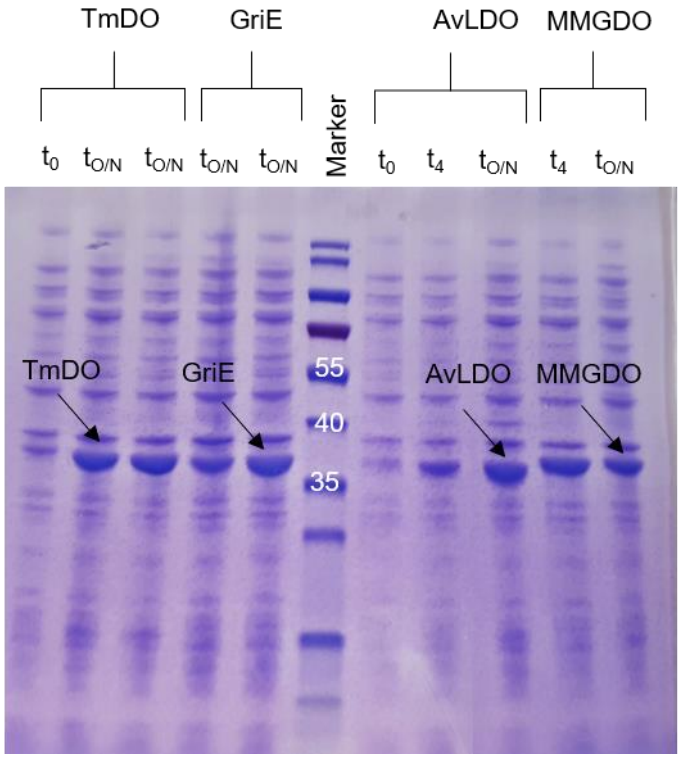

Figure 9: Whole cell profile of TmDO, GriE, AvLDO and MMGDO analyzed by SDS-PAGE. Overnight samples of TmDO and GriE contained a decent amount of protein at a height of 32 kDa. The overnight sample of AvLDO showed a band with high intensity at the expected height of 33 kDa. For MMGDO (35 kDa) the highest amount of protein was present in the t_4 sample. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

The t_0 sample of the dioxygenase from *Burkholderia ambifaria* AMMD (SadA) and the proline-4-hydroxylase (P4H) served again as a standard of native *E. coli* proteins

(Figure 10) For SadA and P4H the highest amount of protein was reached after 4 h of expression. The overnight samples contained lower amounts of protein. Both samples of the cyclic amino acid hydroxylase from *Fusarium oxysporum* c8D (FoPip4H) and the cyclic amino acid hydroxylase from *Aspergillus nidulans* FGSCA4 (AnPip4H) showed overexpression bands with the same intensity at the expected size.

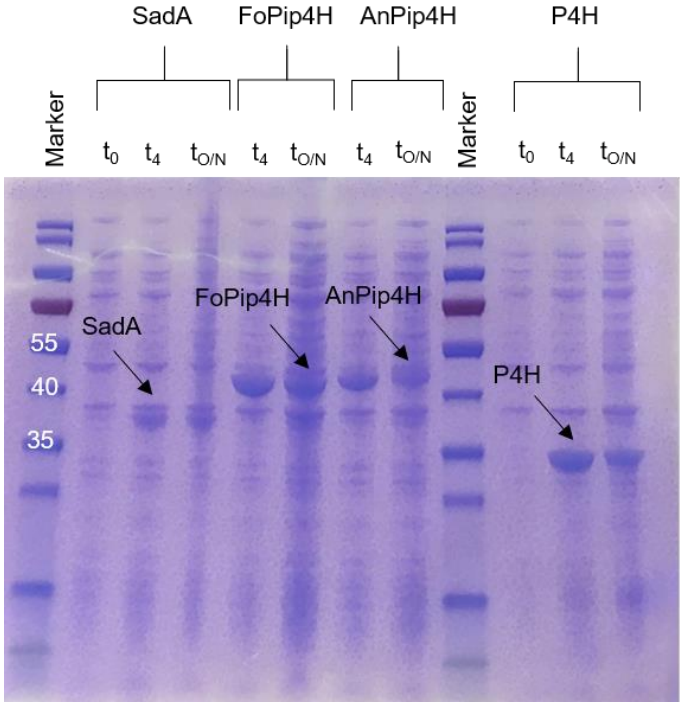

Figure 10: Whole cell profile of SadA, FoPip4H, AnPip4H and P4H analyzed by SDS-PAGE. Overnight and t_4 samples of FoPip4H and AnPip4H contained overall the same amount of protein at the expected height of 40 kDa. For SadA (33 kDa) and P4H (30 kDa) the highest amount of protein was present in the t_4 sample. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

Based on the SDS PAGE, expression levels of the lysine dioxygenase from *Catenulispora acidiphila* (KDO1), BtIDO and MfIDO were not conclusive as no bands of the desired protein were detected (Figure 11).

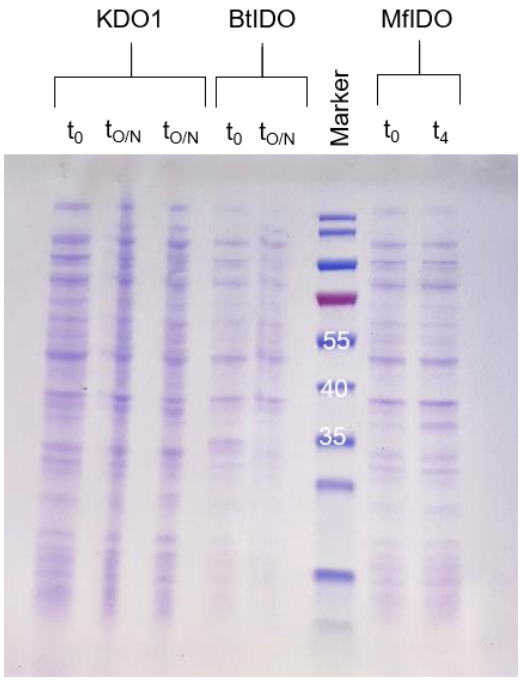

Figure 11: Whole cell profile of KDO1 (43 kDa), BtIDO (29 kDa) and MfIDO (32 kDa) analyzed by SDS-PAGE. For all three enzymes, no visible bands for the desired protein were obtained. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.1.2 Cell disruption with BugBuster®

Cells were disrupted with BugBuster® (see 5.5.5.1) and insoluble and soluble fractions were analyzed by SDS-PAGE. Overnight samples of TmDO showed a high amount of soluble and insoluble protein at the expected size (Figure 12 and Figure 13). For GriE, a decent amount of protein was present in the soluble fraction. This was not the case for the insoluble fraction. No visible bands of soluble protein were obtained for AvLDO. The insoluble fractions on the other hand contained a large amount of protein. For MMGDO, a band with a high intensity was visible in the insoluble as well as in the soluble fraction after 4 h expression. The overnight sample on the other hand contained lower amounts of protein.

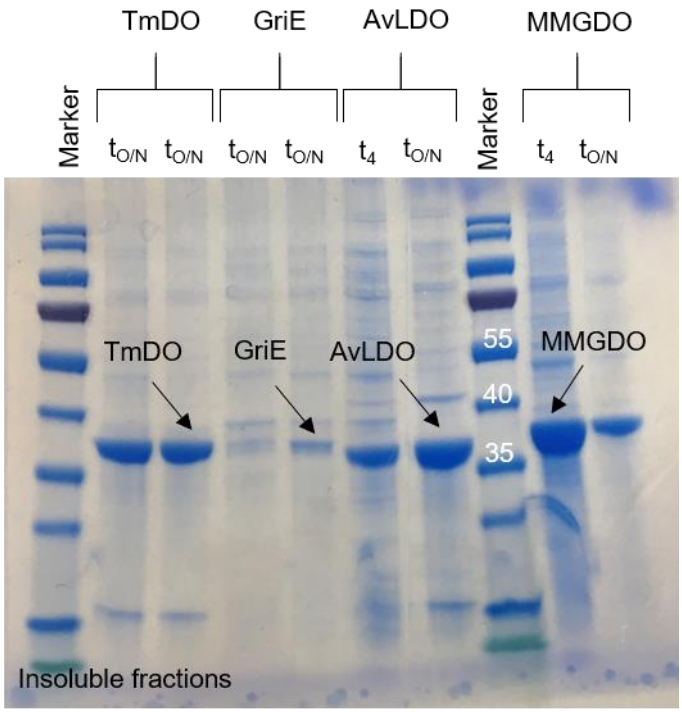

Figure 12: Insoluble fractions of TmDO, GriE, AvLDO and MMGDO analyzed by SDS-PAGE. Overnight samples of TmDO (32 kDa) contained a decent amount of protein. Bands with low intensity were visible for GriE (32 kDa). The second overnight sample of AvLDO (33 kDa) contained a high amount of protein. For MMGDO (35 kDa) the highest amount of protein was present in the t_4 sample. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

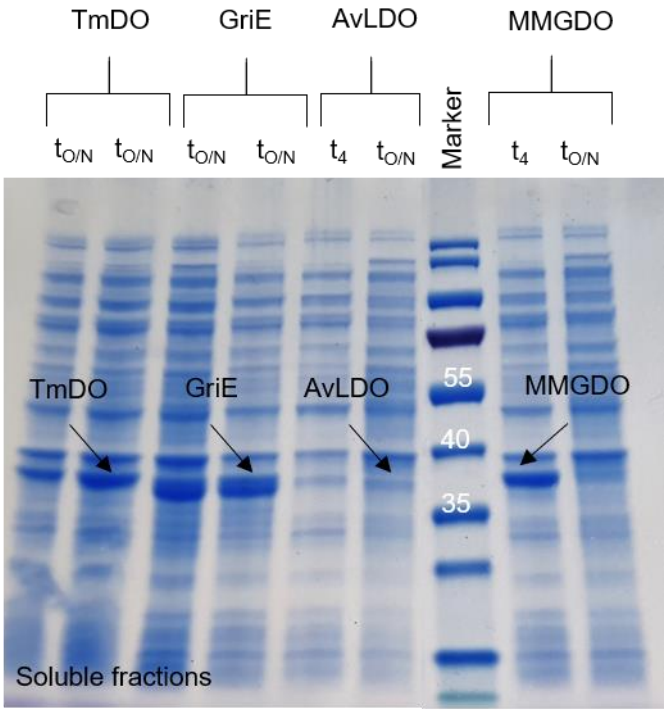

Figure 13: Soluble fractions of TmDO, GriE, AvLDO and MMGDO analyzed by SDS-PAGE. Overnight samples of TmDO and GriE contained a high amount of protein at the expected size of 32 kDa. No visible bands of soluble protein were obtained for AvLDO (33 kDa). For MMGDO (35 kDa) the highest amount of protein was present in the t_4 sample. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

No visible bands of soluble protein were obtained for KDO1, BtIDO and MflDO. Due to pipetting errors, no bands were present in the insoluble fraction (Figure 14).

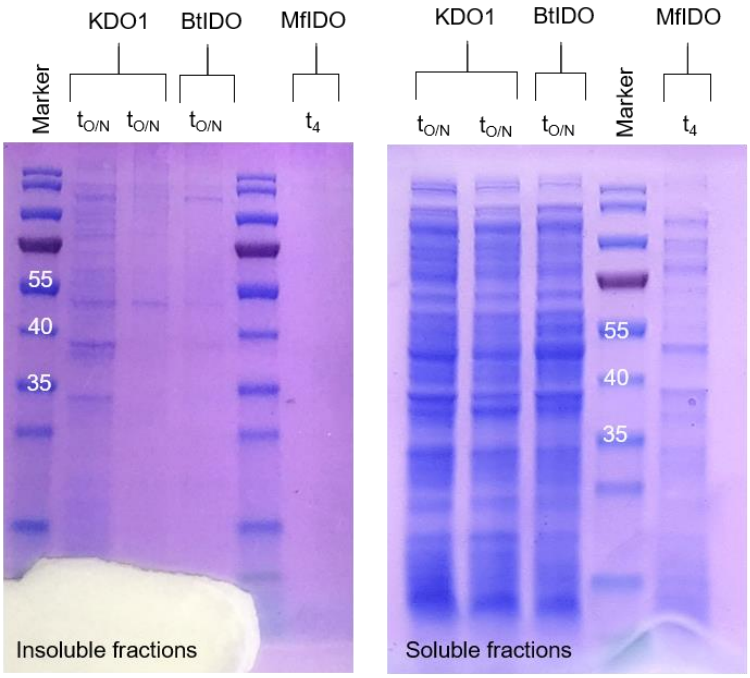

Figure 14: Insoluble and soluble fractions of KDO1, BtIDO and MflDO analyzed by SDS-PAGE. For the insoluble fractions of KDO1 (43 kDa), BtIDO (29 kDa) and MflDO (32 kDa) no bands were visible at the expected height. Samples of BtIDO and MflDO showed no bands due to pipetting errors. For all three enzymes, soluble protein was not obtained at the expected sizes. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

No visible bands of soluble protein were obtained for FoPip4H at the expected height, but a band was slightly visible in the insoluble fraction (Figure 15 and Figure 16). For AnPip4H, a decent amount of protein was present in the t_4 sample of the soluble fraction at the expected height. The insoluble fraction showed a band with low intensity. The t_4 sample of P4H showed a high amount of insoluble protein. However, no bands of soluble protein were obtained. The overnight sample on the other hand showed a higher amount of soluble protein. For SadA, the highest amount of soluble protein was reached after 4 h expression. Only bands with low intensity were visible in the insoluble fraction.

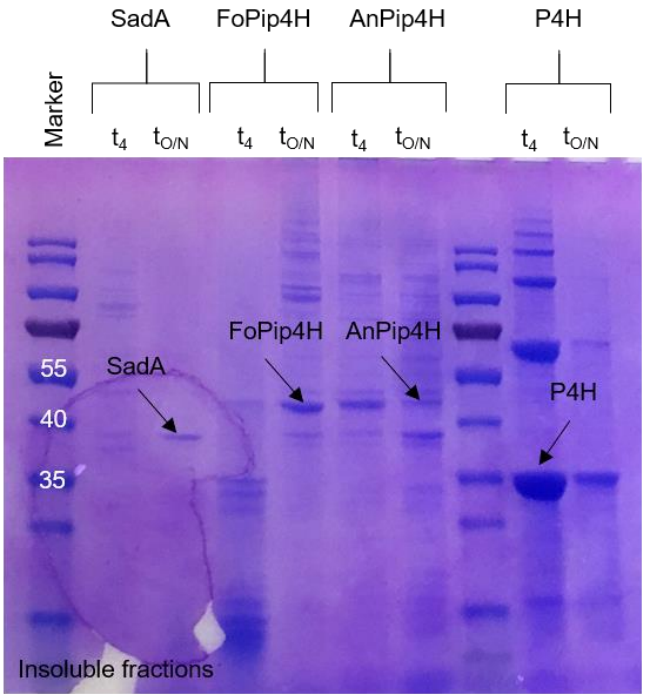

Figure 15: Insoluble fractions of SadA, FoPip4H, AnPip4H and P4H analyzed by SDS-PAGE. Overnight samples of SadA (33 kDa) and FoPip4H (40 kDa) showed bands with low intensity at the expected size. Both samples of AnPip4H (40 kDa) contained a low amount of protein. The t_4 sample of P4H (30 kDa) contained a high amount of protein. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

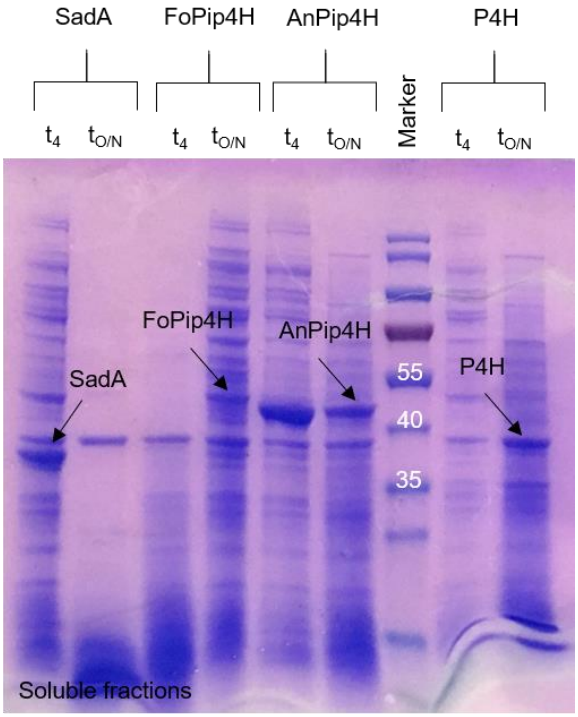

Figure 16: Soluble fractions of SadA, FoPip4H, AnPip4H and P4H analyzed by SDS-PAGE. The t4 sample of SadA (33 kDa) contained a high amount of protein compared to the overnight sample. The overnight sample of FoPip4H (40 kDa) contained a low amount of protein. The t4 sample of AnPip4H showed bands with high intensity at 40 kDa. For P4H (30 kDa) the highest amount of protein was present in the overnight sample. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.2 Protein purification

The CFE of six dioxygenases was subjected to affinity chromatography under the corresponding conditions (see 5.5.3) and each collected purification fraction was analyzed by SDS-PAGE (see 5.5.5.2).

2.2.1 L-leucine dioxygenase from *A. variabilis* – AvLDO

Bands with a light intensity were visible in the first (F1) and second fraction (F2) of wash II at the expected height of 33 kDa. (Figure 17). This indicated, that a small amount of protein was already eluted at an imidazole concentration of 30 mM. The first and second elution fraction (E1, E2) contained the highest amount of protein when using an imidazole concentration of 300 mM. Even though faint bands are visible, the elution fractions are reasonably pure. No bands are visible on the gel for the third washing step (wash III) due to pipetting errors.

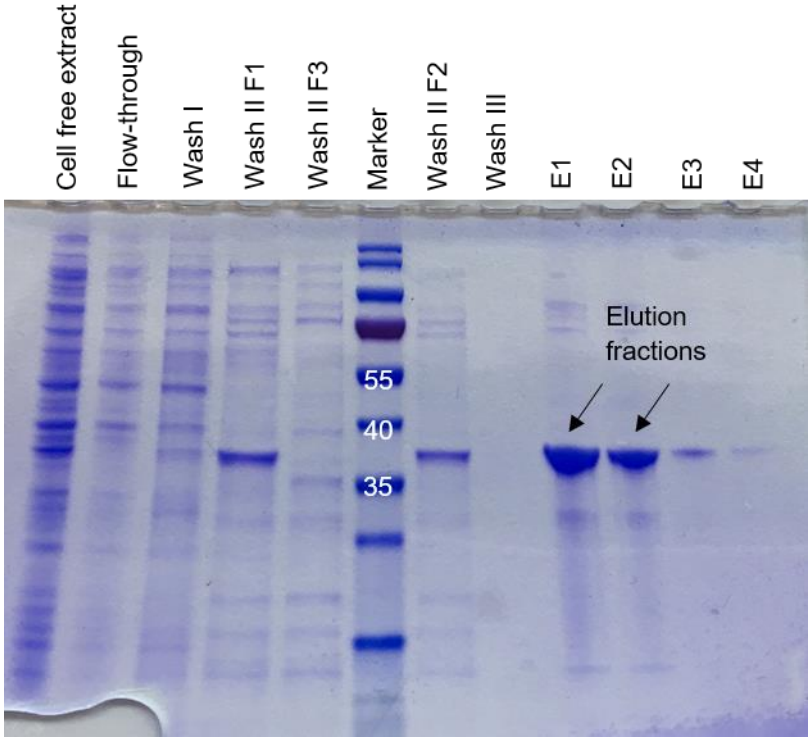

Figure 17: Purification fractions of AvLDO analyzed by SDS-PAGE. Low amounts of protein were present in wash fraction II at 33 kDa. The first and second elution fraction contained the highest amount of protein. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.2.2 L-isoleucine dioxygenase from *B. thuringiensis* - BtIDO

A sufficient amount of protein was obtained in the first and second elution fraction (E1, E2 in Figure 18) at the expected height of 29 kDa. However, unspecific bands were visible as well. These originate from unknown proteins, which bound unspecifically to the column material and were washed off during elution or an insufficient washing procedure of the column. The first two washing fractions are free of desired product, but in the wash III fraction, a band was slightly visible indicating that traces of protein remained on the column after elution.

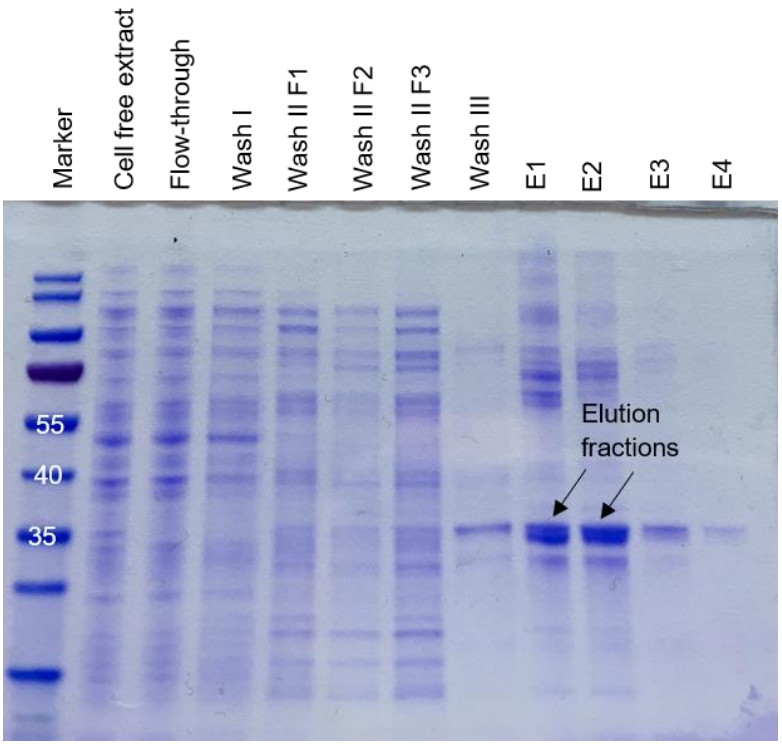

Figure 18: Purification fractions of BtIDO analyzed by SDS-PAGE. The first and second elution fraction contained the highest amount of protein at the expected height of 29 kDa. Traces of enzyme were present after elution (wash III). Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.2.3 L-leucine dioxygenase from *Streptomyces strain* DSM 40835 – GriE

The flow-through contained a high amount of desired protein at the expected height of 32 kDa (Figure 19). Reasons for that might be an insufficient amount of resin applied to the column. Bands with a high intensity were obtained in the first and second washing steps at the expected height indicating that the protein was already eluted at an imidazole concentration of 30 mM. Although a decent amount of protein was received in the first and second elution fraction (E1, E2) traces of protein were visible in the last washing step (wash III).

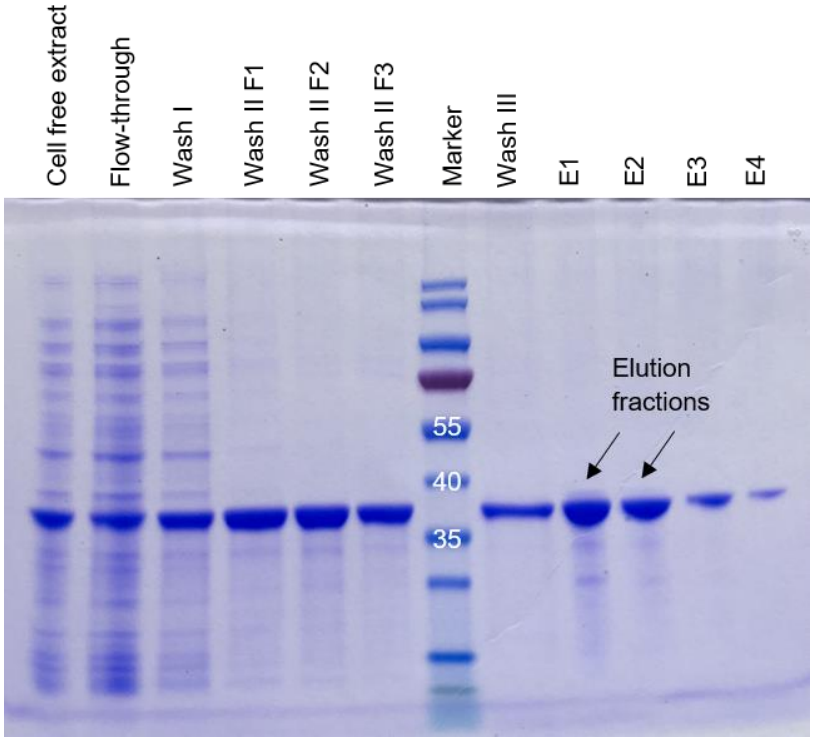

Figure 19: Purification fractions of GriE analyzed by SDS PAGE. A high amount of protein can be seen in almost all fractions with the expected size of 32 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.2.4 N-succinyl amino acid hydroxylase from *B. ambifaria* AMMD – SadA

The CFE and wash I fraction showed faint bands of the protein of interest at the expected height of 33 kDa (Figure 20). A decent amount of protein was obtained in the second washing step with an imidazole concentration of 30 mM. The majority of protein was detached in the first and second elution fraction (E1, E2) when using an imidazole concentration of 300 mM. After elution, a decent amount of protein remained on the column and eluted with wash III (500 mM imidazole). No bands were visible in the flow-through due to pipetting errors.

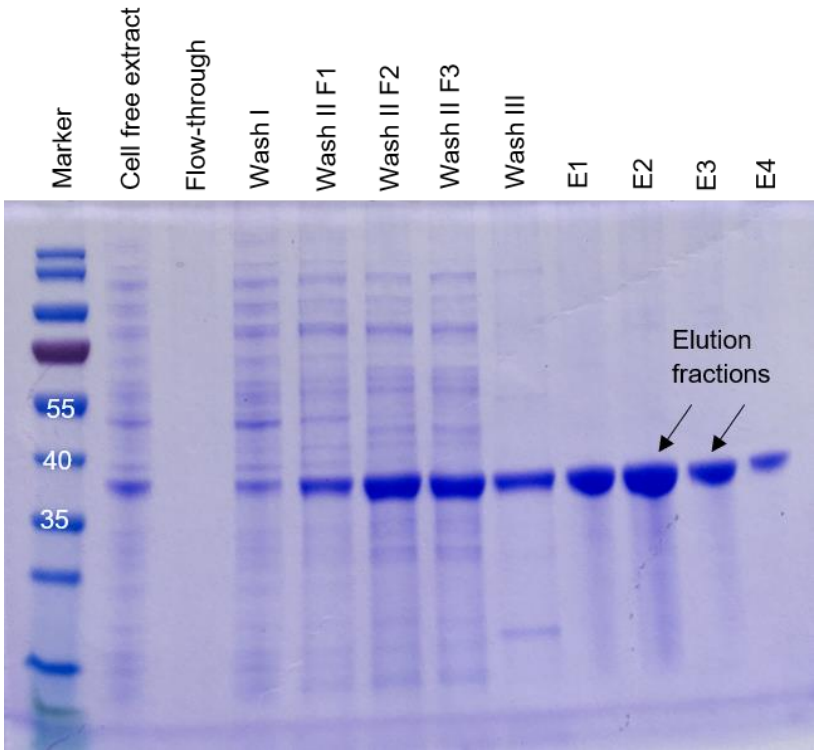

Figure 20: Purification fractions of SadA analyzed by SDS-PAGE. A high amount of protein already eluted with the second washing step at the expected height of 33 kDa. The majority of protein was obtained during elution. After elution, protein was still left on the column (wash III). Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.2.5 Cyclic amino acid hydroxylase from *F. oxysporum* – FoPip4H

Bands with a low intensity were visible in all fractions of the second washing step at the expected height of 40 kDa indicating that a small amount of protein was already eluted at an imidazole concentration of 30 mM (Figure 21). The first and second elution fraction (E1, E2) contained the highest amount of the protein of interest. However, these fractions were not completely pure. Traces of other proteins that bound to the column material were also visible in the form of unspecific bands or insufficient washing of the column was implemented. Some protein remained on the column after elution and was washed off with 500 mM imidazole during the last wash step (wash III).

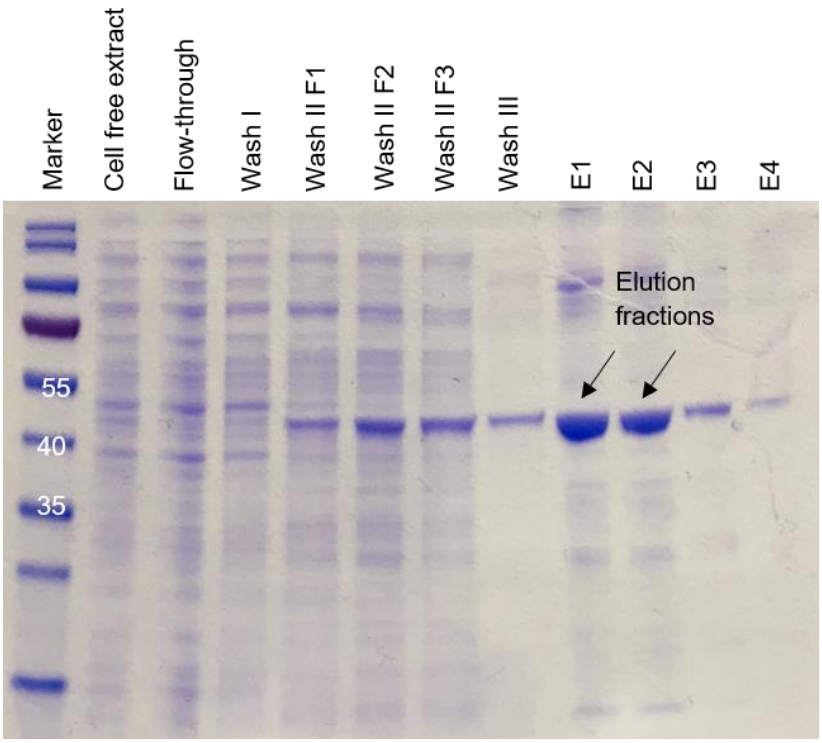

Figure 21: Purification fractions of FoPip4H analyzed by SDS-PAGE. Low amounts of the desired protein were present in the second washing fractions at 40 kDa. The first and second elution fraction contained the highest amount of protein. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.2.6 Cyclic amino acid hydroxylase from *A. nidulans* FGSC A4 – AnPip4H

A decent amount of protein was visible at the expected height of 40 kDa in all fractions of the second washing step due to a high imidazole concentration (Figure 22). The first and second elution fraction (E1, E2) contained the highest amount of desired protein. No bands were visible in the third washing step (wash III) and third elution fraction (E3) due to pipetting errors.

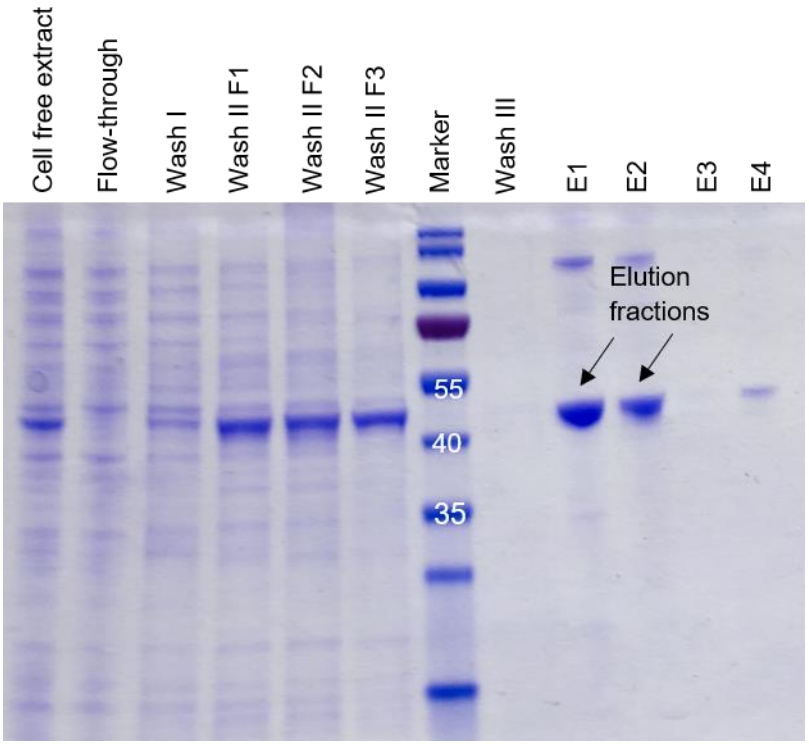

Figure 22: Purification fractions of AnPip4H analyzed by SDS-PAGE. Protein was present in all fractions of the second washing step. The first and second elution fraction contained the highest amount of protein with the expected size of 40 kDa.

2.3 Activity assay

Enzymatic reactions were performed at the corresponding conditions as described in 5.5.6 and hydroxylation activity of the dioxygenases was confirmed by TLC and ninhydrin staining (see 5.5.6.2). Among these eleven dioxygenases (Table 1) four (BtIDO, GriE, MfIDO and FoPip4H) showed activity towards the tested substrates. Biotransformations were set up once with CFE and once with purified enzyme except for TmDO, MMGDO and KDO1. The corresponding L-amino acids and hydroxy isoleucine served as standards. An overview of all R_f-values, calculated as described in 5.5.6.2 of each L-amino acid and standard can be seen in the appendix (see 7.7). For the rest of the dioxygenases, product formation was either not detectable or not conclusive (see appendix 7.7).

Table 1: Overview of all activity assay results for each dioxygenase.

Enzyme	Substrate	Type of protein	Activity as judged by TLC
AvLDO	L-leucine	Lysate: 0.7 mg/mL Purified: 1 mg/mL	Inactive
BtIDO	L-isoleucine	Lysate: 0.9 mg/mL Purified: 1 mg/mL	Active
SadA	N-succinyl-L-leucine	Lysate: 1.9 mg/mL Purified: 2 mg/mL	Inactive
TmDO	L-isoleucine/ L-methionine	Lysate: n.d.	Inactive
MMGDO	L-isoleucine/ L-methionine	Lysate: n.d.	Inactive
MfIDO	L-leucine/ L-methionine	Lysate: 0.7 mg/mL	Active
GriE	L-leucine/ α-methyl-L-leucine	Lysate: 1.7 mg/mL Purified: 2 mg/mL	Active
P4H	L-proline	Lysate: 0.7 mg/mL	Inactive
KDO1	L-lysine	Lysate: 0.12 mg/mL	Inactive
FoPip4H	L-pipecolic acid	Lysate: 1.2 mg/mL Purified: 1.8 mg/mL	Active
AnPip4H	L-pipecolic acid	Lysate: 0.4 mg/mL Purified: 1.2 mg/mL	Inactive

2.3.1 L-isoleucine dioxygenase from *B. thuringiensis* - BtIDO

Enzymatic reactions, which were performed with CFE, showed a conversion of L-isoleucine to hydroxy isoleucine after two hours (Figure 23 [A]). Full conversion was not reached within 30 hours. At position 4 and 5 of Figure 23[A] an additional unspecified spot was observed. Biotransformations set up with purified enzyme showed no activity towards L-isoleucine and only unspecific byproducts were obtained (Figure 23 [B]).

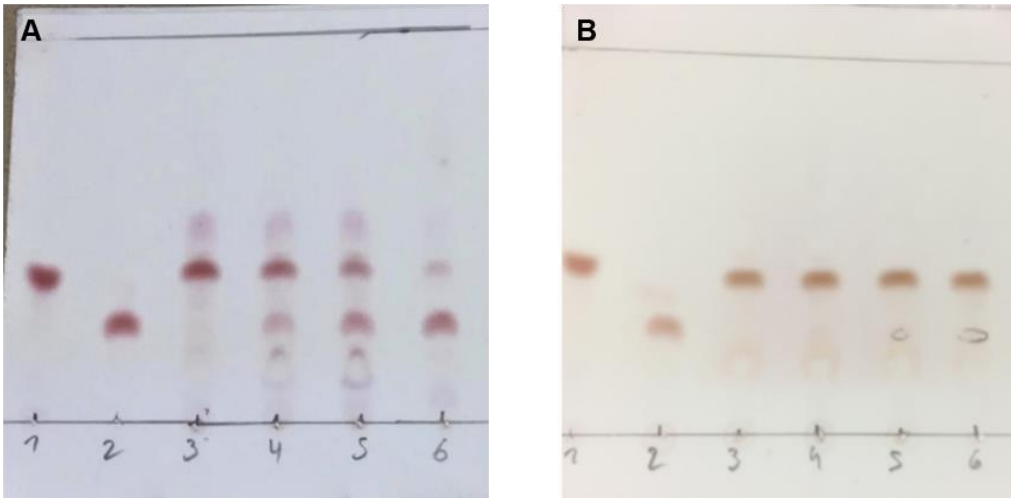

Figure 23: Biotransformations of BtIDO with L-isoleucine as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE and with purified protein. L-isoleucine and hydroxy isoleucine were used as standards. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A]** CFE: 0.9 mg/mL **[B]** purified protein: 1 mg/mL in 500 µL volume at 25 °C. **[A]** [1] L-isoleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_2; [5] t_5; [6] t_{30}; **[B]** [1] L-isoleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_2; [5] t_5; [6] t_{30}.

2.3.2 L-leucine dioxygenase from *Streptomyces* strain DSM 40835 – GriE

Biotransformations set up with CFE and L-leucine as substrate showed product formation after one hour and went up to full conversion within 24 hours (Figure 24 [A]) This was also the case when using purified enzyme, but full conversion was reached already after 18 hours (Figure 24 [B]). The enzyme was unreactive towards α-methyl-L-leucine under the implemented conditions (Figure 24 [C, D]) and only unspecific byproducts were obtained.

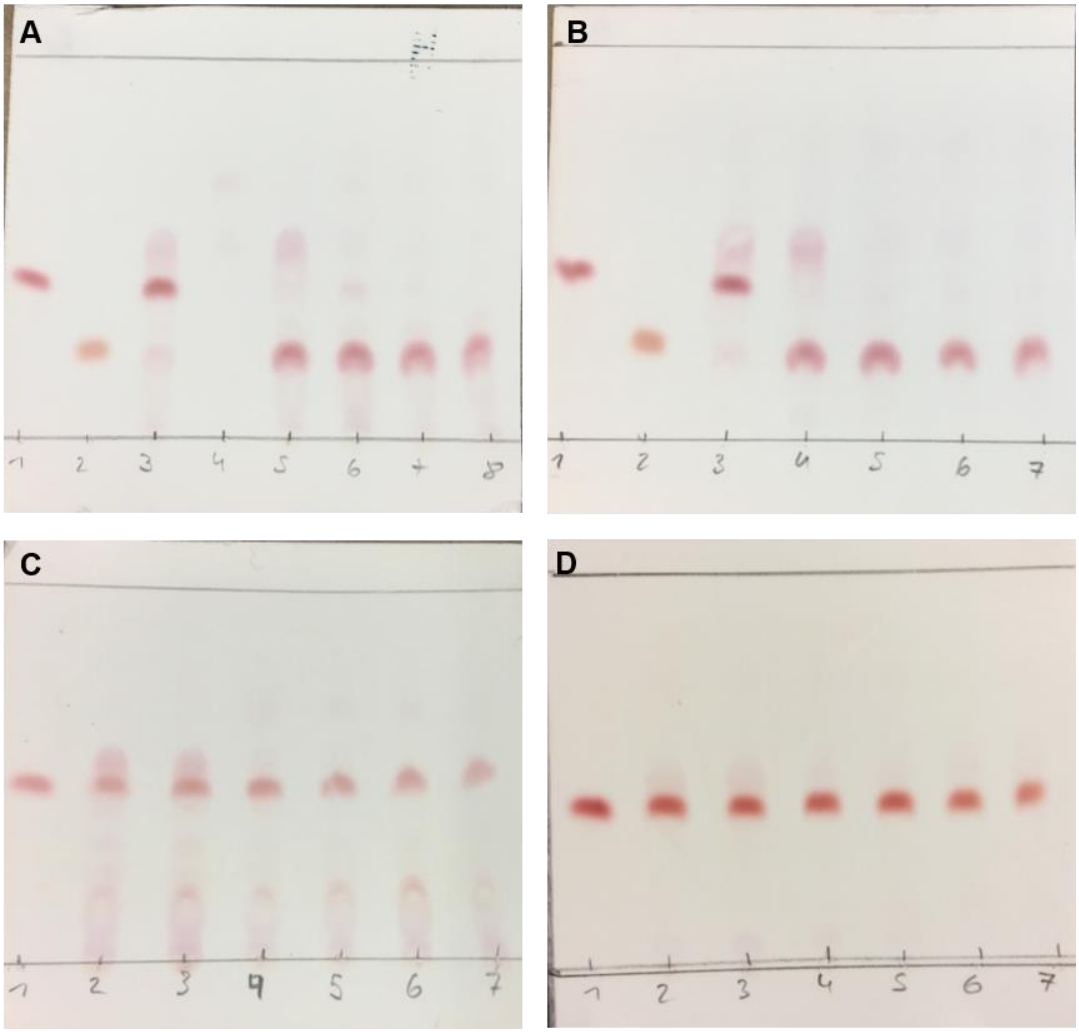

Figure 24: Biotransformations of GriE with L-leucine **[A, B]** and α-methyl-L-leucine **[C, D]** as substrates analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE and with purified protein. L-leucine, hydroxy isoleucine and α-methyl-L-leucine served as standards. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A]** CFE: 1.7 mg/mL **[B]** purified protein: 2 mg/mL **[C]** CFE: 1.7 mg/mL **[D]** purified protein: 2 mg/mL in 500 µL volume at 35 °C. **[A]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] negative control; [5] t_1; [6] $t_{O/N}$; [7] t_{24}; [8] t_{41}; [**B]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] $t_{O/N}$; [6] t_{24}; [7] t_{41}; **[C]** [1] α-methyl-L-leucine; [2] t_0; [3] t_1; [4] $t_{O/N}$; [5] t_{24}; [6] t_{41}; [7] t_{48}; **[D]** [1] α-methyl-L-leucine; [2] t_0; [3] t_1; [4] $t_{O/N}$; [5] t_{24}; [6] t_{41}; [7] t_{48}.

2.3.3 Cyclic amino acid hydroxylase from *F. oxysporum* – FoPip4H

Biotransformations implemented with CFE showed conversion of L-pipecolic acid to *trans*-4-hydroxy-L-pipecolic acid after one hour (Figure 25 [A]). Full conversion was not reached within 19 hours. Activity towards L-pipecolic acid was not detectable when reactions were performed with purified enzyme. (Figure 25 [B]).

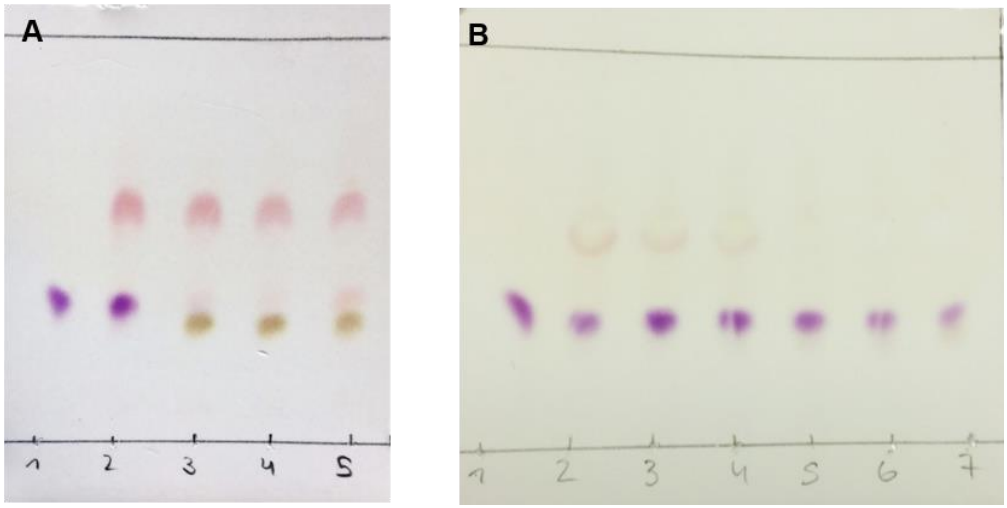

Figure 25: Biotransformations of FoPip4H with L-pipecolic acid as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE and with purified protein. L-pipecolic served as standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A]** CFE: 1.2 mg/mL **[B]** purified protein: 1.8 mg/mL in 500 µL volume at 20 °C. **[A]** [1] L-pipecolic acid; [2] t_0; [3] t_1; [4] t_2; [5] $t_{O/N}$; [**B]** [1] L-pipecolic acid; [2] t_0; [3] t_1; [4] t_2; [5] t_6; [6] t_{24}; [7] t_{68}.

2.3.4 L-leucine dioxygenase from *M. flagellatus* – MflDO

Reactions employing L-leucine showed product formation after one hour (Figure 26 [A]). When using L-methionine as substrate, substrate conversion to L-methionine sulfoxide was slightly detectable after one hour (Figure 26 [B]). For both reactions, full conversion could not be reached within 41 hours. Protein purification was not implemented for time reasons.

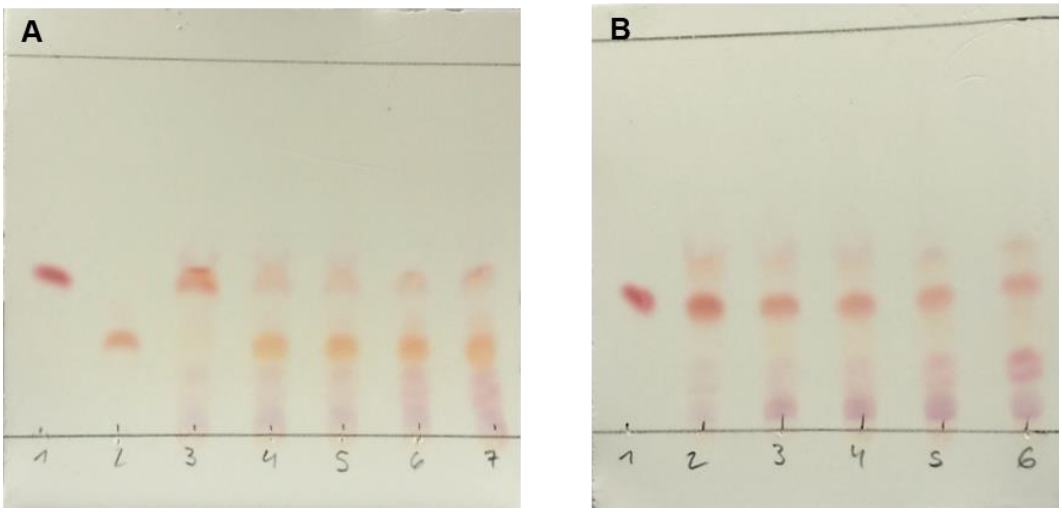

Figure 26: Biotransformations of MflDO with L-leucine **[A]** and L-methionine **[B]** as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE. L-leucine, L-methionine and hydroxy isoleucine served as standards. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, CFE: 0.5 mg/mL in 500 µL volume at 35 °C. **[A]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_3; [6] $t_{O/N}$; [7] t_{41} **[B]** [1] L-methionine; [2] t_0; [3] t_1; [4] t_3; [5] $t_{O/N}$; [6] t_{41}.

2.4 Expression optimization of AvLDO

Heterologous expression of AvLDO was investigated in *E. coli* BL21(DE3) based on expression conditions documented by Correia Cordeiro *et al*. (Correia Cordeiro et al., 2018). After cells were disrupted with BugBuster® (5.5.5.1), insoluble and soluble fractions were analyzed by SDS-PAGE. The t_0 sample served as a standard of native *E. coli* proteins since no protein was overexpressed at this point (Figure 27). The highest amount of soluble protein was reached after six hours of expression. At 24 hours of expression, a depletion of the protein concentration was observed. In the insoluble fraction, bands with a high intensity at the expected height were detected in the t_6 and t_{24} sample indicating inclusion body formation.

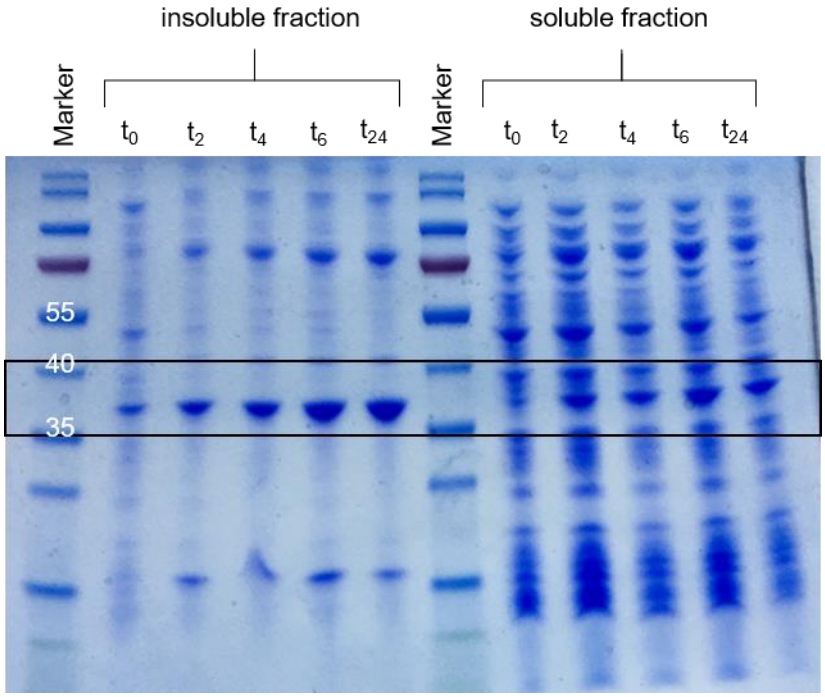

Figure 27: Insoluble and soluble fractions of AvLDO analyzed by SDS-PAGE. The highest amount of soluble protein was obtained after six hours of expression at the expected height 33 kDa. The t_6 and t_{24} sample of the insoluble fraction contained a high amount of protein indicating inclusion body formation.

To improve soluble expression of AvLDO further, different IPTG concentrations and cultivation temperatures were examined (see appendix 7.8). For all approaches, large amounts of the desired protein were visible in the insoluble fractions, which clearly indicates inclusion body formation. Bands with low protein content were obtained in the soluble fractions. However, in most cases those bands occurred already at t_0 indicating no overexpression of AvLDO and correspond to native *E. coli* proteins.

Since high amounts of AvLDO were detected as insoluble protein, co-expression of chaperones was implemented in order to decrease and circumvent inclusion body formation. Expression of the desired protein and the chaperone proteins GroEL/ES was not detectible in both the insoluble fraction as well as in the soluble fraction. Moreover, the addition of $FeSO_4$ was investigated to optimize soluble expression. The insoluble fraction contained again large amounts of protein at the expected height. No bands were visible in the soluble fraction (see appendix 7.8).

An expression study was also performed with autoinduction medium using *E. coli* BL21(DE3) as host organism. SDS-PAGE analysis revealed a high amount of insoluble protein in the t_4 and t_{24} sample. Due to pipetting errors, no bands were visible for t_0 and t_2 sample. Since most of the expressed dioxygenase was obtained as inclusion bodies and overexpression of the soluble protein was not detectable, *E. coli* JM109(DE3) was employed for recombinant expression of AvLDO. In the insoluble fraction, a band with a high protein content was obtained for the t_{24} sample. For the rest of the samples no bands at the expected height were visible due to pipetting errors. Again, overexpression of soluble protein was not detectable (see appendix 7.8).

2.5 Purification of AvLDO

After CFE of AvLDO was subjected to affinity chromatography, each collected purification fraction was analyzed by SDS-PAGE, described in 5.5.5.2. Due to pipetting errors, no bands were visible in the CFE fraction (Figure 28). The majority of the desired protein was obtained in the elution fractions at a height of 33 kDa when using an imidazole concentration of 300 mM, whereas the second elution fraction contained most protein. However, traces of other proteins that bound to the column material were also visible in the form of unspecific bands. The pellet fraction contained a low amount of protein. A faint band belonging to the protein of interest was visible in the wash III fraction indicating that traces of protein still remained on the column after elution.

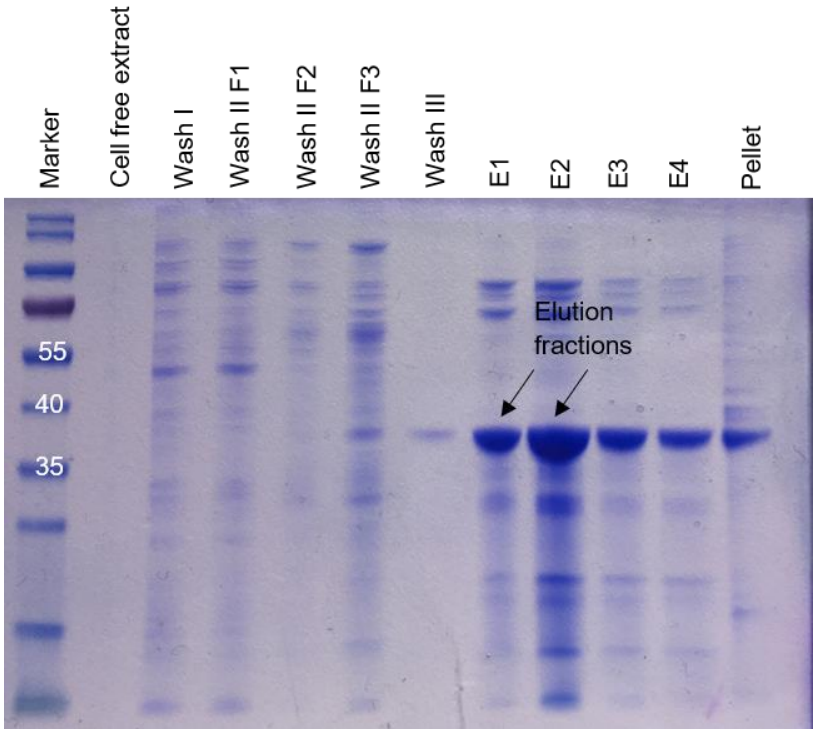

Figure 28: Purification fractions of AvLDO analyzed by SDS-PAGE. The first and second elution fraction contained the majority of the desired protein and can be seen at the expected height of 33 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

2.6 Activity assay with AvLDO

Enzymatic reactions with AvLDO were performed with CFE and purified enzyme evaluated as described in 5.5.6.1. Reactions performed with CFE showed product formation to 5-hydroxyleucine after 18 hours, but full conversion was not reached within 24 hours (Figure 29). For biotransformations set up with 0.25 mg/mL purified protein, product formation was detectable after two hours and went up to full conversion within 24 hours (Figure 30 [C, D]). When using 0.5 mg/mL purified enzyme, full conversion was reached already after 18 hours (Figure 30 [A, B]). Identical results were obtained for reactions performed with the addition of DTT.

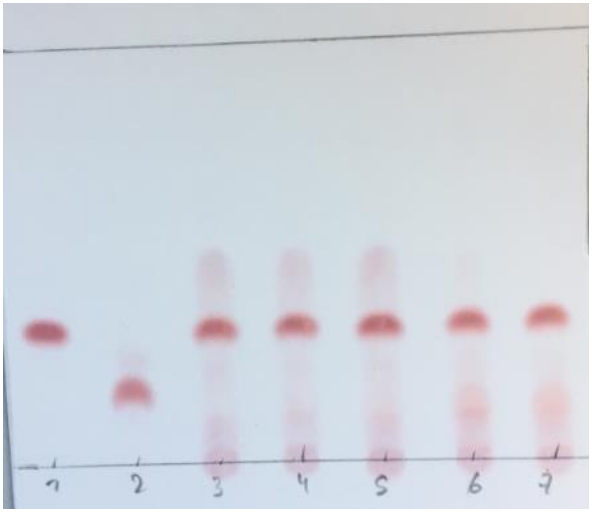

Figure 29: Biotransformations of AvLDO with L-leucine as a substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE. L-leucine was used as a standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, CFE: 1.65 mg/mL in 500 µL volume at 25 °C; [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18}; [7] t_{24}.

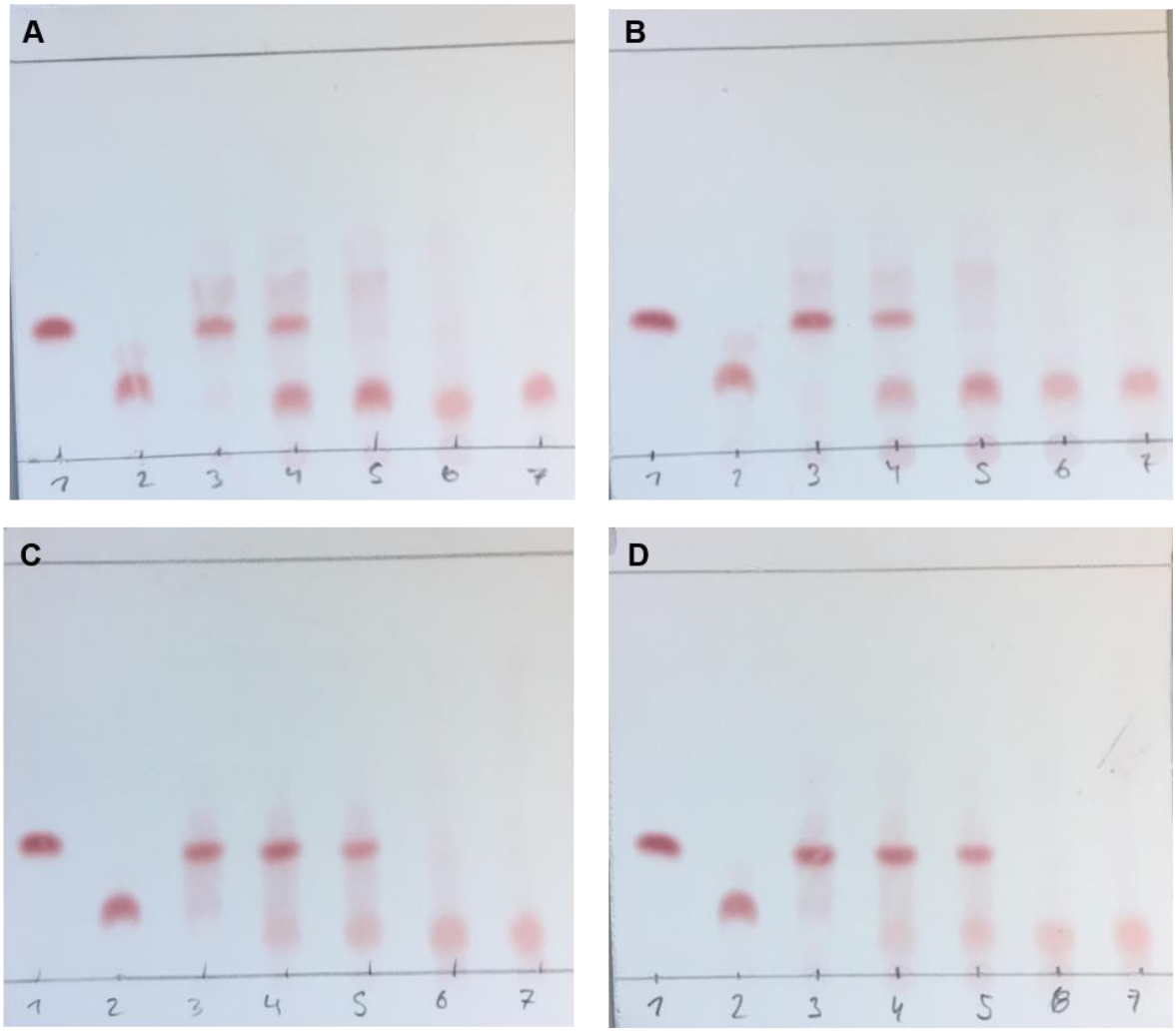

Figure 30: Biotransformations of AvLDO with L-leucine as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with 0.5 mg/mL **[A, B]** and 0.25 mg/mL **[C, D]** purified protein. L-leucine and hydroxy isoleucine were used as standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A C]** with DTT **[B, D]** without DTT in 500 µL volume at 25 °C. **[A]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18}; [7] t_{24}; **[B]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18}; [7] t_{24}; **[C]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18}; [7] t_{24}; **[D]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18} ;[7] t_{24}.

Enzymatic reactions performed with purified enzyme and α-methyl-L-leucine as substrate showed no activity under the implemented conditions, only byproduct formation was visible (Figure 31). This was also the case for biotransformations employing L-isoleucine (see appendix 7.7).

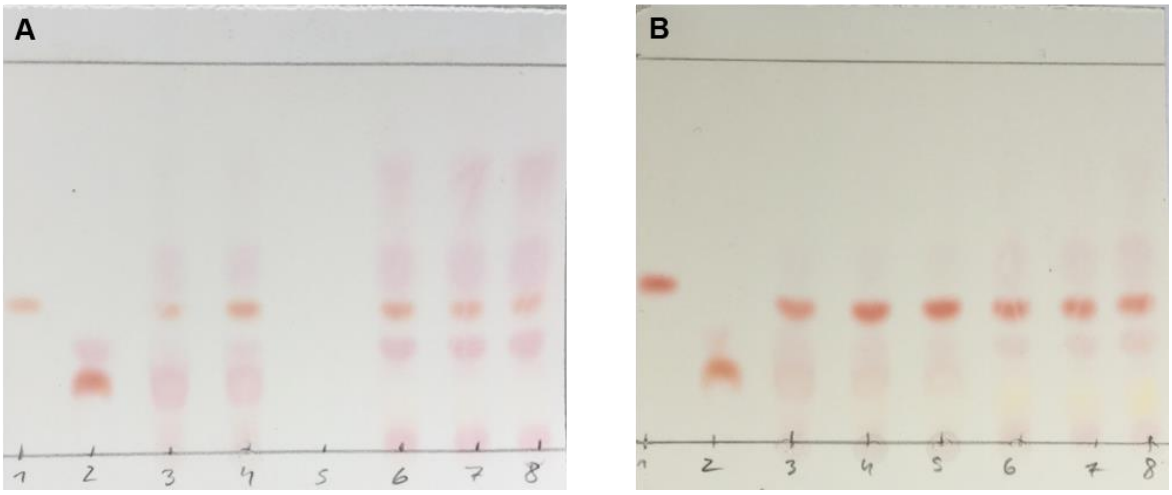

Figure 31: Biotransformations of AvLDO with α-methyl-L-leucine as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with purified protein. α-methylleucine and hydroxy isoleucine were used as standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A]** purified protein: 0.25 mg/mL **[B]** purified protein: 0.5 mg/mL in 500 µL volume at 25°C. **[A]** [1] α-methylleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] nc; [6] t_{18}; [7] t_{24}; [8] t_{48} **[B]** [1] α-methylleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18};[7] t_{24}; [8] t_{48}

2.7 Assessment and determination of the transformation efficiency of *E. coli* 3Δs*ucA*

For the determination of the transformation efficiency of *E. coli* 3Δs*ucA*, chemo-competent cells were initially transformed with plasmid DNA of pET28a(+) as described in 5.4.3. Cells were regenerated for one, two and four hour(s) and afterwards 50 µL, 100 µL and the rest of the cells were streaked out on agar plates (Figure 32). Transformation efficiency was determined with equation (1). The highest transformation efficiency of $5.7x10^3$ cfu/µg was obtained after two hours of regeneration.

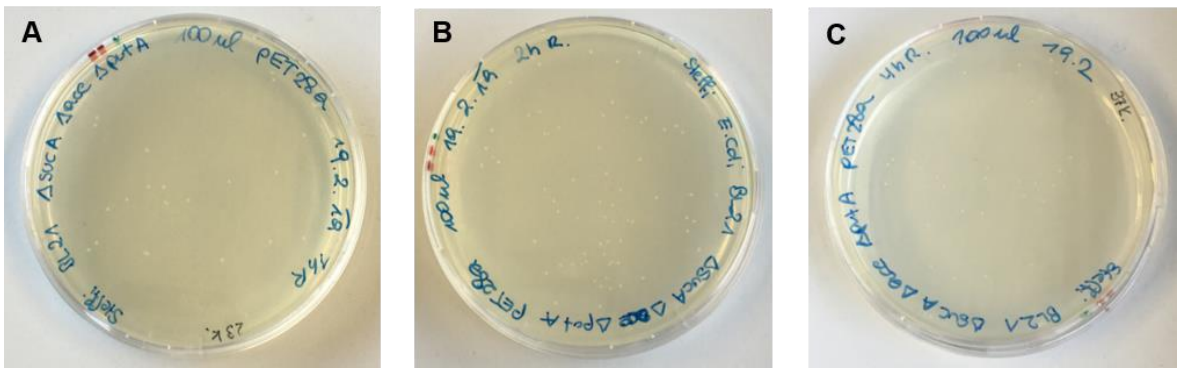

Figure 32: Assessment and determination of the transformation efficiency for *E. coli* 3Δs*ucA*. 100 µL of cells were streaked out on LB agar plates [A] one hour of regeneration with 23 colonies; [B] two hours of regeneration with 85 colonies; [C] four hours of regeneration with 37 colonies.

2.8 Construction of empty vector pET22b(+)

Deletion of the BtIDO gene from the pET22b(+) expression construct was implemented via PCR-mediated plasmid DNA deletion method (Hansson et al., 2008), whereby the backbone of the pET22b_BtIDO was amplified with primers binding outside of the BtIDO gene. The ~15 bp overlapping ends of each primer have a complementary sequence to the opposite strand of the backbone and therefore recombine resulting in a nicked pET22b(+) vector (Figure 33).

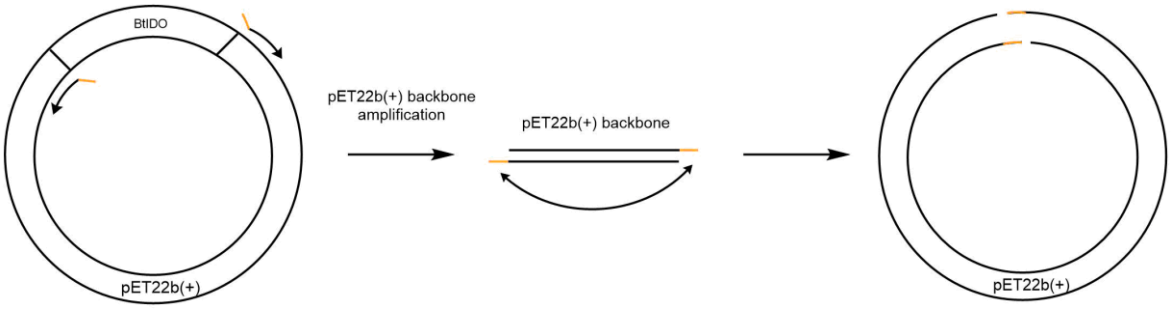

Figure 33: Scheme of establishing the pET22b(+) construct via PCR-mediated plasmid DNA deletion.

After performing gradient PCR, the PCR product was analyzed on a control gel (Figure 34). The expected bands with a size of 5340 bp were obtained for sample 1-3. For sample 4 only a faint band was detected. However, in order to eliminate methylated parental strands, a *Dpn*I digestion was performed followed by transformation of the nicked amplified vector into chemo-competent *E. coli* with the purpose of receiving the desired clones. Four colonies were picked in order to validate the construct by DNA sequencing.

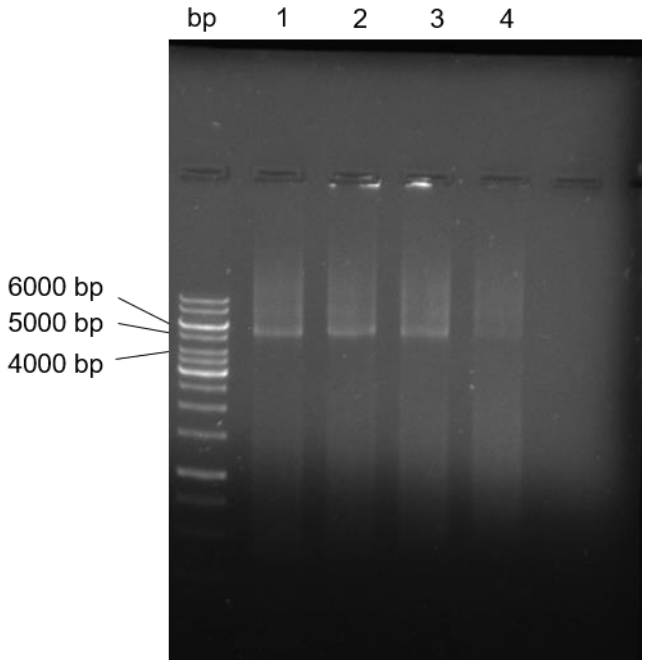

Figure 34: Control gel of gradient PCR for pET22b(+). Bands for [1-3] are visible at the expected size of 5340 bp. Primer annealing was done at temperatures from [1] 64.0 °C;[2] 65.7 °C;[3] 67.3 °C and [4] 69.0 °C.

2.9 Selection of active α-ketoglutarate dependent dioxygenases using *E. coli* 3Δ*sucA* – growth studies

2.9.1 Agar plates

Restoration of the broken TCA cycle was implemented through the hydroxylating activity of GriE, BtIDO and AvLDO. Initially, selection of active α-ketoglutarate dependent dioxygenases was tested out on agar plates. A transformation of *E. coli* 3Δ*sucA* was performed with the constructs pET28a_AvH, pET22b_BtIDO and pET28a_GriE and the empty vector pET28a(+), which served as negative control. After cells were regenerated in SOC medium, they were washed with minimal medium to get rid of any nutrients that may restore cell growth. Colonies were visible on each plate including the negative control (pictures not available).

2.9.2 Liquid medium

For better monitoring of the growth behavior, selection was monitored in liquid medium. In addition to that, the growth behavior of *E. coli* BL21(DE3) and the mutant strain was compared.

2.9.2.1 Growth studies with BtIDO

As expected, growth of the *E. coli* strain BL21(DE3) carrying pET22b(+) was visible for all approaches. Cells which were cultured in LB medium showed growth after about 17 hours and in minimal medium after about 40 hours. For the mutant strain, cell growth was restored earlier when inoculated in LB medium compared to minimal medium (Figure 35 (a-d)). Growth of the empty vector control was visible after 24 hours and for *E. coli* 3Δ*sucA* carrying BtIDO after 40 hours (Figure 35 (a)).

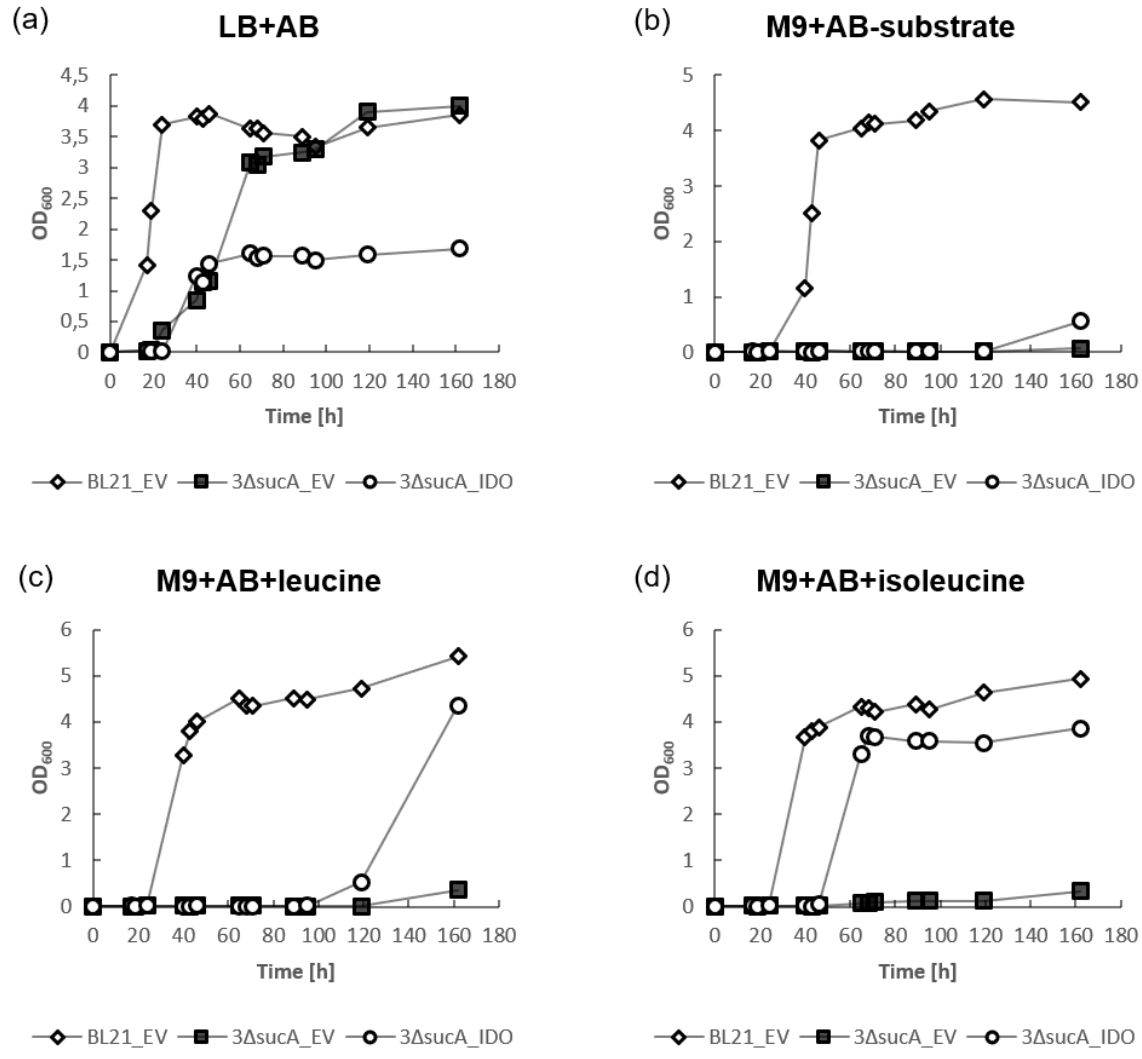

Figure 35: Growth studies with BtIDO. Comparison of the growth behavior of *E. coli* BL21(DE3) carrying pET22b(+) and *E. coli* 3Δ*sucA* carrying pET22b_BtIDO or pET22b(+). Recombinant strains were cultured in different kind of media: **(a)** LB medium supplemented with ampicillin **(b)** minimal medium supplemented with ampicillin but without the addition of a substrate **(c)** minimal medium containing ampicillin and L-leucine **(d)** minimal medium with the addition of ampicillin and L-isoleucine. In all cases, *E. coli* BL21(DE3) showed a quicker growth compared to the mutant strain, as expected. For the mutant strain, a cell growth was earlier restored when inoculated in LB medium compared to minimal medium (a-d). In approach (b), growth of the mutant strain harboring BtIDO was observed after about 7 days (162 hours). Growth of *E. coli* 3Δ*sucA* was restored by the hydroxylating activity of BtIDO after 119 hours (c) and 65 hours (d). For approaches (c) and (d) unexpected growth was observed for the empty vector controls.

However, growth of the mutant strain harboring BtIDO was even observed without external substrate addition after about 7 days (162 hours) (Figure 35 (b)). In the presence of L-isoleucine, growth of *E. coli* 3Δ*sucA* was visible after 65 hours and for the empty vector control after 71 hours (Figure 35 (c)). In case of L-leucine a delayed growth after about five days (119 hours) was obtained which coincide with Hibi *et al.*(Hibi *et al*., 2011) where they demonstrated a higher substrate specificity of BtIDO towards L-isoleucine than for L-leucine. Besides, unexpected growth was observed for the empty vector control after about seven days (162 hours) (Figure 35 (d)).

2.9.2.2 Growth studies with GriE

For *E. coli* BL21(DE3) harboring pET28a(+), growth was observed after 17 hours when cultivated in LB medium and in minimal medium after about 41 hours (Figure 36 (a)). Cell growth was again earlier observed for *E. coli* 3Δ*sucA* when set up in LB medium than in minimal medium. Growth was already visible after about 41 hours (Figure 36 (a)). For both mutant strains carrying either the empty vector pET28a(+) or pET28a_GriE, growth was even observed without the addition of L-leucine or L-isoleucine after about five days (120 hours) (Figure 36 (b)). In case when L-leucine was added to the minimal medium, hydroxylating activity of GriE restored growth of the mutant strain after about 65 hours. However, the empty vector control also showed growth after six days (144 hours) (Figure 36 (c)). Growth of *E. coli* 3Δ*sucA* was also observed in the presence of L-isoleucine hydroxylation after six days (144 hours). Again, unexpected growth was visible for the empty vector control (Figure 36 (d)).

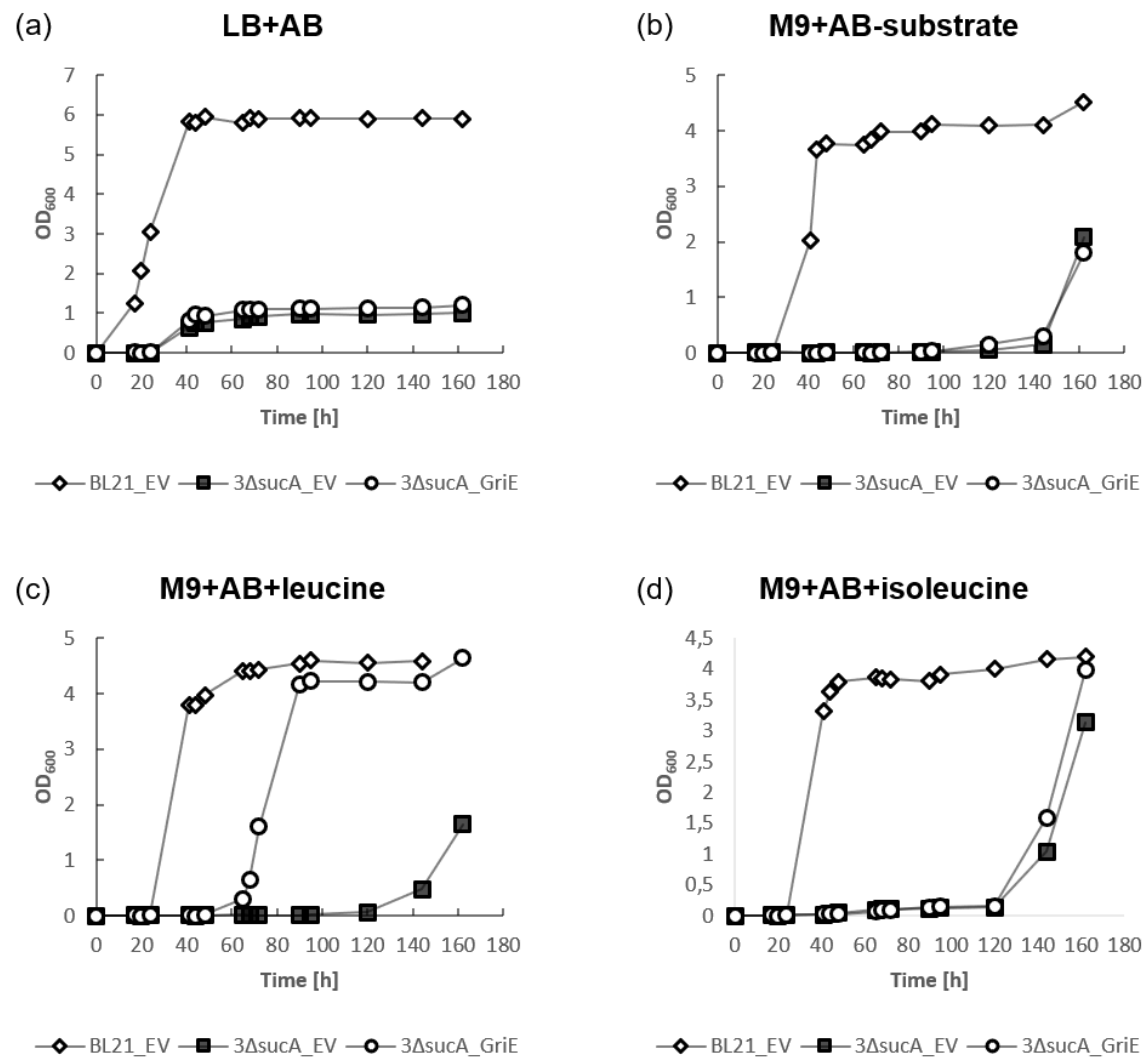

Figure 36: Growth studies with GriE. Comparison of the growth behavior of *E. coli* BL21(DE3) carrying pET28a(+) and *E. coli* 3Δ*sucA* carrying pET28a_GriE or pET28a(+). Recombinant strains were cultured in different kind of media: **(a)** LB medium supplemented with kanamycin **(b)** minimal medium supplemented with kanamycin but without the addition of a substrate. **(c)** minimal medium containing kanamycin and L-leucine **(d)** minimal medium with kanamycin and L-isoleucine. For *E. coli* BL21(DE3), higher growth rates were observed compared to the mutant strain. For *E. coli* 3Δ*sucA*, growth was faster restored when inoculated in LB medium compared to minimal medium (a-d). Both mutant strains harboring either pET_GriE or pET28a(+) showed growth after about five days (120 hours) (b). The introduction of GriE restored the TCA cycle and growth of the mutant strain in minimal medium was observed 65 hours (c) and 144 hours (d). Figure (c) and (d) show an unexpected growth of the empty vector controls.

2.9.2.3 Growth studies with AvLDO

The *E. coli* strain BL21(DE3) carrying pET28a(+), which was cultured in LB medium showed growth after about 19 hours and in minimal medium after about 42 hours (Figure 37 (a)). *E. coli* 3Δ*sucA* cultures, which were set up in LB medium showed a faster cell growth restoration compared to minimal medium, as expected. Growth was observed already after 42 hours (Figure 37 (a)). However, the mutant strain carrying AvLDO was able to grow in minimal medium in the absence of AvLDO

activity after about seven days (161 hours). For the empty vector control, on the other hand, cell growth was not visible (Figure 37 (b)).

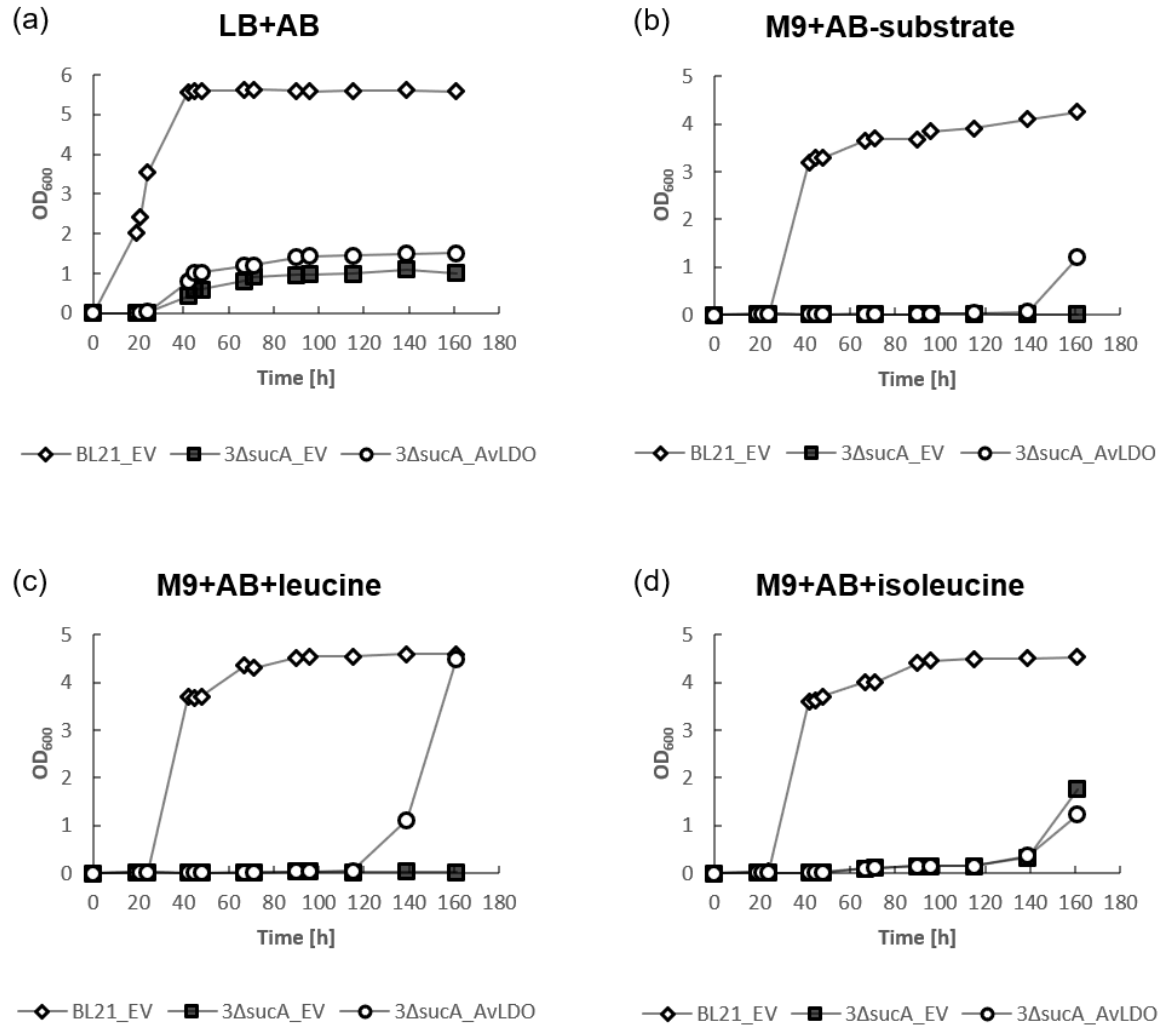

Figure 37: Growth studies with AvLDO. Comparison of the growth behavior of *E. coli* BL21(DE3) carrying pET28a(+) and *E. coli* 3Δ*sucA* carrying pET28a_AvLDO or pET28a(+). Recombinant strains were cultures in different kinds of media **(a)** LB medium supplemented with kanamycin **(b)** minimal medium containing kanamycin but no substrate. **(c)** minimal medium supplemented with kanamycin and L-leucine **(d)** minimal medium with kanamycin and L-isoleucine. For all approaches, *E. coli* BL21(DE3) showed a higher growth rate compared to the mutant strain and for the mutant strain cell growth occurred faster when inoculated in LB medium compared to minimal medium (a-d). *E. coli* 3Δ*sucA* carrying AvLDO showed growth after about seven days (162 hours) (b). Growth of *E. coli* 3Δ*sucA* was restored by the hydroxylating activity of AvLDO after 139 hours (c) and 71 hours (d). Unexpected growth was observed for the empty vector controls (c,d).

The introduction of AvLDO restored the TCA cycle and growth of the mutant strain in minimal medium was observed after about six days (139 hours (Figure 37 (c)). Interestingly, growth was even observed with the addition of L-isoleucine after about 71 hours, although Correia Cordeiro *et al.* (Correia Cordeiro et al., 2018) stated that AvLDO is inactive towards L-isoleucine (Figure 37 (d)). Overall, restoration of cell

growth occurred earlier in case of L-isoleucine than compared with L-leucine. The empty vector control showed unexpected growth after about 67 hours (Figure 37 (d)).

3 Discussion

Non-heme Fe(II)/α-ketoglutarate dependent oxygenases, responsible for the CH-activation of small molecules, play an important role in the production of optically pure hydroxy amino acids which serve as precious building blocks in various industries. Due to the high regio- and stereoselectivities non-heme Fe(II)/α-KG-dependent oxygenases can perform and their great substrate acceptance, these enzymes find increasing interest in industrial applications and research projects (Hüttel, 2013; Peters and Buller, 2019). The aim of this master's thesis can be divided into three sections comprising a) the establishment of an Fe(II)/α-ketoglutarate dependent dioxygenase toolbox including the characterization and screening of in total eleven dioxygenases, b) optimization of soluble expression levels of AvLDO and c) the implementation of a modified *E. coli* strain for the selection of active α-ketoglutarate dependent dioxygenases.

3.1 Expression of different dioxygenases

Initially, the heterologous expression of eleven dioxygenase constructs was investigated in *E. coli* BL21(DE3) based on descriptions from literature (Table 2) with some adaptions in expression conditions. Whole cell profiles of each dioxygenase were analyzed by SDS-PAGE and revealed overall satisfying expression levels, whereas for KDO1, BtIDO and MfIDO no bands of the protein of interest were obtained. Moreover, soluble and insoluble protein fraction were analyzed and SDS-PAGE analysis demonstrated that most of the dioxygenases were expressed in a soluble form in varying expression levels. The best expression levels were obtained for GriE after an overnight expression and for AnPip4H after 4 h of expression. For SadA, TmDO and MMGDO, lower levels of soluble protein were visible on the SDS gel. In case of SadA, the highest amount of soluble enzyme was received during an expression of four hours at 28 °C. In previous works, expression of the recombinant dioxygenase was rather preferred with the expression vector pQE80 in the host organism *E. coli* JM109 or *E. coli* Rosetta(DE3) cultivating cells in LB medium overnight at 28 °C (Hibi et al., 2012; Qin et al., 2013). To optimize soluble expression, different *E. coli* strains such as *E. coli* JM109(DE3) or *E. coli* Rosetta(DE3), other expression vectors and an adaption in IPTG concentrations should be considered. The putative dioxygenase TmDO showed satisfying expression levels after 4 h of expression at 20 °C and for MMGDO the highest amount of soluble protein was observed after an overnight expression at

30 °C. Expression levels could be increased again by investigating different expression vectors, host organisms or IPTG concentrations.

No visible bands or bands with a low intensity were obtained for the rest of the examined dioxygenases including AvLDO, BtIDO, MflDO, FoPip4H and KDO1. Nevertheless, in case of BtIDO, MflDO and FoPip4H, soluble protein expression was verified by purification or evidence of activity. Reasons for non-detected bands might be errors during pipetting or adapted conditions in expression. Instead of overexpressing KDO1 at 20 °C overnight, an implementation as documented in Baud *et al*. (Baud et al., 2014) should be investigated by cultivating cells in TB medium supplemented with 0.5 M sorbitol and 5 mM betaine until an OD_{600} of 2 is reached. Furthermore, expression of AvLDO was implemented based on the documented conditions by Meisborn (Meisborn, 2014) at 30°C overnight but resulted only in low levels of soluble protein. Instead, expression conditions by Correia Cordeiro *et al* (Correia Cordeiro et al., 2018) should be considered performing heterologous expression in *E. coli* BL21(DE3) at 20 °C overnight.

3.2 Protein purification

The CFE of six dioxygenases was subjected to affinity chromatography and SDS-PAGE analysis confirmed that all dioxygenases were successfully purified yielding high amounts of protein in the elution fractions. However, minor problems remain for instance, in case of GriE and SadA the majority of the desired protein was already detached from the column during the washing steps. A significant amount of protein was present in the flow-through indicating that insufficient nickel resin was applied on the column on which the protein of interest could bind. An adaption of the imidazole concentration and a higher amount of column material should be taken into account in order to minimize the loss of desired protein. In case of Fopip4H, AnPip4h, AvLDO and BtIDO, low amounts of protein were lost during washing steps. Additionally, it was observed that elution fractions of BtIDO and FoPip4H were not completely pure due to unknown proteins, which bound unspecifically to the column material and were detached during elution. A further reason for this issue is that the column material was washed insufficiently, as the last washing step (wash II F3) still showed *E. coli* protein background. Adjusting the imidazole concentration during the washing steps could reduce these problems.

3.3 Activity assay

A screening of in total eleven dioxygenases was examined in order to reveal the substrate acceptance of these enzymes. Among the investigated enzymes, four (BtIDO, GriE, MfIDO and FoPip4H) showed product formation. Initially, it is important to mention that enzymatic reactions were performed with unoptimized conditions based on the reaction conditions by Meisborn (Meisborn, 2014) including 50 mM buffer, 5 mM substrate, 7.5 mM α-ketoglutarate, 0.5 mM $FeSO_4x7H_2O$ and 5 mM ascorbate. Moreover, in most cases the authentic product standard was not available leading to inconclusive results. However, the substrate conversion observed on TLC was verified by the depletion of the substrate spot resulting in the formation of a product spot. Occasionally, unspecific byproducts were obtained which were not further characterized for reasons of time, but it is assumed that these products occur as the prepared reaction mixture is of complex composition.

Only for GriE, product formation (full conversion) was obtained for both CFE and purified enzyme within 24 hours. Biotransformations performed with CFE of BtIDO, FoPip4H and MflDO showed product formation after one or two hours of reaction time. However, full conversion could not be detected in the investigated period. Since enzyme activity was detected in the CFEs of BtIDO and FoPip4H, it can be assumed that the implemented purification conditions were unsuitable for the proteins and rendered them inactive due to un- or misfolding. The folding state of the protein could be checked with differential scanning fluorimetry in order to monitor the folding state of the protein in different buffers and purification conditions (Niesen et al., 2007; Vivoli et al., 2014). Instead of using 50 mM potassium phosphate buffer for cell lysis 20 mM Tris-HCl buffer (pH 7) might be preferable. Phosphates are able to chelate bivalent metals, such as Fe(II), leading to precipitation and might therefore render the activity of the enzyme (Enoki et al., 2016; Good and Izawa, 1972; Hibi et al., 2012, 2016; Lukat et al., 2017).

For the rest of the dioxygenases, product formation was either not detectable or not conclusive. For instance, in case of SadA the substrate N-succinyl-L-valine was not detectable because the substrate lacks a free amine group and is therefore inaccessible for ninhydrin. Bromocresol green was used instead as a pH indicator to detect acid groups, but still no activity towards N-succinyl-L-valine was observed. Instead, other methods such as NMR or HPLC are recommended to confirm product

formation; or the use of another branched L-amino acid. For example, Hibi *et al.* (Hibi et al., 2012) successfully observed via NMR the conversion of N-succinyl L-leucine into N-succinyl L-threo-β-hydroxyleucine with high diastereoselectivity. KDO1 was probably unreactive towards L-lysine. However, several problems can be a reason for wrong interpretation of the result. Based on the SDS-PAGE results, no or very low amounts of soluble protein were detected. If any soluble protein was obtainend, biotransformations were set up with a too low concentration (0.12 mg/mL) to detect any conversion with TLC. Furthermore, the TLC results were not conclusive since no authentic product standard was available. Moreover, spots were not detectable with the implemented mobile phase due to the high polarity of hydroxy lysine. Besides optimizing expression conditions, more reliable analytics such as HPLC should be considered. The substrate specificity of the putative dioxygenases TmDO and MMGDO is still unknown (Meisborn, 2014). Biotransformations performed with L-leucine and L-methionine revealed no product formation under the implemented conditions. The substrate scope of both enzymes needs therefore further investigation. When analyzing enzymatic reactions of P4H with TLC, detected compounds changed in color from yellow to pink. It is not conclusive whether a conversion was established since an authentic product standard was not available.

The reason for the inactivity of AnPip4H remains unknown. Significant expression levels of soluble protein were observed which leads to the assumption that the implemented purification conditions were damaging the protein and the adaptation of reaction conditions might cause AnPip4H to be inactive. Again, it is recommended that the folding state of the protein is verified with differential scanning fluorimetry in order to monitor the folding state of the protein. Furthermore, employing the exact conditions from Hibi *et al*. (Hibi et al., 2016) might diminish this issue. They documented reaction conditions in 50 mM MES buffer (pH 6.5) including 10 mM L-Pip, 15 mM, α-KG, 0.5 mM $FeSO_4x7H_2O$, 5 mM ascorbate and 1 mM DTT implemented at 20 °C. Biotransformations performed with AvLDO revealed no activity towards L-leucine. This can be explained by the insufficient amount of soluble protein obtained during heterologous expression. Furthermore, lysis and protein purification conditions were unoptimized and might have damaged the protein.

3.4 Optimization in expression, purification and reaction conditions of AvLDO

Due to low soluble expression levels of AvLDO, which were obtained with the performed conditions, expression studies were implemented based on conditions stated by Correia Cordeiro *et al* (Correia Cordeiro et al., 2018). The highest amount of soluble protein was detected after 6 h of expression. The duration of the expression should be kept under 24 hours as a depletion of AvLDO was observed. In addition, different IPTG concentrations and cultivation temperatures were investigated and revealed no effects on improving soluble expression levels. The majority of the desired protein was detected as inclusion bodies. In order to overcome this issue, co-expression of the chaperonin GroEL and its cofactor GroES was examined. Chaperonins are known for assisting the protein folding machinery by preventing the misfolding and aggregation of non-native polypeptides. Therefore, GroEL binds to non-native proteins through hydrophobic residues and in the cavity of GroES correct folding of the protein is facilitated through ATP hydrolysis (Hayer-Hartl et al., 2016; Xu and Sigler, 1998). Expression of AvLDO and the chaperone proteins GroEL/ES was not detectible in both the insoluble fraction as well as the soluble fraction. Furthermore, auto-induced protein expression of AvLDO was investigated in the host organisms *E. coli* BL21(DE3) and *E. coli* JM109(DE3). Autoinduction leads in several cases to a tighter controlled protein induction, higher cell densities and soluble expression levels compared to an expression induced with IPTG (Jia and Jeon, 2016). Here, in both cases the amount of soluble protein was not increased.

Besides all the implemented approaches to improve soluble expression, several other factors such as the absence of disulfide bonds or the redox potential of the cytoplasm plays an important role for the correct folding of the proteins. Therefore, host strains with thioredoxin reductase deficiency and fusion proteins find great utilization in improving folding and solubilization of the desired protein. However, recombinant protein which aggregated into inclusion bodies can be recovered and refolded into its bioactive form under mild conditions in a four-step procedure. The process itself is rather complex and time consuming and therefore not the number one choice (Hannig and Makrides, 1998; Jia and Jeon, 2016; Kaur et al., 2018). Nevertheless, due to the fact that the expression implemented with the conditions from Correia Cordeiro *et al*. (Correia Cordeiro et al., 2018) revealed the best results,

the obtained cell pellet was employed in order to set up enzymatic reactions. Interestingly, this time, AvLDO showed activity towards L-leucine and product formation (full conversion) was obtained for both CFE and purified enzyme within 24 hours. Several adaptions were implemented during the procedure including the resuspension of the cell pellet was dissolved in 20 mM Tris-HCl buffer (pH 7.4) instead of 50 mM potassium phosphate buffer (pH 7.4). Moreover, sonication was implemented at lower intensity settings and besides stock solutions for $FeSO_4x7H_2O$ and ascorbate were prepared right before use and not frozen in order to prevent oxidation. It is important to mention that the obtained results are not completely reproducible, as procedures need further optimization. For example, cell lysis should be implemented with non-mechanical disruption methods such as chemical or biological lysis as they perform a more gentle and quantitative disruption than sonication (Shehadul Islam et al., 2017). Optimized purification conditions yielded high amounts of desired protein during elution. However, to increase sample purity and maintain biological activity fast protein liquid chromatography is recommended such as the automated purification system ÄKTA (Chapman, 2005; Sigrell et al., 2003).

Furthermore, in case of AvLDO and GriE, enzymatic reactions were employed with α-methyl-L-leucine with the aim to identify substrate tolerance towards α-methyl amino acids. Biotransformations were performed with these enzymes as they employ a similar hydroxylation activity towards L-leucine which additionally serve as a positive control. Both enzymes showed no activity towards α-methyl-L-leucine under the implemented conditions, only byproduct formation was visible. However, these results should encourage investigating further enzymes as the hydroxylation of α-methyl amino acids find tremendous interest providing several appealing chemical building blocks.

3.5 Selection of active α-ketoglutarate dependent dioxygenases

Growth of the *E. coli* 3Δ*sucA* mutant could be restored with TCA re-cyclization via dioxygenase catalyzed L-amino acid hydroxylation, therefore selecting active α-ketoglutarate dependent hydroxylases. In order to illustrate the impact that scarcity of succinyl-CoA and the absence of a glyoxylate shunt might cause, the growth behavior of the mutant strain and *E. coli* BL21(DE3) was investigated. Growth of *E. coli* BL21(DE3) showed a faster restoration of cell growth compared to *E. coli*

3Δ*sucA* in both LB medium and minimal medium. This can be explained by the fact, that succinyl-CoA plays an important role in the biosynthesis of methionine, lysine and diaminopimelate (DAP), which are employed in cell wall formation. Therefore, a reduced succinyl-CoA pool combined with the absence of a glyoxylate shunt invokes a reduced cell size, longer growth rates and overall lower biomass rates (Theodosiou et al., 2017). Studies of Theodosiou *et al*. (Theodosiou et al., 2017) demonstrated that a supplementation of DAP, lysine and methionine has no influence on the growth rates or biocatalytic efficiency but result in 16% increased cell size.

Growth studies with BtIDO, GriE and AvLDO were investigated and revealed that the hydroxylating activity of each dioxygenase is able to shunt the TCA cycle and therefore restore growth of the mutant strain. A survival of the mutant strains in minimal medium is only guarantied if the strain possesses the ability to produce α-ketoglutarate dependent dioxygenases and hydroxylate the favored substrate. In case of BtIDO, cell growth of *E. coli* 3Δ*sucA* (pET22b_BtIDO) was visible at an earlier stage when cultivated in L-isoleucine containing minimal medium than with L-leucine. The higher specific activity towards L-isoleucine coincide with results of Hibi *et al*. (Hibi et al., 2011) documenting a specific activity of 0.71 ± 0.29 U/mg for L-isoleucine and 0.31 ± 0.04 U/mg for L-leucine. For GriE and AvLDO, restoration of the TCA cycle occurred faster when the mutant strain was cultivated in L-leucine compared to L-isoleucine. Although both enzymes are unreactive towards L-isoleucine cell growth was restored (Correia Cordeiro et al., 2018; Zwick and Renata, 2018a).

However, it was examined that cell growth of *E. coli* 3Δ*sucA* carrying the dioxygenase gene was even restored in minimal medium without the external addition of a substrate. The same phenomenon was also documented by Theodosiou *et al*. (Theodosiou et al., 2017) explaining that the endogenous substrate hydroxylation rate might be sufficient to shunt the TCA and restore cell growth of the mutant strain. Further reasons for visible cell growth may be the biosynthesis pathways of L-leucine and L-isoleucine in *E. coli*, which both use pyruvate as a precursor, and thus might be retrieved from the organism before entering the (broken) TCA cycle. Synthesis of those amino acids is therefore independent from the TCA cycle (Amorim Franco and Blanchard, 2017). Furthermore, Li *et al*. (Li et al., 2014) discovered that *E. coli* is also capable of

generating high levels of enzymes for the biosynthesis of amino acids when cultured in minimal medium. This leads to the assumption, that the rather low activity of the dioxygenases is indeed able to close the TCA cycle by hydroxylating L-leucine or L-isoleucine, which is provided by the organism itself and thus assures growth of *E. coli* 3Δ*sucA*. In this case, observed growth is independent from the substrate supplied in the medium. Moreover, *E. coli* 3Δ*sucA* produces acetate when cultured on fermentable sugars, which could be additionally employed as sole carbon source besides glucose. Although the mutant strain has a glyoxylate deficiency and therefore cannot assimilate acetate, the catabolic pathway via TCA cycle could lead to a formation of α-ketoglutarate. The so-generated α-ketoglutarate could function as cosubstrate during substrate hydroxylation and shunt the broken TCA cycle (Kornberg, 1966; Paliy and Gunasekera, 2007; Theodosiou et al., 2017).

In most cases unexpected growth of *E. coli* 3Δ*sucA* bearing the empty vector pET22b(+) or pET28a(+) was observed. There are a number of reasons responsible for this issue. For instance, succinate could be synthesized through other side pathways leading to restoration of the TCA cycle and visible cell growth. Studies reported that *E. coli* 3Δ*sucA* mutants excrete little amounts of L-glutamate which could then be converted via aminobutyrate catabolic pathway into succinate (Li et al., 2006; Zhang et al., 2018). Another pathway for succinate synthesis is the reductive branch of the TCA cycle under anaerobe conditions (Thakker et al., 2012; Zhang et al., 2018; Zhu and Tang, 2017). It is also important to point out that the risk of contaminations during handling cannot be excluded. Furthermore, the inactivation of ampicillin through hydrolysis of the β-lactam ring catalyzed by *E. coli* secreted β-lactamases and the instability of ampicillin and kanamycin at longer cultivation lead to decrease of the antibiotic concentration over time and elevate the risk for contaminations (Briñas et al., 2002).

For further investigations in the future, several factors should be considered. First, bacterial cultures should be supplemented with a higher antibiotic concentration and additionally supplemented during the procedure in order to minimize the risk of contaminations. Moreover, optical cell density sensors or cultivating cells in microtiter plates instead of shake flasks could also reduce this issue. As growth studies were implemented at unoptimized conditions it remains unknown if the desired expression levels were reached in order to couple the hydroxylation activity of the hydroxylases with cell growth. It is recommended to pursue the expression

levels via SDS-PAGE analysis and employ the corresponding expression conditions for each dioxygenase. Additionally, main cultures could be inoculated at a specific optical density to ensure equal cell densities in each culture. Besides, the duration of the cultivation of main cultures should be increased until cultures reach stationary phase as the mutant strain might employ other carbon sources and reveal an ability for diauxic growth. Furthermore, as L-leucine and L-isoleucine are both provided by the organism itself substrate dependent growth can be excluded. In this case, a different substrate scope should be tested out.

4 Conclusion

This study provided insight in the establishment of a non-heme Fe(II)/α-ketoglutarate dependent oxygenase toolbox investigating expression levels and screening the substrate scope of eleven dioxygenases. Ten of the dioxygenases were successfully expressed in *E. coli* BL21(DE3), with varying levels of soluble protein depending on the enzyme. Among the investigated dioxygenases, four showed substrate consumption including BtIDO, GriE, FoPip4H and MfIDO. Enzymatic reaction performed with GriE revealed conversion of L-leucine to 5-hydroxyleucine for both CFE and purified enzyme within 24 hours. In case of BtIDO, MfIDO and FoPip4H product formation was observed after one or two hours of reaction time when employing reactions with CFE. However, biotransformations with purified enzyme were unreactive towards the tested substrates as the implemented purification conditions might have damaged the protein. Furthermore, expression studies were performed in order to improve soluble expression levels of AvLDO. Several approaches were tested out such as the co-expression of chaperones GroEL/ES, inducing bacterial cultures with different IPTG concentrations, lowering the cultivation temperature and auto-induced protein expression in autoinduction medium. None of the approaches led to an expression optimization of AvLDO. The best results were obtained when heterologous expression of AvLDO was performed based on the conditions by Correia Cordeiro *et al*. (Correia Cordeiro et al., 2018). A 6 h expression revealed the highest amount of desired protein. Enzymatic reactions with purified protein showed conversion of L-leucine to 5-hydroxyleucine within 24 hours, but were inactive towards L-isoleucine and α-methyl-L-leucine. Furthermore, the implementation of an engineered *E. coli* strain was investigated in order to generate a selection tool for active α-ketoglutarate dependent dioxygenases. Additionally, the growth behavior of the mutant strain *E. coli* 3Δ*sucA* was compared with *E. coli* BL21(DE3). Growth studies with BtIDO, GriE and AvLDO revealed that the hydroxylating activity of the dioxygenases shunts the TCA cycle by coupling the hydroxylation of a substrate and the decarboxylation of α-ketoglutarate into succinate, which eventually restores cell growth. Nevertheless, cell growth was also observed for *E. coli* 3Δ*sucA* carrying the gene for the recombinant dioxygenase in minimal medium without substrate addition due to the endogenous hydroxylating activity. Succinate production via the γ-aminobutyrate catalytic pathway employing acetate as a further carbon source might

have led to restoration of the broken TCA cycle resulting in an unexpected growth of *E. coli* 3Δ*sucA* bearing the empty vector. Moreover, due to the instability of ampicillin and kanamycin at the implemented cultivation temperatures the required antibiotic concentrations were not existent and enhanced the risk of contaminations.

5 Materials and Methods

5.1 Chemicals and Labware

All chemicals and labware used in this thesis are listed in the appendix (see 7.1). Unless otherwise stated, standard laboratory devices were used for the experiments performed in this thesis.

5.2 Enzymes used in this thesis

All dioxygenases used in this study are listed below in Table 2.

Table 2: Overview of all dioxygenases used in this study; 'RUB' = Ruhr Universität Bochum.

Enzyme	Short name	Literature	Source
Leucine dioxygenase from *A. variabilis*	AvLDO	Meisborn 2014	RUB
Isoleucine dioxygenase from *B. thuringiensis*	BtIDO	Kodera 2009, Hibi 2011, Enoki 2016	RUB
Putative dioxygenase from marine metagenome	MMGDO	-	RUB
Leucine dioxygenase from *M. flagellatus*	MfIDO	Smirnov 2012	RUB
Putative dioxygenase from *T. mobilis*	TmDO	-	RUB
Leucine dioxygenase from *Streptomyces* strain DSM 40835	GriE	Lukat and Katsuyama 2017	GenScript Biotech B.V. (Netherlands)
Lysine dioxygenase from *C. acidiphila*	KDO1	Baud 2014	GenScript Biotech B.V. (Netherlands)
Proline-4-hydroxylase	P4H	Klein 2011, Falcioni 2013	RUB
N-succinyl amino acid hydroxylase from *B. ambifaria* AMMD	SadA	Hibi 2012	RUB
Cyclic aa hydroxylase from *F. oxysporum*	FoPip4H	Hibi 2016	GenScript Biotech B.V. (Netherlands)
Cyclic amino acid hydroxylase from *A. nidulans* FGSC A4	AnPip4H	Hibi 2016	GenScript Biotech B.V. (Netherlands)

5.3 Microbiological Methods

5.3.1 Strains

All *E. coli* strains used for cloning and selection experiments and for the expression of recombinant dioxygenases are listed in Table 3.

Table 3: List of bacterial strains used in this study.

Strain	Genotype	Use	Source
E. coli JM109(DE3)	endA1, recA1, gyrA96, thi, hsdR17 (r_k^-, m_k^+), relA1, supE44, λ–, Δ(lac-proAB), [F´, traD36, proAB, lacI^qZΔM15], lDE3.	Expression of recombinant proteins	Promega
E. coli TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80(lacZ)ΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG	Cloning experiments	Thermo Fisher Scientific
E. coli BL21(DE3)	F- ompT hsdSB (rBmB-) gal dcm (DE3)	Expression of recombinant proteins	Thermo Fisher Scientific
E. coli BL21 *ΔsucAΔaceAΔputA* (DE3) (pLysS) (3Δ*sucA*)	Deletion of *sucA*, *aceA*, and *putA* encoding a-KG dehydrogenase E1 subunit, isocitrate lyase, and proline dehydrogenase	Selection experiments	Helmholtz centre for environmental research - UFZ, Leipzig

5.3.2 Over-night cultures

Over-night cultures (ONC's) were prepared by supplementing 5-10 mL of LB medium (see appendix 7.3) with the corresponding antibiotics: kanamycin with a final concentration of 40 µg/mL for pET28a(+)- and ampicillin with a final concentration of 100 µg/mL for pET22b(+)-based expression constructs. Inoculation was performed with a single colony or 5 µL of glycerol stock and cultures were incubated overnight at 37 °C and 120 rpm.

5.3.3 Glycerol stocks

Glycerol stocks were prepared by mixing 1 mL of an ONC and 1 mL sterile glycerol (60%) in cryotubes. Afterwards, the cells were stored at -80 °C until further use.

5.4 Molecular biological methods

5.4.1 Plasmid list

All plasmids used in this study are listed below in Table 4.

Table 4: Overview of all plasmids used in this study**.** 'Km' = kanamycin, 'Amp' = ampicillin.

Construct	Plasmid	Resistance marker	Insert	His-tag
1	pET28a(+)	Km	AvLDO	N-term
2	pET22b(+)	Amp	BtIDO	N-term
3	pET28a(+)	Km	MMGDO	N-term
4	pET22b(+)	Amp	MfLDO	N-term
5	pET28a(+)	Km	TmDO	N-term
6	pET28a(+)	Km	GriE	N-term
7	pET22b(+)	Amp	KDO1	N-term
8	pET28a(+)	Km	P4H	N-term
9	pET28a(+)	Km	SadA	N-term
10	pET28a(+)	Km	FoPip4H	N-term
11	pET28a(+)	Km	AnPip4H	N-term
12	pET28a(+)	Km	-	N/C-term
13	pET22b(+)	Amp	-	N/C-term

5.4.2 Preparation of chemical competent *E. coli* cells

For the preparation of chemo-competent cells, 100 mL of LB medium were inoculated with 5 mL of ONC and incubated at 37 °C and 200 rpm. Once an OD_{600} of 0.4- 0.5 was reached, cells were harvested by centrifugation (4.000 g, 4 °C, 15 min). The pellet was resuspended in 30 mL Tfb1 and 3.2 mL 1M $MgCl_2$ followed by an incubation of 15 min on ice. After centrifugation (4.000 g, 4 °C, 10 min) the supernatant was discarded and the pellet was resuspended in 4 mL Tfb2 and incubated for 15 min on ice. Aliquots of 50 µL were shock frozen in liquid nitrogen and stored at -80 °C. All components of the solutions can be seen in the appendix (see 7.3.).

5.4.3 Transformation of chemical competent *E. coli* cells

Plasmid DNA was added to thawed chemo-competent *E. coli* cells and incubated for 30 min on ice. A heat shock at 42 °C for 42 s was performed before adding 400 µL LB-SOC (see appendix 7.3) for regeneration and incubation for 60-90 min at 37 °C. Afterwards, the cells were streaked out on LB plates, containing the corresponding antibiotic for selection of positive transformants. For the

transformation of chemo-competent *E. coli* 3Δ*sucA*, cells were transformed with 50 ng of plasmid DNA via heat shock method. After a regeneration of two hours at 37 °C cells were streaked out on LB plates supplemented with the corresponding antibiotic and incubated at 28 °C until colonies were visible.

5.4.4 Assessment and determination of the transformation efficiency of *E. coli* 3Δ*sucA*

Chemo-competent *E. coli* 3Δ*sucA* were transformed as described in 5.4.3. After a regeneration time of one, two and four hour(s), cells were streaked out on LB agar plates containing the corresponding antibiotic. The transformation efficiency was calculated with equation (1).

$$Transformation\ efficiency\ \left[\frac{cfu}{\mu g}\right] = \frac{Number\ of\ transformants}{Amount\ of\ plasmid\ DNA[\mu g]} \qquad (1)$$

5.4.5 Plasmid isolation

Plasmid isolation was performed with the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) according to the manufacturer's protocol.

5.4.6 DNA sequencing

For validation of a construct, the isolated plasmid was sequenced by Sanger Sequencing (Microsynth AG). Sequencing primers were either available in the in-house primer collection or provided by the sequencing company (see appendix 7.5).

5.4.7 Construction of the empty vector pET22b(+)

The insert of the construct pET22b_BtIDO was deleted based on the PCR-mediated method demonstrated by Hansson *et al.* (Hansson et al., 2008) in order to obtain the empty vector pET22b(+). For that, a gradient PCR was performed as described in Table 5 and Table 6. The T-gradient was adjusted to 64.0 °C, 65.7 °C, 67.3 °C and 69.0 °C. After amplification of the backbone, 3 µL of PCR product was analyzed by gel electrophoresis on a 0.8 % agarose gel in 1 x TAE buffer at 200 Volt and 400 mA for about 35 min. Therefore, 0.5 µL of loading dye (6x) were added to 3 µL PCR product and 5 µL of a 1 kb DNA-ladder (Thermo Fisher Scientific, USA) was used. The 0.8 % agarose gel was supplemented with 5 µL of GelRed® Nucleic Acid

Gel Stain (Biotium Inc.) for visualization of the DNA. Gels were visualized under UV light using the G: BOX F3 gel documentation system (Syngene, United Kingdom). The rest of the PCR product was digested by *Dpn*I in order to remove methylated parental strands. For that, 0.5 µL of *Dpn*I was added to the PCR product and incubated for 2.5 hours at 37 °C followed by an inactivation of the enzyme at 80 °C for 20 min. Afterwards, chemo-competent *E. coli* TOP10 were transformed with 5 µL of PCR mixture via heat shock method (see 5.4.3). Cells were streaked out on LB plates supplemented with ampicillin with a final concentration of 100 µg/mL and incubated at 37 °C. For validation of the construct, four colonies were picked and sequenced (see 5.4.6).

Table 5: Reaction mixture of the gradient PCR.

Components	Final volume 25 µL
Forward primer [10 µM]	0.5 µL
Reverse primer [10 µM]	0.5 µL
5x Phusion HF buffer	5 µL
dNTPs [10 mM each]	0.5 µL
Phusion polymerase [2U/µL]	0.2 µL
Template (50 ng)	1 µL
ddH_2O	17.3 µL

Table 6: Gradient PCR temperature program.

PCR program	Temperature	Time
Initial denaturation	98 °C	30 s
30 cycles:		
Denaturation	98 °C	10 s
Annealing	64-69 °C	15 s
Extension	72 °C	2.5 min
Final extension	72 °C	5 min
Hold	4 °C	∞

5.5 Biochemical Methods

5.5.1 Expression of recombinant dioxygenases

Expression (Table 7) was performed in *E. coli* BL21(DE3) carrying the plasmid with the genes of the recombinant protein. For that, 400 mL of TB medium supplemented with the corresponding antibiotic were inoculated with 5-10 mL of ONC and incubated at 37 °C and 120 rpm. Once an OD_{600} of 0.6-0.8 was reached, overexpression was induced by addition of isopropanyl β-D-1 thiogalactopyranoside (IPTG) and cultivated for four hours or overnight at corresponding temperatures

(Table 7). Samples were taken right after induction (t_0), after four hours (t_4) and on the next day ($t_{O/N}$) to analyze the expression of recombinant protein. Cells were harvested by centrifugation (8.000 g, 15 min, 4 °C) and the cell pellet was resuspended in 0.9 % NaCl. After centrifugation (4000 g, 15 min, 4 °C), the biomass was stored at -20 °C until further use.

Table 7: Overview of all expression conditions.

Enzyme	Temperature [°C]	Duration of expression	IPTG concentration [mM]
AvLDO	30	O/N	1
BtIDO	30	O/N	1
MMGDO	30	4h	1
MfIDO	37	2h	1
TmDO	20	O/N	1
GriE	20	O/N	0.1
KDO1	20	O/N	0.5
P4H	20	O/N	1
SadA	28	4h	1
FoPip4H	28	O/N	1
AnPip4H	28	4h	1

5.5.2 Expression optimization of AvLDO

Due to low soluble expression levels of AvLDO, which were obtained with the implemented conditions, an expression study was conducted. Expression was performed based on the conditions stated by Correia Cordeiro *et al*. (Correia Cordeiro *et al*., 2018). In addition to that, different IPTG concentrations and cultivation temperatures were investigated. Moreover, the addition of $FeSO_4$, a Chaperone Plasmid Kit (Takara Bio Inc., France) and expression with autoinduction medium (Studier, 2005) was examined.

5.5.2.1 IPTG concentration and cultivation temperature

Heterologous expression of AvLDO was performed in *E. coli* BL21(DE3). For that, 80 mL of TB medium supplemented with kanamycin (final concentration 40 µg/mL) were inoculated with an ONC to an OD_{600} of 0.1 and incubated at 37 °C and 120 rpm. Once an OD_{600} of 0.6 or 1.2 was reached, overexpression was induced by addition of IPTG and cultivated for 24 hours at 20 °C (Table 8). Samples were taken right after induction (t_0), after two hours, after four hours (t_4) and after 24 hours (t_{24}) according to equation (2) to analyze the expression of recombinant protein. Cells

were harvested by centrifugation (4.000 g, 15 min, 4 °C) and the biomass was stored at -20 °C until further use.

$$V_{sample} = \frac{7}{OD_{600}} \quad (2)$$

Table 8: Overview of all expression conditions.

Approach	Cultivation temperature	Expression temperature	OD_{600} (induction)	IPTG concentration [mM]
1	37 °C	20 °C	0.6	1
2				0.5
3				0.1
4	37 °C	20 °C	1.2	1
5				0.5
6				0.1

The same procedure was implemented with a cultivation temperature of 28 °C and 0.5 mM IPTG (Table 9).

Table 9: Overview of all expression conditions.

Approach	Cultivation temperature	Expression temperature	OD_{600} (induction)	IPTG concentration [mM]
7	28 °C	20 °C	0.6	0.5
8			1.2	0.5

5.5.2.2 Co-expression of chaperones

Expression was conducted in *E. coli* BL21(DE3) carrying the chaperone plasmid Gro7 and the expression plasmid with the target protein. For this purpose, 80 mL of TB medium supplemented with the corresponding antibiotics were inoculated with an ONC to an OD_{600} of 0.1 and incubated at 37 °C and 120 rpm. After an OD_{600} of 0.4-0.5 was reached, L-arabinose (0.5 mg/mL) was added to induct chaperone expression and cultures were incubated at 20 °C. At an OD_{600} of 0.7-0.8 overexpression of the target protein was induced by addition of 0.5 mM IPTG and cultivated for 24 hours at 20 °C. Sampling and further procedures of the experiment were implemented as stated above in section 5.5.2.1.

5.5.2.3 Addition of $FeSO_4$

Another approach to improve soluble expression of AvLDO was the addition of $FeSO_4$. Heterologous expression was implemented as described in 5.5.2.1. At an OD_{600} of 0.6, 0.5 mM IPTG and 1 mM $FeSO_4$ was added at the time of induction, respectively.

5.5.2.4 Autoinduction

For auto induced protein expression, ZYP-5052 medium was prepared as described by Studier (Studier, 2005). All medium components as well as the composition of ZYP-0.8G and ZYP-5052 medium can be seen in the appendix (see 7.3). Autoinduction was implemented with *E. coli* BL21(DE3) and *E. coli* JM109(DE3) carrying the plasmid with the gene of interest**.** Prior expression, ONC's were prepared by inoculating 15 mL of ZYP-0.8G medium with a single colony containing kanamycin with a final concentration of 40 µg/mL. For expression, 200 mL ZYP-5052 medium were inoculated with 400 µL ONC and incubated for four hours at 37 °C and 120 rpm. The glucose concentration was pursued with glucose test stripes (Macherey-Nagel, Germany). Once the glucose was fully depleted, cultures were incubated for 20 hours at 20 °C and 120 rpm. Sampling and further procedures of the experiment were implemented as stated above in section 5.5.2.1.

5.5.3 Purification of recombinant dioxygenases

The cell pellet was resuspended in lysis buffer (10 mL/g pellet) (Table 10) and disrupted by sonication (Branson sonifier: output control 7-8, duty cycle 80 %, 5 min) twice. To avoid heat accumulation, cells were kept on ice in between. After removal of cell debris by centrifugation (15.000 g, 30 min, 4 °C), the target enzyme was purified from the supernatant by Ni-affinity chromatography (Bio-Rad, USA). The affinity column was equilibrated with five column volumes (cv) buffer containing 10 mM imidazole. The cell free extract (CFE) was applied to the column and incubated on ice for 20 minutes with slow shaking. The column was washed with 10 cv buffer containing 10 mM imidazole. A second wash was implemented with 10 cv buffer containing 30 mM imidazole. 3 mL fractions were collected. The protein was eluted with 5-7 cv buffer containing 300 mM imidazole and 1.5 mL fractions were collected. In the end the column was washed with 5 cv buffer containing 500 mM imidazole and 1.5 mL fractions were kept. Each collected purification

fraction was analyzed with SDS-PAGE (see 5.5.5.2). The protein concentration of the eluate was measured spectrophotometrically. Eluate fractions which contained sufficient amounts of protein were pooled and concentrated to a final volume of 2.5 mL (4000 rpm, 4 °C) using vivaspin® 20 centrifugal concentrators (Sartorius, Germany). For buffer exchange and imidazole removal the eluate was applied on a PD-10 Desalting Column (GE Healthcare, USA) according to the manufacturer's protocol. For each enzyme the corresponding equilibration buffer was used (Table 10). The protein concentration of the eluate was measured spectrophotometrically and concentrated using vivaspin® 20 centrifugal concentrators to a final volume of 2 mL. The purified protein was directly used for setting up biotransformations or stored at 4 °C until further use.

Table 10: Overview of the purification conditions for each enzyme.

Enzyme	MW [kDa]	Lysis Buffer	Buffer Exchange	Abs 0.1 %
AvLDO	33	50 mM potassium phosphate buffer (KPi), pH 7.4 500 mM NaCl 10 mM imidazole 10 % glycerol	50 mM sodium acetate buffer, pH 4 10 % glycerol; 500 mM NaCl	1.70
BtlDO	29		50 mM KPi, pH 6 10 % glycerol; 500 mM NaCl	0.81
SadA	33		50 mM KPi, pH 6.5 10 % glycerol; 500 mM NaCl	1.1
FoPip4H	40		50 mM KPi, pH 6.5 10 % glycerol; 500 mM NaCl	0.72
AnPip4H	40		50 mM KPi, pH 6.5 10 % glycerol; 500 mM NaCl	1.06
P4H	30		Not purified	-
GriE	32	50 mM KPi, pH 8; 500 mM NaCl 10 mM imidazole; 10 % glycerol	50 mM KPi, pH 8 10 % glycerol; 500 mM NaCl	0.83
KDO1	43	50 mM KPi, pH 8; 500 mM NaCl 10 mM imidazole; 10 % glycerol	Not purified	-
MflDO	32	50 mM KPi, pH 7.4 500 mM NaCl 10 % glycerol	Not purified	-
TmDO	32			
MMGDO	35			

*Abs 0.1% corresponds to the protein absorbance of a 0.1 % solution, which corresponds to mg/mL. The absorbance values were calculated from the respective amino acid sequences at 280 nm using ProtParam.

5.5.3.1 Optimized lysis and purification of AvLDO

The cell pellet was resuspended in 20 mM Tris-HCl (pH 7.4) containing 300 mM NaCl and disrupted by sonication (Branson sonifier: output control 5, duty cycle 50 %, 2 min) three times. To avoid heat accumulation, cells were kept on ice in between. After centrifugation (15.000 g, 30 min, 4 °C) the supernatant filtrated and purified with a Ni-affinity column (Bio-Rad, USA) as described in 5.5.3 or stored at 4 °C for further use. In case the purification was implemented on the next day it is suggested to add imidazole (final concentration 10 mM) right before the procedure starts. With a PD-10 Desalting Column (GE Healthcare, USA) the buffer was replaced after purification with 50 mM sodium acetate (pH 4) containing 10 % glycerol and the remaining imidazole was removed according to the manufacturer's protocol.

5.5.4 Determination of protein concentration

The protein concentration of each collected purification fraction was determined spectrophotometrically with the nanodrop 2000c spectrophotometer (Peqlab, Germany). The measured protein concentration (based on 1 Abs = 1 mg/mL) is then divided by the protein absorbance value of a 0.1 % solution (Abs 0.1 %) given in Table 10, which were calculated based on the amino acid sequence using ProtParam (Gasteiger et al., 2005).

For a more accurate determination of the protein concentration a bicinchoninic acid assay was performed (BCA assay) (Thermo Fisher Scientific, USA) in microplates according to the manufacturer's protocol. The absorption was measured in triplicates with a plate reader (BioTek Instruments, USA).

$$Protein\ concentration\ \left[\frac{mg}{mL}\right] = \frac{Protein\ concentration\ \left[\frac{mg}{mL}\right]\ at\ 280\ nm}{Abs\ 0.1\ \%} \qquad (3)$$

5.5.5 Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

5.5.5.1 SDS-PAGE of protein expression samples

Soluble and insoluble protein fractions of protein expression samples were analyzed by SDS-PAGE with ExpressPlus™ PAGE gels (12%) (GenScript, USA). For that,

cell pellets were disrupted with BugBuster®. The following equation (4) was used to calculate the specific amount of BugBuster® for each sample.

$$V_{BugBuster}[\mu L] = \frac{OD_{600} * V_{sample}[mL]}{30} \quad (4)$$

The mixture was shaken at 300 rpm in a ThermoMixer® for 20 minutes at room temperature. Afterwards debris was pelleted by centrifugation (16.000 rpm, 20 min, 4 °C) and the supernatant was transferred into a fresh Eppendorf tube. Potassium phosphate buffer was added in the same amount as Bugbuster® and the pellet was resuspended. 5 µL of SDS sample buffer (4x) and 10 µL distilled water were added to 5 µL of sample, denatured at 95 °C for 10 minutes and loaded on the gel. As standard, 10 µL of PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA) was used. SDS gels were run in Tris-MOPS running buffer (1x) (GenScript, USA) at 120 V and 200 mA for about one hour. They were stained with Coomassie Brillant Blue overnight and destained with destaining solution (see appendix 7.2) until bands were visible.

5.5.5.2 SDS-PAGE of protein purification fractions

After determining the protein concentration of each collected purification fraction by BCA assay, proteins were analyzed by SDS-PAGE. For that, 5 µg of sample were analyzed on the gel. The calculated volume was filled up with distilled water to a final volume of 15 µL. Afterwards 5 µl of SDS sample buffer (4x) were added to 15 µL sample and denatured at 95 °C for 10 min. 18 µL sample and 10 µL of PageRuler™ Prestained Protein Ladder were analyzed with ExpressPlus™ PAGE gels (12%) (GenScript, USA). Gels were run in Tris-MOPS running buffer (1x) at 120 V and 200 mA for about one hour. Gels were stained with Coomassie Brillant Blue overnight and destained with destaining solution (see appendix 7.2) until bands were visible.

5.5.6 Activity assay

All reaction components of the activity assay are shown below in Table 11. Stock solutions of the substrate, α-ketoglutarate, ascorbate and $FeSO_4x7H_2O$ were prepared. All components were dissolved in 10 mL distilled water and 2.5 mL of the corresponding buffer stock (Table 11 and Table 12) except for $FeSO_4x7H_2O$ which

is only dissolved in water. The pH value was adjusted with HCl or NaOH, and the volume filled up with water to 50 mL. 2 mL aliquots were stored at -21 °C. The biotransformations were performed in 1.5 mL glass vials with a reaction volume of 0.5 mL at the corresponding conditions shown in Table 12 at 750 rpm. 50 µL of enzymatic reaction were quenched by addition of 25 µL 50 % v/v acetic acid. Debris was removed by centrifugation (16.000 rpm, 1 min, 4 °C) and the supernatant was stored at -21 °C. Product formation was detected with thin-layer chromatography and ninhydrin staining.

Table 11: Composition of the standard reaction mixture; X = amount of protein.

Components	Concentration of stock solution	Final concentration
Substrate	50 mM	5 mM
α-ketoglutarate	100 mM	7.5 mM
Ascorbate	5 mM	0.5 mM
$FeSO_4x7H_2O$	50 mM	5 mM
Protein		X
Buffer	1 M	50 mM

Table 12: Overview of all biotransformation conditions for each dioxygenase; n.d: not determined

Enzyme	Substrate	pH	Buffer	Amount enzyme	Reaction temperature
AvLDO	L-leucine	4	sodium acetate buffer	Lysate: 0.7 mg/mL Purified: 1 mg/mL	25 °C
BtIDO	L-isoleucine	6	potassium phosphate buffer (KPi)	Lysate: 0.9 mg/L Purified: 1 mg/mL	25 °C
SadA	N-succinyl-L-leucine	6.5	KPi	Lysate: 1.9 mg/mL Purified: 2 mg/mL	30 °C
TmDO	L-isoleucine/ L-methionine	6.5	KPi	n.d.	30 °C
MMGDO	L-isoleucine/ L-methionine	6.5	KPi	n.d.	30 °C
MfIDO	L-leucine/ L-methionine	6.5	KPi	Lysate: 0.7 mg/mL	35 °C
GriE	L-leucine/ α-methyl-L-leucine	8	KPi	Lysate: 1.7 mg/mL Purified: 2 mg/mL	35 °C
P4H	L-proline	6.5	KPi	Lysate: 0.7 mg/mL	35 °C
KDO1	L-lysine	7.4	KPi	Lysate: 0.12 mg/mL	20 °C
FoPip4H	L-pipecolic acid	6.5	KPi	Lysate: 1.2 mg/mL Purified: 1.8 mg/mL	20 °C
AnPip4H	L-pipecolic acid	6.5	KPi	Lysate: 0.4 mg/mL Purified: 1.2 mg/mL	20 °C

5.5.6.1 Optimized activity assay for AvLDO

All reaction components of the activity assay are shown in Table 11. All stock solutions were freshly prepared when required as described in 5.5.6 and were not stored to prevent oxidation or decomposition of $FeSO_4$ and ascorbate. The biotransformations were performed in 1.5 mL glass vials with a reaction volume of 0.5 mL at 25 °C and 750 rpm with the corresponding amount of protein (Table 13). Sampling and further procedure can be seen in 5.5.6.

Table 13: Overview of all biotransformation conditions for AvLDO.

Enzyme	Substrates	Buffer	Amount of protein	Reaction temperature
AvLDO	L-leucine	Sodium acetate, pH 4	Lysate: 1.65 mg/mL Purified enzyme: 0.25 mg/mL and 0.5 mg/mL	25 °C
	L-isoleucine		Purified enzyme: 0.25 mg/mL and 0.5 mg/mL	
	α-methyl-L-leucine		Purified enzyme: 0.25 mg/mL and 0.5 mg/mL	

5.5.6.2 Thin-layer chromatography (TLC)

The activity in hydroxylation of natural substrates was confirmed by TLC after stained with ninhydrin. For that, 2 µL of reaction mixture were applied on a silica gel TLC plate (Merck Millipore, USA) and chromatograms were developed with 1-butanol: acetic acid: ddH_2O (22:3:5), referred to as mobile phase A. Plates were dried and spots were detected with staining solution (a) consisting of 0.2 % w/v ninhydrin. L-pipecolic acid and L-lysine were developed with 1-butanol: acetic acid: ddH_2O (5:3:2), referred to as mobile phase B, for better separation. Spots were detected with staining solution (b) containing 1 % w/v ninhydrin and 500 µL acetic acid in 100 mL acetone. L-amino acids and hydroxy isoleucine served as standards. These were prepared by mixing 5 µL of 50 % v/v acetic acid with 10 µL of 5 mM standard mixture. The retention factor (R_f) was calculated for each compound with equation (5).

$$Rf = \frac{distance\ spot\ moved}{distance\ solvent\ moved} \quad (5)$$

5.5.7 Selection of active α-ketoglutarate dependent dioxygenases using *E. coli* 3Δ*sucA* – growth studies

5.5.7.1 Growth assay on agar plates

Initially, selection of active α-ketoglutarate dependent dioxygenases was investigated on agar plates. For that, chemo competent *E. coli* 3Δ*sucA* were transformed with 50 ng plasmid DNA of pet28a_GriE, pet28a_AvLDO, pET22b_BtIDO and pET28a(+) via heat shock method and regenerated for 2 hours (see 5.4.3). Cells were centrifuged (12.000 rpm, 24 °C, 3 min) and the supernatant was discarded. Afterwards, the cell pellet was resuspended in minimal medium (appendix see 7.3) followed by a second centrifugation step. The pellet was resuspended in 300 µL minimal medium and cells were streaked out on four different kinds of agar plates, namely minimal medium without the addition of substrate, minimal medium containing L-leucine or L-isoleucine and LB medium with the corresponding antibiotic. Plates were incubated at 28 °C for about two to three days.

5.5.7.2 Growth assay in liquid medium

For better monitoring of the growth behavior, an investigation in liquid medium was implemented. Initially, precultures of the recombinant *E. coli* strains were prepared. For that, 10 mL of the corresponding medium and antibiotics (Table 14) were inoculated with 25 µL of ONC and incubated at 30 °C and 120 rpm for 6 h. Afterwards, main cultures (Table 14) of 50 mL were inoculated with 125 µL preculture and incubated at 30 °C and 120 rpm. For the induction of the protein expression, 1 mM IPTG was added to both precultures and main cultures at the time of inoculation, respectively. Bacterial cell growth was monitored by measuring the optical density at 600 nm.

Table 14: Overview of reaction conditions used for growth studies**.** EV: empty vector; DO: dioxygenase.

Approach	Strain	Medium for pre- and main culture
1a	BL21_EV	LB+kanamycin/ampicillin
1b	3Δ*sucA*_EV	
1c	3Δ*sucA*_DO	
2a	BL21_EV	M9-substrate+kanamycin/ampicillin
2b	3Δ*sucA*_EV	
2c	3Δ*sucA*_DO	
3a	BL21_EV	M9+leucine+kanamycin/ampicillin
3b	3Δ*sucA*_EV	
3c	3Δ*sucA*_DO	
4a	BL21_EV	M9+isoleucine+kanamycin/ampicillin
4b	3Δ*sucA*_EV	
4c	3Δ*sucA*_DO	

6 References

Amorim Franco, T.M., and Blanchard, J.S. (2017). Bacterial Branched-Chain Amino Acid Biosynthesis: Structures, Mechanisms, and Drugability. Biochemistry *56*, 5849–5865.

Aronson, J.N., Borris, D.P., Doerner, J.F., and Akers, E. (1975). γ-Aminobutyric Acid Pathway and Modified Tricarboxylic Acid Cycle Activity During Growth and Sporulation of Bacillus thuringiensis. Appl Microbiol *30*, 489–492.

Baud, D., Saaidi, P.-L., Monfleur, A., Harari, M., Cuccaro, J., Fossey, A., Besnard, M., Debard, A., Mariage, A., Pellouin, V., et al. (2014). Synthesis of Mono- and Dihydroxylated Amino Acids with New α-Ketoglutarate-Dependent Dioxygenases: Biocatalytic Oxidation of C□H Bonds. ChemCatChem *6*, 3012–3017.

Bollinger, J.M., Price, J.C., Hoffart, L.M., Barr, E.W., and Krebs, C. (2005). Mechanism of Taurine: α-Ketoglutarate Dioxygenase (TauD) from Escherichia coli. European Journal of Inorganic Chemistry *2005*, 4245–4254.

Briñas, L., Zarazaga, M., Sáenz, Y., Ruiz-Larrea, F., and Torres, C. (2002). Beta-lactamases in ampicillin-resistant Escherichia coli isolates from foods, humans, and healthy animals. Antimicrob. Agents Chemother. *46*, 3156–3163.

Bugg, T.D.H. (2003). Dioxygenase enzymes: catalytic mechanisms and chemical models. Tetrahedron *59*, 7075–7101.

Chapman, T. (2005). Pure but not simple. Nature *434*, 795–796.

Choi, J.-M., Han, S.-S., and Kim, H.-S. (2015). Industrial applications of enzyme biocatalysis: Current status and future aspects. Biotechnol. Adv. *33*, 1443–1454.

Correia Cordeiro, R.S., Enoki, J., Busch, F., Mügge, C., and Kourist, R. (2018). Cloning and characterization of a new delta-specific l-leucine dioxygenase from Anabaena variabilis. J. Biotechnol. *284*, 68–74.

Dong, J., Fernández-Fueyo, E., Hollmann, F., Paul, C.E., Pesic, M., Schmidt, S., Wang, Y., Younes, S., and Zhang, W. (2018). Biocatalytic Oxidation Reactions: A Chemist's Perspective. Angew. Chem. Int. Ed. Engl. *57*, 9238–9261.

Duncan, T., Trewick, S.C., Koivisto, P., Bates, P.A., Lindahl, T., and Sedgwick, B. (2002). Reversal of DNA alkylation damage by two human dioxygenases. PNAS *99*, 16660–16665.

Enoki, J., Meisborn, J., Müller, A.-C., and Kourist, R. (2016). A Multi-Enzymatic Cascade Reaction for the Stereoselective Production of γ-Oxyfunctionalyzed Amino Acids. Front. Microbiol. *7*.

Faber, K. (2018). Biotransformations in Organic Chemistry: A Textbook (Springer International Publishing).

Gamenara, D., Seoane, G., Méndez, P.S., and María, P.D. de (2012). Redox Biocatalysis: Fundamentals and Applications (John Wiley & Sons).

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook, J.M. Walker, ed. (Totowa, NJ: Humana Press), pp. 571–607.

Good, N.E., and Izawa, S. (1972). Hydrogen ion buffers. Meth. Enzymol. *24*, 53–68.

Hannig, G., and Makrides, S.C. (1998). Strategies for optimizing heterologous protein expression in Escherichia coli. Trends Biotechnol. *16*, 54–60.

Hansson, M.D., Rzeznicka, K., Rosenbäck, M., Hansson, M., and Sirijovski, N. (2008). PCR-mediated deletion of plasmid DNA. Anal. Biochem. *375*, 373–375.

Hausinger, R.P. (2004). FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. Crit. Rev. Biochem. Mol. Biol. *39*, 21–68.

Hayer-Hartl, M., Bracher, A., and Hartl, F.U. (2016). The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. Trends Biochem. Sci. *41*, 62–76.

Herr, C.Q., and Hausinger, R.P. (2018). Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. Trends in Biochemical Sciences *43*, 517–532.

Hibi, M., and Ogawa, J. (2014). Characteristics and biotechnology applications of aliphatic amino acid hydroxylases belonging to the Fe(II)/α-ketoglutarate-dependent dioxygenase superfamily. Appl Microbiol Biotechnol *98*, 3869–3876.

Hibi, M., Kawashima, T., Kodera, T., Smirnov, S.V., Sokolov, P.M., Sugiyama, M., Shimizu, S., Yokozeki, K., and Ogawa, J. (2011). Characterization of Bacillus thuringiensis L-isoleucine dioxygenase for production of useful amino acids. Appl. Environ. Microbiol. *77*, 6926–6930.

Hibi, M., Kawashima, T., Kasahara, T., Sokolov, P.M., Smirnov, S.V., Kodera, T., Sugiyama, M., Shimizu, S., Yokozeki, K., and Ogawa, J. (2012). A novel Fe(II)/α-ketoglutarate-dependent dioxygenase from Burkholderia ambifaria has β-hydroxylating activity of N-succinyl l-leucine. Lett. Appl. Microbiol. *55*, 414–419.

Hibi, M., Kawashima, T., Sokolov, P.M., Smirnov, S.V., Kodera, T., Sugiyama, M., Shimizu, S., Yokozeki, K., and Ogawa, J. (2013). L-leucine 5-hydroxylase of Nostoc punctiforme is a novel type of Fe(II)/α-ketoglutarate-dependent dioxygenase that is useful as a biocatalyst. Appl. Microbiol. Biotechnol. *97*, 2467–2472.

Hibi, M., Mori, R., Miyake, R., Kawabata, H., Kozono, S., Takahashi, S., and Ogawa, J. (2016). Novel Enzyme Family Found in Filamentous Fungi Catalyzing trans-4-Hydroxylation of l-Pipecolic Acid. Appl Environ Microbiol *82*, 2070–2077.

Hüttel, W. (2013). Biocatalytic Production of Chemical Building Blocks in Technical Scale with α-Ketoglutarate-Dependent Dioxygenases. Chemie Ingenieur Technik *85*, 809–817.

Hutton, J.J., Tappel, A.L., and Udenfriend, S. (1967). Cofactor and substrate requirements of collagen proline hydroxylase. Archives of Biochemistry and Biophysics *118*, 231–240.

Islam, M.S., Leissing, T.M., Chowdhury, R., Hopkinson, R.J., and Schofield, C.J. (2018). 2-Oxoglutarate-Dependent Oxygenases. Annu. Rev. Biochem. *87*, 585–620.

Jia, B., and Jeon, C.O. (2016). High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. Open Biol *6*.

Kaur, J., Kumar, A., and Kaur, J. (2018). Strategies for optimization of heterologous protein expression in E. coli: Roadblocks and reinforcements. Int. J. Biol. Macromol. *106*, 803–822.

Kodera, T., Smirnov, S.V., Samsonova, N.N., Kozlov, Y.I., Koyama, R., Hibi, M., Ogawa, J., Yokozeki, K., and Shimizu, S. (2009). A novel l-isoleucine hydroxylating enzyme, l-isoleucine dioxygenase from Bacillus thuringiensis, produces (2S,3R,4S)-4-hydroxyisoleucine. Biochem. Biophys. Res. Commun. *390*, 506–510.

Kornberg, H.L. (1966). The role and control of the glyoxylate cycle in Escherichia coli. Biochem J *99*, 1–11.

Li, M., Ho, P.Y., Yao, S., and Shimizu, K. (2006). Effect of sucA or sucC gene knockout on the metabolism in Escherichia coli based on gene expressions, enzyme activities, intracellular metabolite concentrations and metabolic fluxes by 13C-labeling experiments. Biochemical Engineering Journal *30*, 286–296.

Li, Z., Nimtz, M., and Rinas, U. (2014). The metabolic potential of Escherichia coli BL21 in defined and rich medium. Microbial Cell Factories *13*, 45.

Lukat, P., Katsuyama, Y., Wenzel, S., Binz, T., König, C., Blankenfeldt, W., Brönstrup, M., and Müller, R. (2017). Biosynthesis of methyl-proline containing griselimycins, natural products with anti-tuberculosis activity †Electronic supplementary information (ESI) available. See DOI: 10.1039/c7sc02622f Click here for additional data file. Chem Sci *8*, 7521–7527.

Martín, J.F., Casqueiro, J., Kosalková, K., Marcos, A.T., and Gutiérrez, S. (1999). Penicillin and cephalosporin biosynthesis: mechanism of carbon catabolite regulation of penicillin production. Antonie Van Leeuwenhoek *75*, 21–31.

Martinez, S., and Hausinger, R.P. (2015). Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-dependent Oxygenases. J. Biol. Chem. *290*, 20702–20711.

Meisborn, J. (2014). Untersuchungen von Multi-Enzym-Systemen zur umweltfreundlichen Herstellung von optisch reinen Hydroxyaminosäuren.

Niesen, F.H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols *2*, 2212–2221.

Ogawa, J., Kodera, T., Smirnov, S.V., Hibi, M., Samsonova, N.N., Koyama, R., Yamanaka, H., Mano, J., Kawashima, T., Yokozeki, K., et al. (2011). A novel L-isoleucine metabolism in Bacillus thuringiensis generating (2S,3R,4S)-4-hydroxyisoleucine, a potential insulinotropic and anti-obesity amino acid. Appl. Microbiol. Biotechnol. *89*, 1929–1938.

Ohnishi, K., Hasegawa, A., Matsubara, K., Date, T., Okada, T., and Kiritani, K. (1988). Cloning and nucleotide sequence of the brnQ gene, the structural gene for a membrane-associated component of the LIV-II transport system for branched-chain amino acids in Salmonella typhimurium. Jpn. J. Genet. *63*, 343–357.

Paliy, O., and Gunasekera, T.S. (2007). Growth of E. coli BL21 in minimal media with different gluconeogenic carbon sources and salt contents. Appl. Microbiol. Biotechnol. *73*, 1169–1172.

Peters, C., and Buller, R.M. (2019). Industrial Application of 2-Oxoglutarate-Dependent Oxygenases. Catalysts *9*, 221.

Pratter, S.M., Ivkovic, J., Birner-Gruenberger, R., Breinbauer, R., Zangger, K., and Straganz, G.D. (2014). More than just a Halogenase: Modification of Fatty Acyl Moieties by a Trifunctional Metal Enzyme. ChemBioChem *15*, 567–574.

Qin, H.-M., Miyakawa, T., Jia, M.Z., Nakamura, A., Ohtsuka, J., Xue, Y.-L., Kawashima, T., Kasahara, T., Hibi, M., Ogawa, J., et al. (2013). Crystal Structure of a Novel N-Substituted L-Amino Acid Dioxygenase from Burkholderia ambifaria AMMD. PLOS ONE *8*, e63996.

Sauer, U., and Eikmanns, B.J. (2005). The PEP—pyruvate—oxaloacetate node as the switch point for carbon flux distribution in bacteria: We dedicate this paper to Rudolf K. Thauer, Director of the Max-Planck-Institute for Terrestrial Microbiology in Marburg, Germany, on the occasion of his 65th birthday. FEMS Microbiol Rev *29*, 765–794.

Schofield, C.J., and Zhang, Z. (1999). Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. Curr. Opin. Struct. Biol. *9*, 722–731.

Shehadul Islam, M., Aryasomayajula, A., and Selvaganapathy, P.R. (2017). A Review on Macroscale and Microscale Cell Lysis Methods. Micromachines (Basel) *8*.

Sheldon, R.A., Arends, I., and Hanefeld, U. (2007). Green Chemistry and Catalysis (John Wiley & Sons).

Sigrell, J.A., Eklund, P., Galin, M., Hedkvist, L., Liljedahl, P., Johansson, C.M., Pless, T., and Torstenson, K. (2003). Automated multi-dimensional purification of tagged proteins. J. Struct. Funct. Genomics *4*, 109–114.

Smirnov, S.V., Kodera, T., Samsonova, N.N., Kotlyarova, V.A., Rushkevich, N.Y., Kivero, A.D., Sokolov, P.M., Hibi, M., Ogawa, J., and Shimizu, S. (2010). Metabolic engineering of Escherichia coli to produce (2S, 3R, 4S)-4-hydroxyisoleucine. Appl Microbiol Biotechnol *88*, 719–726.

Smirnov, S.V., Sokolov, P.M., Kodera, T., Sugiyama, M., Hibi, M., Shimizu, S., Yokozeki, K., and Ogawa, J. (2012). A novel family of bacterial dioxygenases that catalyse the hydroxylation of free L-amino acids. FEMS Microbiol. Lett. *331*, 97–104.

Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expression and Purification *41*, 207–234.

Thakker, C., Martínez, I., San, K.-Y., and Bennett, G.N. (2012). Succinate production in Escherichia coli. Biotechnol J *7*, 213–224.

Theodosiou, E., Breisch, M., Julsing, M.K., Falcioni, F., Bühler, B., and Schmid, A. (2017). An artificial TCA cycle selects for efficient α-ketoglutarate dependent hydroxylase catalysis in engineered Escherichia coli. Biotechnol. Bioeng. *114*, 1511–1520.

Townsend, C.A. (2002). New reactions in clavulanic acid biosynthesis. Current Opinion in Chemical Biology *6*, 583–589.

Vaillancourt, F.H., Yin, J., and Walsh, C.T. (2005). SyrB2 in syringomycin E biosynthesis is a nonheme FeII alpha-ketoglutarate- and O2-dependent halogenase. Proc. Natl. Acad. Sci. U.S.A. *102*, 10111–10116.

Vivoli, M., Novak, H.R., Littlechild, J.A., and Harmer, N.J. (2014). Determination of Protein-ligand Interactions Using Differential Scanning Fluorimetry. Journal of Visualized Experiments.

Wohlgemuth, R. (2010). Biocatalysis--key to sustainable industrial chemistry. Curr. Opin. Biotechnol. *21*, 713–724.

Wu, L.-F., Meng, S., and Tang, G.-L. (2016). Ferrous iron and α-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products. Biochim. Biophys. Acta *1864*, 453–470.

Xu, F. (2005). Applications of oxidoreductases: Recent progress. Industrial Biotechnology *1*, 38–50.

Xu, Z., and Sigler, P.B. (1998). GroEL/GroES: structure and function of a two-stroke folding machine. J. Struct. Biol. *124*, 129–141.

Zhang, C., Ma, J., Li, Z., Liang, Y., Xu, Q., Xie, X., and Chen, N. (2018). A strategy for L-isoleucine dioxygenase screening and 4-hydroxyisoleucine production by resting cells. Bioengineered *9*, 72–79.

Zhu, L.-W., and Tang, Y.-J. (2017). Current advances of succinate biosynthesis in metabolically engineered Escherichia coli. Biotechnol. Adv. *35*, 1040–1048.

Zwick, C.R., and Renata, H. (2018a). Remote C–H Hydroxylation by an α-Ketoglutarate-Dependent Dioxygenase Enables Efficient Chemoenzymatic Synthesis of Manzacidin C and Proline Analogs. J. Am. Chem. Soc. *140*, 1165–1169.

Zwick, C.R., and Renata, H. (2018b). A one-pot chemoenzymatic synthesis of (2S, 4R)-4-methylproline enables the first total synthesis of antiviral lipopeptide cavinafungin B. Tetrahedron *74*, 6469–6473.

7 Appendix

7.1 Chemicals and labware

Table 15: List of chemicals used in this study.

Chemicals	Source	CAS number
Acetic acid	Roth, Germany	64-19-7
Acetone	Roth, Germany	67-64-1
Agar-Agar	Roth, Germany	9002-18-0
Ammonium chloride	AppliChem, Germany	12125-02-9
Ammonium sulfate	Roth, Germany	7783-20-2
Ampicillin sodium salt	Roth, Germany	69-52-3
Anti-foaming solution	Sigma-Aldrich, USA	
Ascorbate	Roth, Germany	50-81-7
Biotin	Roth, Germany	58-85-5
Boric acid	Roth, Germany	10043-35-3
BugBuster®	Merck Millipore, USA	
Butanol	Roth, Germany	71-36-3
Calcium chloride dihydrate	Roth, Germany	10035-04-8
Cobalt (II) Chlorate Dihydrate	Roth, Germany	7646-79-9
Coomassie Brilliant Blue	Sigma-Aldrich, USA	6104-59-2
Dipotassiumhydrogen phosphate	Roth, Germany	7758-11-4
Disodium EDTA dihydrate	Roth, Germany	6381-92-6
Disodiumhydrogen phosphate	Roth, Germany	7558-79-4
Ethanol	Chem-Lab, Belgium	64-17-5
Glucose	Roth, Germany	50-99-7
Glycin	Roth, Germany	56-40-6
Glycerol	Roth, Germany	56-81-5
Hydrochloric acid	Roth, Germany	32862-91-2
(2*S*,3*R*,4*S*)-Hydroxyisoleucine	Sigma-Aldrich, USA	781658-23-9
Imidazole	Roth, Germany	288-32-4
Iron (II) sulphate heptahydrate	Roth, Germany	7782-63-0
L-isoleucine	TCI, Japan	73-32-5
Isopropyl β-D-1-thiogalactopyranosid	Roth, Germany	367-93-1
Kanamycin sulfate	Roth, Germany	25389-94-0
α-ketoglutarate	Roth, Germany	328-50-7
α-lactose	Roth, Germany	63-42-3
L-leucine	TCI, Japan	61-90-5
L-lysine	Fluorochem, UK	56-87-1
Magnesium chloride tetrahydrate	Sigma-Aldrich, USA	7791-18-6
Magnesium sulfate heptahydrate	Roth, Germany	10034-99-8
Manganese (II) Chloride tetrahydrate	Sigma-Aldrich, USA	13446-34-9
β-mercaptoenthanol	Sigma-Aldrich, USA	60-24-2
L-methionine	Roth, Germany	63-68-3
2-methyl-L-leucin	Flourchem, UK	144-24-1
Sodium molybdate dihydrate	Roth, Germany	10102-40-6

(3-(*N*-morpholino) propanesulfonic acid)	Sigma-Aldrich, USA	1132-61-2
Ninhydrin	Roth, Germany	485-47-2
L-pipecolic acid	Fluorochem, UK	3105-95-1
L-proline	TCI, Japan	147-85-3
Potassium acetate	Roth, Germany	127-08-2
Potassium chloride	Roth, Germany	7447-40-7
Potassium dihydrogenphosphate	Roth, Germany	7778-77-0
Sodium acetate	Roth, Germany	127-09-3
Sodium chloride	Roth, Germany	7647-14-5
Sodium dodecyl sulfate	Roth, Germany	151-21-3
Thiamine hydrochloride	Roth, Germany	67-03-8
TRIS	Roth, Germany	1185-53-1
Tryptone	Roth, Germany	91079-40-2
Yeast extract	Roth, Germany	8013-01-2
Zinc sulfate heptahydrate	Roth, Germany	7447-20-0

Table 16: List of labware used in this study.

Labware	Model	Source
Shaking incubator		Infors HT, Switzerland
Centrifuges	5415 R	Eppendorf, Germany
	5810 R	Eppendorf, Germany
	Avanti® J-20 XP	Beckman coulter, USA
Rotors	JA-25.50	Beckman coulter, USA
	JA-10	Beckman coulter, USA
ThermoMixers®	ThermoMixer® C	Eppendorf, Germany
	ThermoMixer® comfort	Eppendorf, Germany
	HTM 2	HTA-BioTec, Germany
Analytical balance	Practum®	Sartorius, Germany
Precision balances	AY-6000	Sartorius, Germany
pH-electrode	inoLab® pH 720	VWR International, USA
Vortex	Vortex-Genie 2	Scientific industries, USA
Spectrophotometer	BioPhotometer 6131	Eppendorf, Germany
	NanoDrop™ 2000	Thermo Scientific, USA
Branson Ultrasonics™ Sonifier	S-250A	Thermo Scientific, USA

7.2 Solutions for SDS-PAGE

Staining solution:

Table 17: Staining solution used for SDS-PAGE gels

Component	Amount for 1 L
Coomassie Brilliant Blue R-250	1 g
Glacial acetic acid (conc.)	100 mL
Ethanol	300 mL
dH_2O	600 mL

Filtrate and store at room temperature.

Destaining solution:

Table 18: Solution used for destaining of SDS-PAGE gels.

Component	Amount for 1 L
Glacial acetic acid (conc.)	100 mL
Ethanol	300 mL
dH_2O	600 mL

Store at room temperature.

7.3 Recepies

LB medium:

Table 19: Composition of LB medium.

Component	Amount of 1 L
Tryptone	10 g
Yeast extract	5 g
Sodium chloride	10 g
dH_2O	Fill up to 1 L

Autoclave and store at room temperature.

TB medium:

Table 20: Composition of TB medium.

Component	Amount of 1 L
Tryptone	12 g
Yeast extract	24 g
Glycerol	4 mL
dH_2O	Fill up to 1 L

Autoclave and store at room temperature.

0.17 M KH_2PO_4/0.72 M K_2HPO_4 buffer:

Table 21: KH_2PO_4/K_2HPO_4 buffer used for TB medium.

Component	Amount of 1 L
KH_2PO_4	23.1 g
K_2HPO_4	125.4 g
dH_2O	Fill up to 1 L

Autoclave and store at room temperature.

SOC (10x):

Table 22: Composition of SOC (10x) medium.

Component	Amount of 50 mL
KCl	0.1 g
$MgCl_2$	1 g
$MgSO_4$	1 g
Glucose monohydrate	2 g
ddH_2O	50 mL

10x SOC is diluted with LB medium to create a 1x LB-SOC mixture.

TfB1:

Table 23: Composition of TfB1 buffer used for preparation of chemo-competent cells.

Component	Amount for 500 mL
KCH_3COO	1.47 g
KCl	3.72 g
$CaCl_2$	0.55 g
Glycerol	65 mL

Autoclave and store at 4 °C.

TfB2:

Table 24: Composition of TfB2 buffer used for preparation of chemo-competent cells.

Component	Amount for 500 mL
KCl	0.37 g
$CaCl_2$	5.55 g
MOPS	1.05 g
Glycerol	65 mL

Autoclave and store at 4 °C.

$MgCl_2$ solution:

Table 25: Composition of TfB2 buffer used for preparation of chemo-competent cells.

Component	Amount for 50 mL
$MgCl_2x4H_2O$	4.76 g
dH_2O	Fill up to 50 mL

Sterilize over 0.22 µm filter and stored at 4 °C.

Autoinduction medium:

NPS (20x):

Table 26: NPS (20x) solution used for preparation of autoinduction medium.

Component	Amount for 1 L
$(NH_4)_2SO_4$	66 g
KH_2PO_4	136 g
Na_2HPO_4	142 g
dH_2O	900 mL

Autoclave and store at room temperature.

5052 (50x):

Table 27: 5052 (50x) solution used for preparation of autoinduction medium.

Component	Amount for 1 L
Glycerol	250 g
dH_2O	730 mL
Glucose	25 g
α-lactose	100 g

Autoclave and store at room temperature.

Glucose (40%):

Table 28: Glucose solution (40 %) used for preparation of autoinduction medium.

Component	Amount for 100 mL
Glucose	40 g
dH_2O	74 mL

Autoclave and store at room temperature.

$MgSO_4$ (1M):

Table 29: $MgSO_4$ (1M) used for preparation of autoinduction medium.

Component	Amount for 100 mL
$MgSO_4x7H_2O$	24.65 g
dH_2O	Fill up to 100 mL

Sterilize over 0.22 µm filter and stored at room temperature.

ZYP-0.8G:

Table 30: Composition of ZYP-0.8G medium used for preparation of ONC in autoinduction experiments.

Component	Amount for 100 mL
LB medium	93 mL
$MgSO_4$ (1M)	100 µL
Glucose (40%)	2 mL
NPS (20x)	5 mL

Autoclave 100 mL of LB medium and add all other sterilized solutions.

ZYP-5052:

Table 31: Composition of ZYP-5052 medium used for preparation of main cultures in autoinduction experiments.

Component	Amount for 1 L
LB medium	928 mL
$MgSO_4$ (1M)	1 mL
5052 (50x)	20 mL
NPS (20x)	50 mL

Autoclave 1 L of LB medium and add all other sterilized solutions.

Minimal medium:

Table 32: Composition of minimal medium used in growth studies.

Components	Final concentration	Stock	Stock concentration	Comments	Stock volume for 1 L medium
$Na_2HPO_4x2H_2O$	8.5 g/L	5x	42.5 g/L	Adjust to pH 7.0. Autoclave and store at room temperature	200 mL
KH_2PO_4	3.0 g/L	5x	15 g/L		
NaCl	0.5 g/L	5x	2.5 g/L		
NH_4Cl	1.0 g/L	1M	5.0 g/L		
$MgSO_4x7H_2O$	2 mL/L	1.000x	24.6 g/100 mL	Autoclave and store at room temperature	2 mL
ThiaminexHCl	1 mg/L	20.000x	1 mg/mL	Sterilize over 0.22 µm filter and store aliquots at -20 °C.	1 mL
Biotin	5 g/L	1.000x	0.1 mg/mL	Sterilize over 0.22 µm filter. Store aliquots at -20 °C.	50 µL
Us^{Fe} trace element solution	1 mL/L				1 mL
Glucose	5 g/L	20%	200 g/L	Sterilize over 0.22 µm filter. Store at room temperature.	25 mL
Substrate	5mM	10x	50 mM	Sterilize over 0.22 µm filter.	100 mL
ddH_2O				Sterilized by autoclaving	670 mL (w/o substrate: 770 mL)

US^{Fe} trace element solution (1000x):

Table 33: Composition of US^{Fe} trace element solution (1000x) used for minimal medium preparation.

Component	Amount for 1 L
$FeSO_4x7H_2O$	8.87 g/L
$CaCl_2x2H_2O$	4.12 g/L
$MnCl_2x2H_2O$	1.23 g/L
$ZnSO_4x7H_2O$	1.87 g/L
H_3BO_3	0.30 g/L
$Na_2MoO_4x2H_2O$	0.25 g/L
$CuCl_2x2H_2O$	0.15 g/L
Disodium $EDTAx2H_2O$	0.84 g/L
Dissolve everything in 1M HCl (82.8 mL 37 % HCl/L)	

Autoclave and store aliquots at -20 °C.

M9 agar plates:

15 g/L agar and the respective amount of ddH2O are mixed and autoclaved. All other sterilized solutions are added after the agar cooled down to 60 °C.

7.4 Protein sequences

Table 34: Protein sequences for all dioxygenases used in this study.

Enzyme	Accession no.
AvLDO	ABA21237.1
BtIDO	EEM30089.1
SadA	WP_011660927.1
MflDO	WP_011480845.1
TmDO	WP_014747361.1
MMGDO	EBP75668.1
GriE	AKC91859.1
FoPip4H	BAU50539.1
KDO1	WP_012784479.1
AnPip4H	XP_659994
AvLDO	ABA21237.1
BtIDO	EEM30089.1
SadA	WP_011660927.1
MflDO	WP_011480845.1
P4H	R9UTQ8.1

7.5 Primer sequences

Table 35: Primer sequences used in this study.

Name	Sequence 5'-3'
T7_Fw	TAATACGACTCACTATAGG
T7_Term	TGCTAGTTATTGCTCAGCG
pET22bxIDO_Fw	GTTAGCAGCCGGATCATGTATATCTCCTTCTTAAAGTTAA ACAAAATTATTTCTAGAG
pET22bxIDO_Rv	GAAGGAGATATACATGATCCGGCTGCTAACAAAGC

7.6 Plasmid maps

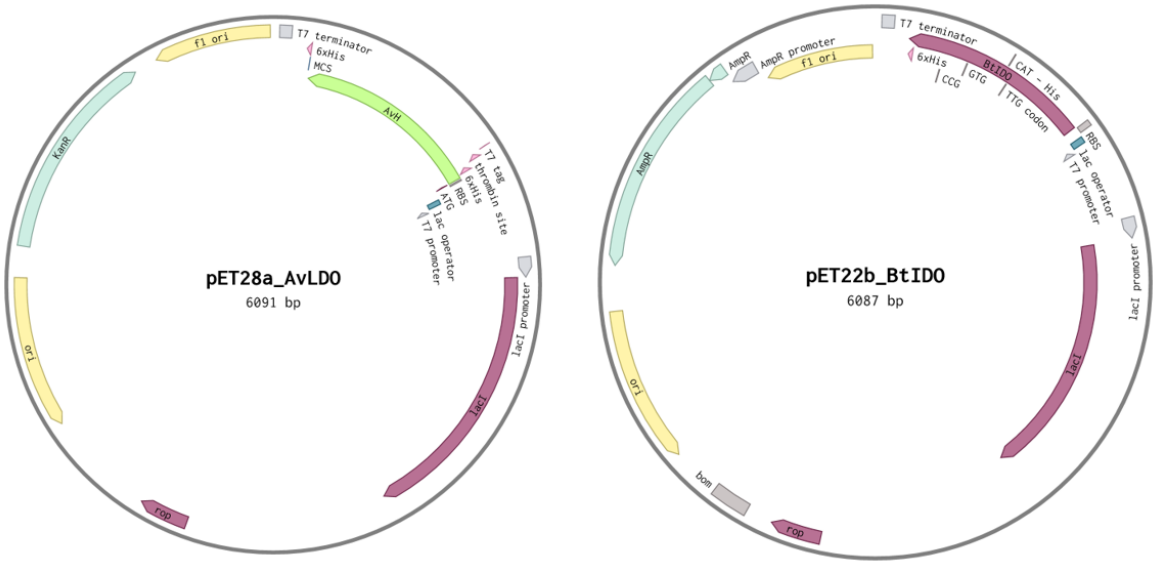

Figure 38:Vector map of construct pET28a_AvLDO and pET22b_BtIDO created with Benchling.

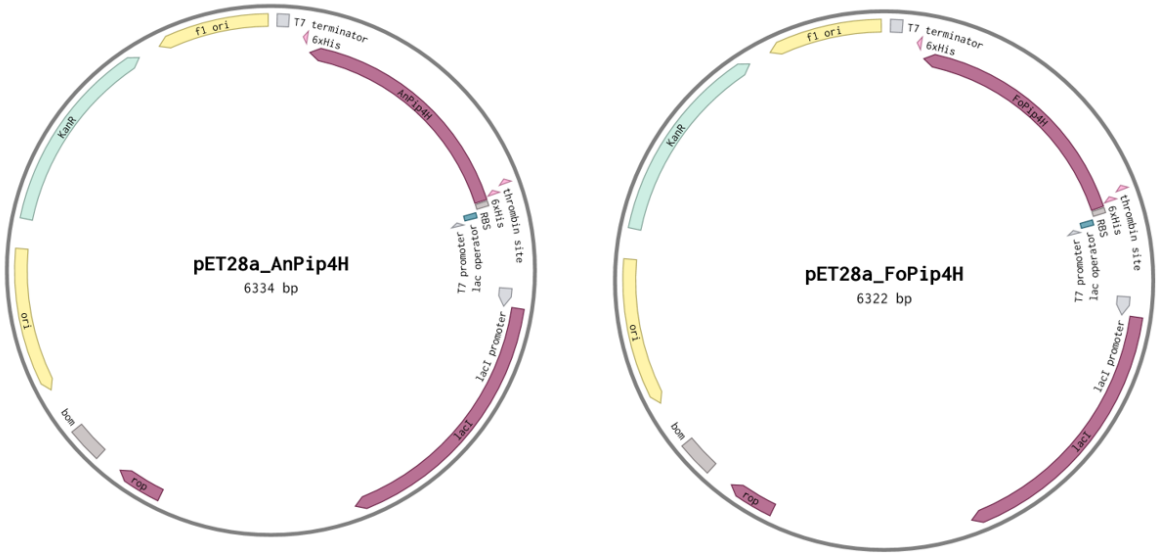

Figure 39: Vector map of construct pET28a_AnPip4H and pET28a_FoPip4H created with Benchling.

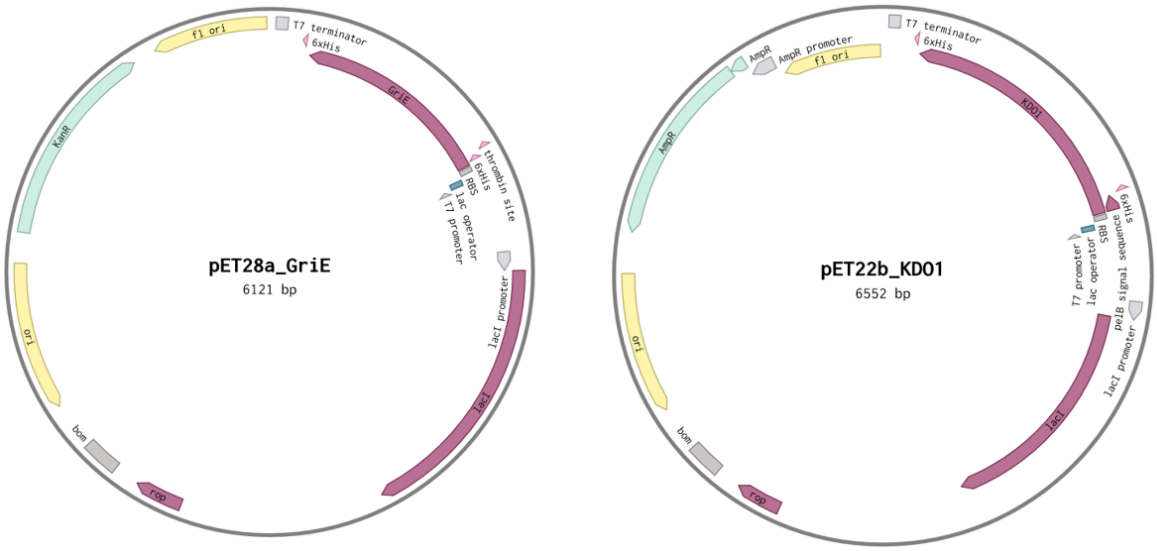

Figure 40: Vector map of construct pET28a_GriE and pET22b_KDO1 created with Benchling.

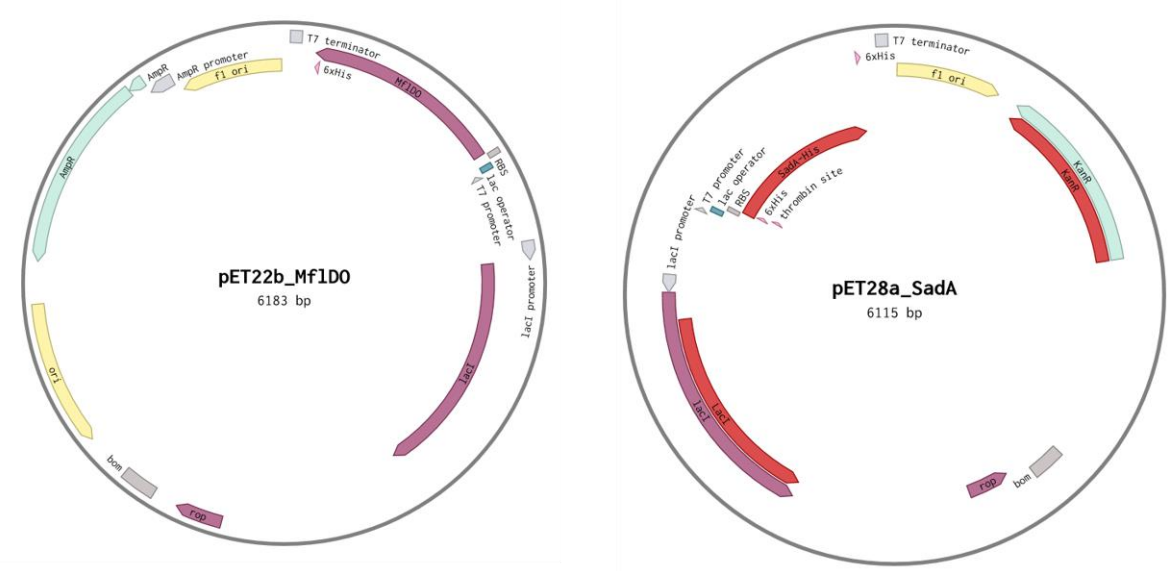

Figure 41: Vector map of construct pET22b_MflDO and pET28a_SadA created with Benchling.

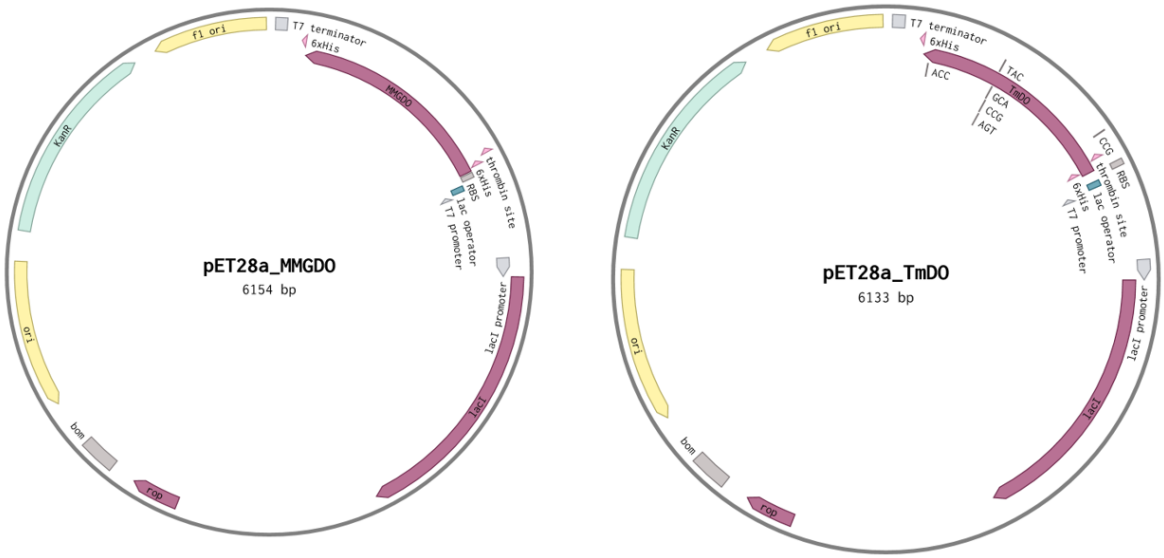

Figure 42: Vector map of construct pET28a_MMGDO and pET28a_TmDO created with Benchling.

7.7 Activity assay of different dioxygenases

AvLDO:

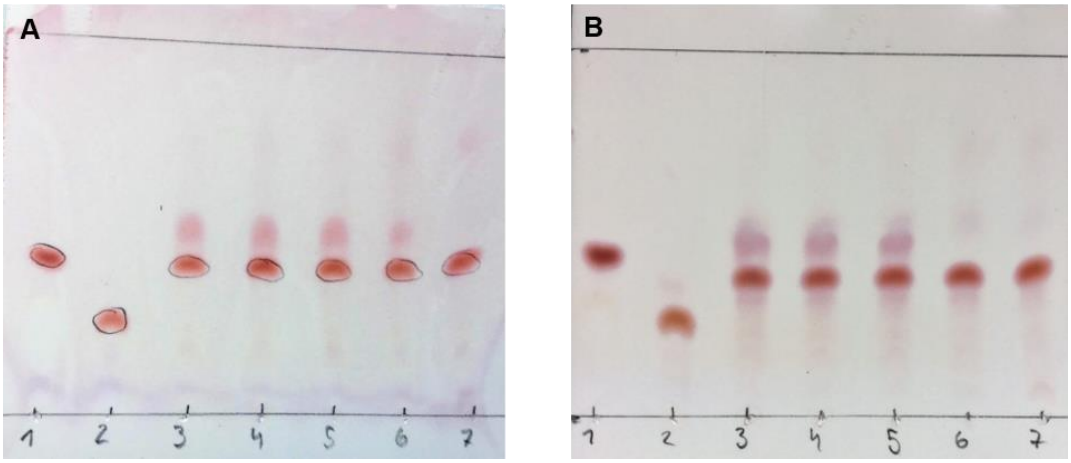

Figure 43: Biotransformations of AvLDO with L-leucine as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE and with purified protein. L-leucine and hydroxy isoleucine were used as standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A]** CFE: 0.7 mg/mL **[B]** purified protein: 1 mg/mL in 500 µL volume at 25 °C. **[A]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_2; [5] t_5; [6] t_{30}; **[B]** [1] L-leucine; [2] hydroxy isoleucine; [3] t_0; [4] t_2; [5] t_5; [6] t_{30}.

MMGDO:

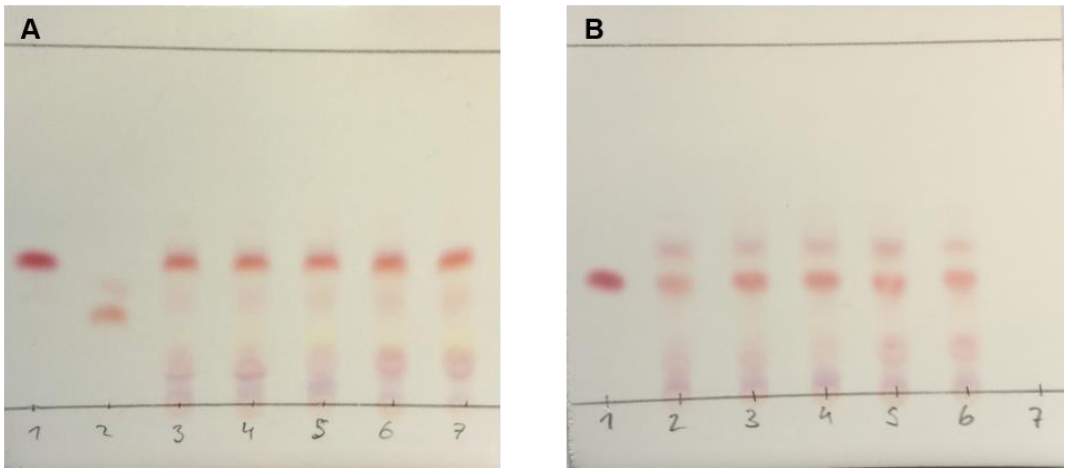

Figure 44: Biotransformations of MMGDO with L-isoleucine **[A]** and L-methionine **[B]** as substrate analyzed by TLC and detected with ninhydrin. Reactions were performed with CFE. L-isoleucine, L-methionine and hydroxy isoleucine served as standards. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$ in 500 µL volume at 30 °C. **[A]** [1] L-isoleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] $t_{O/N}$; [7] t_{43} **[B]** [1] L-methionine; [2] t_0; [3] t_1; [4] t_2; [5] $t_{O/N}$; [6] t_{43}; [7] negative control

TmDO:

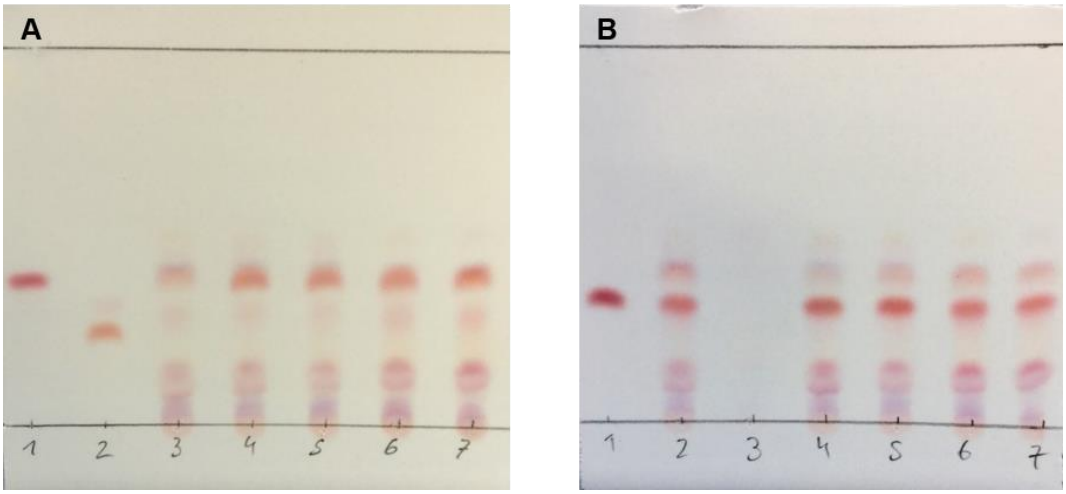

Figure 45: Biotransformations of TmDO with L-isoleucine **[A]** and L-methionine **[B]** as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE. L-isoleucine, L-methionine and hydroxy isoleucine served as standards. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$ in 500 µL volume at 30 °C. **[A]** [1] L-isoleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] $t_{O/N}$; [7] t_{43} **[B]** [1] L-methionine; [2] t_0; [3] negative control; [4] t_1; [5] t_2; [6] $t_{O/N}$; [7] t_{43}.

SadA:

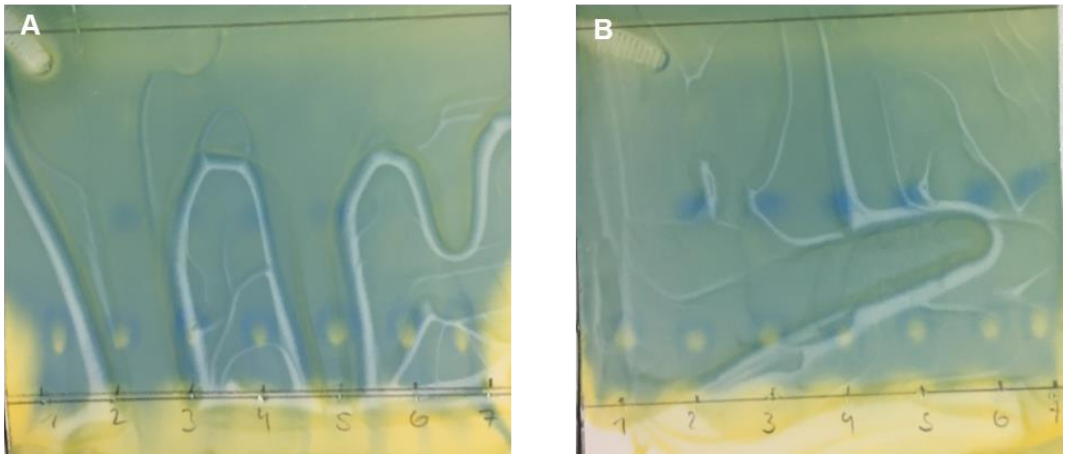

Figure 46: Biotransformations of SadA with N-succinyl-valine as substrate analyzed by TLC and detected with bromocresol green. Reactions were performed with CFE and with purified protein. N-succinyl-valine was served as a standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM FeSO4x7H2O, **[A]** CFE: 1.9 mg/mL **[B]** purified protein: 2 mg/mL in 500 µl volume at 35 °C. **[A]** [1] N-succinyl-valine; [2] t_0; [3] t_{30}; [4] t_1; [5] $t_{O/N}$; [6] t_{24}; [7] t_{48}; [**B]** [1] N-succinyl-valine; [2] t_0; [3] t_{30}; [4] t_1 [5] $t_{O/N}$; [6] t_{24}; [7] t_{48}.

KDO1:

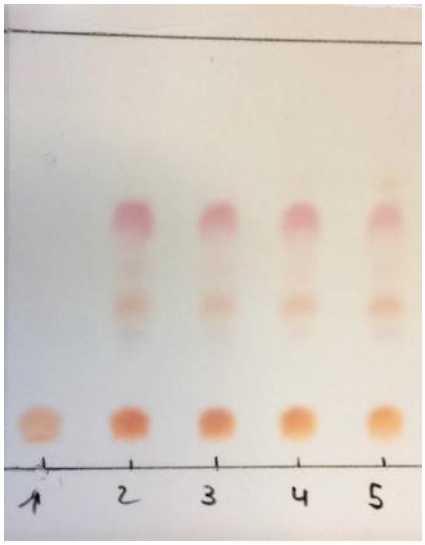

Figure 47: Biotransformations of KDO1 with L-lysine as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE. L-lysine served as a standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, CFE: 0.12 mg/ml in 500 µL volume at 20 °C. [1] L-lysine; [2] t_0; [3] t_1; [4] t_2 [5] $t_{O/N}$;

P4H:

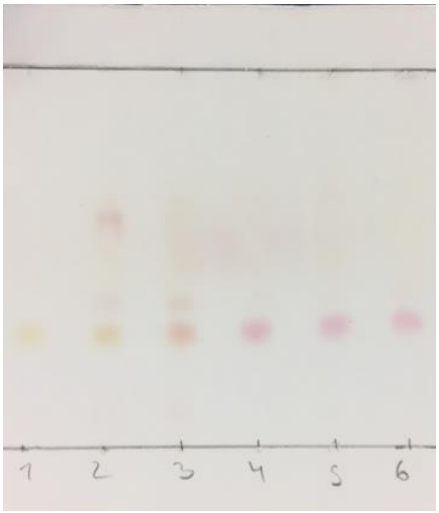

Figure 48: Biotransformations of P4H with L-proline as a substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with CFE. L-proline was used as a standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, CFE: 0.14 mg/mL in 500 µL volume at 35 °C. [1] L-proline; [2] t_0; [3] t_1; [4] t_2 [5] $t_{O/N}$; [6] t_{24};

AnPip4H:

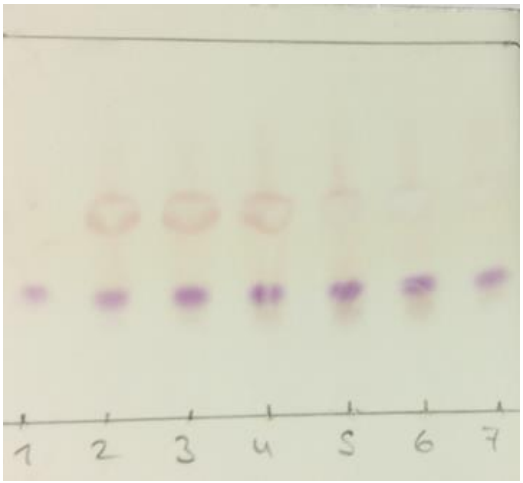

Figure 49: Biotransformations with AnPip4H with L-pipecolic acid as a substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with purified protein. L-pipecolic acid served as standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, purified protein: 1.2 mg/mL in 500 µL volume at 20 °C. [1] L-pipecolic acid; [2] t_0; [3] t_1; [4] t_2 [5] t_6; [6] t_{24}; [7] t_{68}.

Table 36: R_f-values of all L-amino acids and standards used in this work.

Compound	TLC condition	R_f-value
L-isoleucine	Mobile phase A	0.62
	Mobile phase B	0.32
L-leucine	Mobile phase A	0.62
	Mobile phase B	0.36
Hydroxy isoleucine	Mobile phase A	0.48
	Mobile phase B	0.20
L-lysine	Mobile phase A	0.07
L-pipecolic acid	Mobile phase A	0.32
N-succinyl-L-valine	Mobile phase B	0.13
L-methionine	Mobile phase B	0.34

7.8 Expression optimization of AvLDO

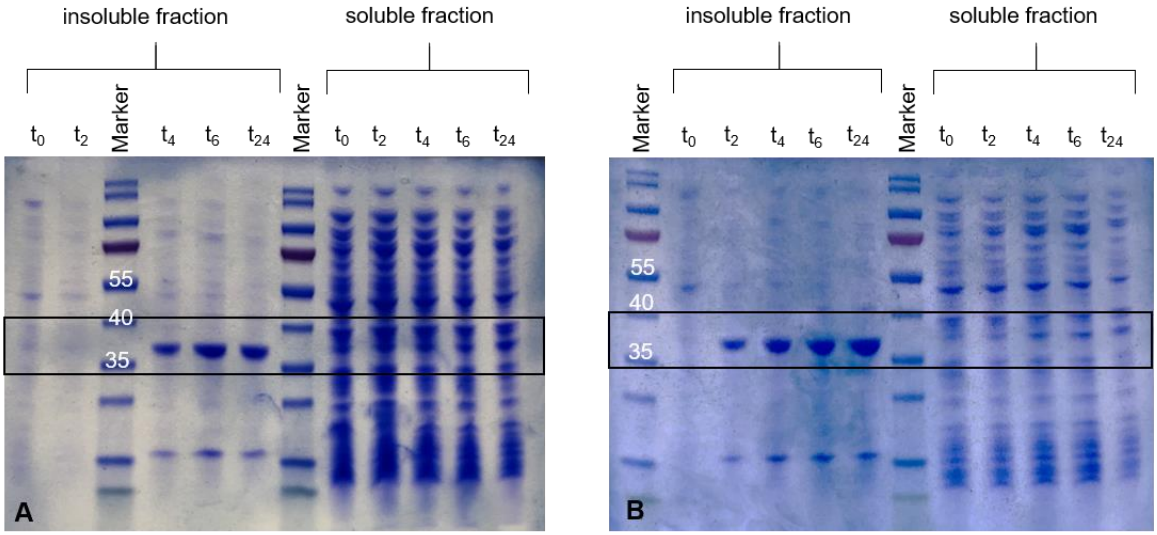

Figure 50: Expression of AvLDO (33 kDa) at a cultivation temperature of 37 °C and induction at OD_{600} of 0.6 **[A]** and 1.2 **[B]** with 0.1 mM IPTG. PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

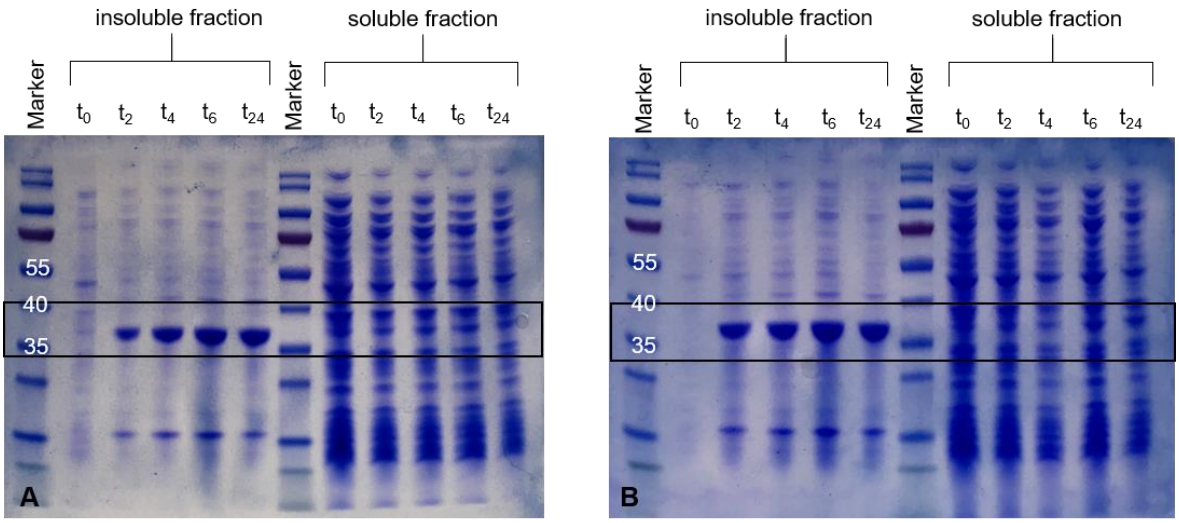

Figure 51: Expression of AvLDO (33 kDa) at a cultivation temperature of 37 °C and induction at OD_{600} of 0.6 **[A]** and 1.2 **[B]** with 1 mM IPTG. PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

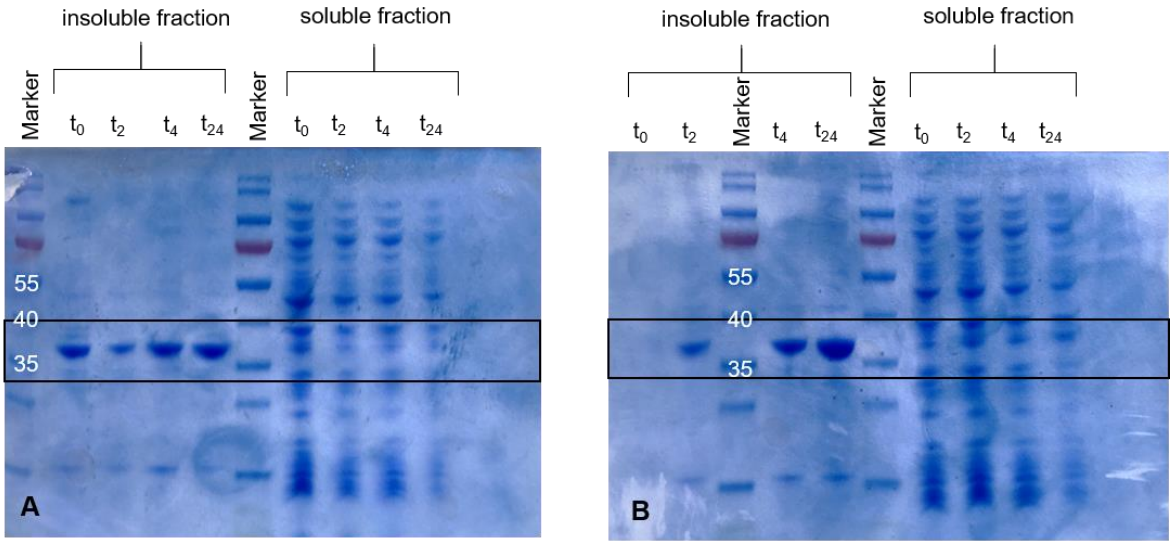

Figure 52: Expression of AvLDO (33 kDa) at a cultivation temperature of 28 °C and induction at OD_{600} of 0.6 **[A]** and 1.2 **[B]** with 0.5 mM IPTG. PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

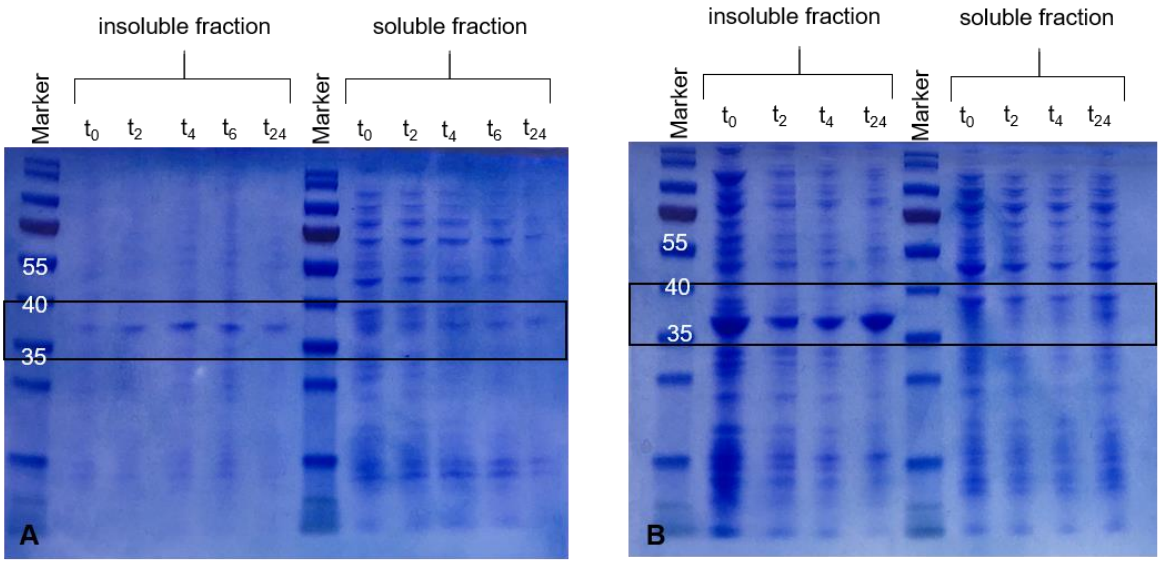

Figure 53: [A] Expression of AvLDO (33 kDa) in *E. coli* BL21(DE3) with co-expression of groES-groEL. **[B]** Expression of AvLDO with the addition of $FeSO_4$. PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

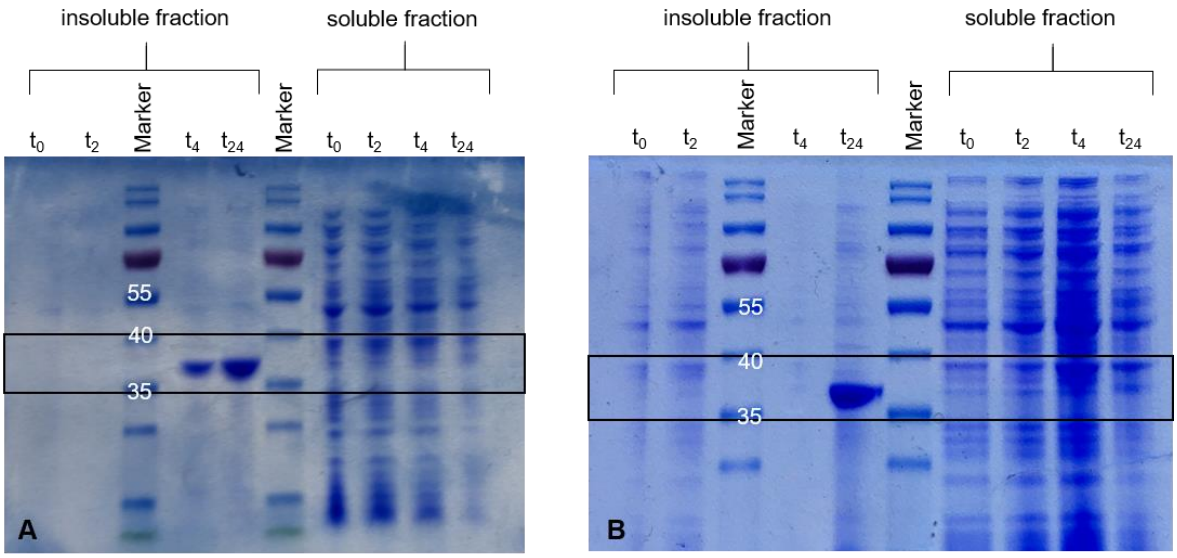

Figure 54: Autoinduction of AvLDO (33 kDa) in **[A]** *E. coli* BL21(DE3) and **[B]** *E. coli* JM109(DE3). PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

7.9 Activity assay of AvLDO

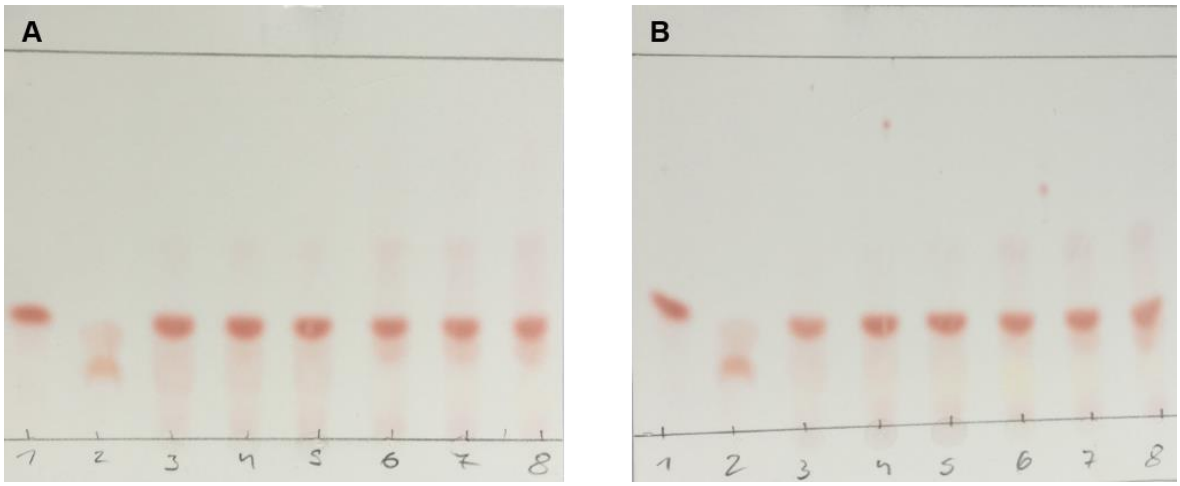

Figure 55: Biotransformations of AvLDO with L-isoleucine as substrate analyzed by TLC and detected with ninhydrin staining. Reactions were performed with purified protein. L-isoleucine and hydroxy isoleucine were used as standard. Final reaction conditions: 5 mM substrate, 7.5 mM α-ketoglutarate, 5 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, **[A]** purified protein: 0.25 mg/mL **[B]** purified protein: 0.5 mg/mL in 500 µL volume at 25 °C. **[A]** [1] L-isoleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18}; [7] t_{24}; [8] t_{48} **[B]** [1] L-isoleucine; [2] hydroxy isoleucine; [3] t_0; [4] t_1; [5] t_2; [6] t_{18} ;[7] t_{24}; [8] t_{48}