Selina Pieber, BSc.

Engineering of Dioxygenases

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, August 2020

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

At this point, I would like to express my special gratitude to my supervisor Univ.-Prof. Dr.rer.nat Robert Kourist for offering me this project at the Institute of Molecular Biotechnology in his working group Biocatalysis and Protein Engineering. Hereby, he enabled me to conduct my master thesis in a very interesting field of biotechnology and to gain valuable insights into an exciting topic of research.

My honest thanks go to my supervisor Anna Schweiger, MSc who accompanied me throughout my entire time at the institute. Besides her professional guidance and her patience, she always encouraged me to work independently and make my own suggestions. Especially, during the Covid-19-required shift work she was available for answering all my questions around the clock effortlessly and also enabled me to conduct all my planned experiments despite this exceptional situation. Furthermore, I really appreciate the time she invested in proofreading my thesis and all the useful advices and scientific input she gave to me.

I am also very grateful for the support of the whole working group. Special thanks go to Maria Schabhüttl for always assisting me in the everyday lab routine. Further, I want to thank Bastian Kremser who started at the same time with his master thesis at the institute and always encouraged me on my journey.

Finally, I gratefully thank my boyfriend Manuel Schnetzinger and my parents Annemarie and Andreas Pieber who encouraged me during my whole studies, always believed in my skills and supported me in any way they could.

Thank you all.

Abstract

α-Ketoglutarate dependent dioxygenases belong to the class of non-heme iron(II) dependent dioxygenases and catalyse a variety of chemical reactions important for industrial applications. A common reaction catalysed by this enzyme superfamily is the stereoselective hydroxylation of L-amino acids. As enantiopure hydroxy amino acids are valuable building blocks for the organic synthesis of pharmaceuticals, a variety of these enzymes has been discovered in recent years.

Within this study, recombinant production of the Leucine dioxygenase from *Anabaena variabilis* (AvLDO), the Isoleucine dioxygenase from *Bacillus thuringiensis* (BtIDO) and the Leucine dioxygenase from *Streptomyces* strain DSM40835 (GriE) were studied in detail employing the triple-knockout strain *Escherichia coli* Δ*sucA*Δ*aceA*Δ*putA* (3Δ*sucA*).

The knockout of the α-ketoglutarate dehydrogenase and isocitrate lyase encoding genes prevents succinate production and leads to an interruption of the TCA cycle, which renders the resulting strain nonviable. The desired hydroxylation reaction performed by Fe(II)/α-KG dependant dioxygenases can restore growth by the formation of succinate as a co-product. As part of this work, a growth-based selection assay employing the respective 3Δ*sucA* strain could by established and verified by using three dioxygenases AvLDO, BtIDO and GriE with different activity and substrate scope.

In this study, the established assay was applied for the selection of improved AvLDO variants. In detail, a mutant library of AvLDO was created to extend the narrow substrate scope of the enzyme towards L-isoleucine. Simultaneous saturation mutagenesis of two positions in close proximity to the substrate binding site led to the creation of 35 different amino acid variants. An iterative cultivation approach in liquid minimal medium supplemented with L-isoleucine was followed by solid cultivation on minimal medium agar plates. It enabled the selection of clones with larger colony size, which is assumed to correspond to a higher enzyme activity. In total, 14 clones were picked and analysed including sequencing, re-screening, and expression in minimal medium. One specific genotype was detected eight times. It showed only one mutation, specifically an exchange of isoleucine by threonine. Three variants, including the highly abundant variant described before, showed earlier exponential growth and at least 3-fold higher specific growth rates than the

wild type in the re-screening. Therefore, these variants are assumed to be higher active towards L-isoleucine. Further, verification of L-isoleucine conversion was attempted by an in-vitro assay and qualitative thin layer chromatography analysis.

Zusammenfassung

α-Ketoglutarat-abhängige Dioxygenasen gehören zu der Klasse der Nicht-Häm-Fe(II)-abhängigen Dioxygenasen und katalysieren eine Vielfalt chemischer Reaktionen die für industrielle Anwendungen von Bedeutung sind. Eine häufig katalysierte Reaktion dieser Enzym-Superfamilie ist die stereoselektive Hydroxylierung von L-Aminosäuren. Da enantiomerenreine Hydroxy-Aminosäuren wertvolle Bausteine für die organische Synthese von Pharmazeutika darstellen, wurde in den letzten Jahren eine Vielfalt dieser Enzyme entdeckt.

Im Rahmen dieser Arbeit wurden die rekombinante Produktion der Leucin Dioxygenase von *Anabaena variabilis* (AvLDO), der Isoleucin Dioxygenase von *Bacillus thuringiensis* (BtIDO) sowie der Leucin Dioxygenase vom Stamm *Streptomyces* DSM40835 (GriE) im Detail untersucht. Dafür wurde der dreifache Knockout-Stamm *Escherichia coli* Δ*sucA*Δ*aceA*Δ*putA* (3Δ*sucA*) verwendet.

Das Ausschalten von α-Ketoglutarat Dehydrogenase und Isocitrat Lyase codierenden Genen verhindert die Succinat-Produktion und führt zu einer Unterbrechung des Citratzyklus, was den resultierenden Stamm nicht lebensfähig macht. Die erwünschte Hydroxylierung, die von Fe(II)/α-Ketoglutarat-abhängigen Dioxygenasen katalysiert wird, kann den Citratzyklus durch die Bildung des Co-Produkts Succinat wiederherstellen. Im Rahmen dieser Arbeit konnte ein wachstumsbasierter Selektions-Assay, unter Einsatz des entsprechenden 3Δ*sucA* Stamms entwickelt werden. Durch die Verwendung von drei verschiedenen Dioxygenasen AvLDO, BtIDO und GriE, die eine unterschiedliche Aktivität sowie ein unterschiedliches Substratspektrum aufweisen, konnte der Assay verifiziert werden.

In dieser Arbeit wurde der entwickelte Selektions-Assay verwendet, um verbesserte AvLDO-Varianten zu detektieren. Im Einzelnen wurde eine Bibliothek der AvLDO erstellt, um das schmale Substratspektrum zu erweitern. Simultane Sättigungsmutagenese zweier Positionen, die sich in der Nähe der Substrat-Bindungsstelle befinden, führten zu Bildung 35 verschiedener Aminosäurevarianten. Ein iterativer Kultivierungsansatz in flüssigem Minimalmedium wurde auf Platten aus festen Minimalmedium überstrichen. Dies ermöglichte die Selektion von größeren Klonen, von welchen vermutet wird, dass sie Varianten mit höheren Enzymaktivitäten tragen. Insgesamt wurden 14 Klone ausgewählt und im Detail analysiert. Diese Analyse beinhaltete Sequenzierung, Re-

Screening sowie Expression in Minimalmedium. Ein spezifischer Genotyp wurde achtmal entdeckt. Dieser wies nur eine Mutation, genauer, einen Austausch von Isoleucin durch Threonin auf. Neben dieser Variante wurden noch zwei weitere Varianten entdeckt, welche im durchgeführten Re-Screening ein früheres exponentielles Wachstum und eine mindestens dreifach höhere spezifische Wachstumsrate im Vergleich zum Wildtyp-Enzym gezeigt haben. Von diesen Varianten wird angenommen, dass sie eine höhere Enzymaktivität gegenüber L-Isoleucin aufweisen. Weiters wurde versucht, die Umsetzung von L-Isoleucin mittels eines in-vitro-Assays und einer qualitativen Dünnschichtchromatographie-Analyse zu verifizieren.

List of abbreviations

3DM	Three-dimensional molecular class specific information systems
(2*S*,3*R*,4*S*)-HIL	(2*S*,3*R*,4*S*)-4-hydroxyisoleucine
2-OG	2-Oxoglutarate
AceA	Isocitrate lyase
AceK	Isocitrate dehydrogenase kinase/phosphatase
AMKP	(2S,3*R*)-2-amino-3-methyl-4-ketopentanoate
Amp	Ampicillin
AvLDO	Leucine dioxygenase from *Anabaena variabilis*
bp	Base pairs
BrnQ	Branched-chain amino acid transporter
BtIDO	Isoleucine dioxygenase from *Bacillus thuringiensis*
CASTing	Combinatorial active-site saturation testing
***cis*-3-Hypip**	*cis*-3-hydroxy-L-proline
***cis*-5-Hypip**	*cis*-5-hydroxy-L-proline
CFE	Cell free extract
CFU	Colony forming units
ddH_2O	Double-distilled water
dH_2O	Distilled water
dNTPs	Nucleoside triphosphate
DSBH	Double-stranded β-helix fold
EC	Enzyme class
EDTA	Ethylenediaminetetraacetic acid
fwd	forward
g	Relative centrifugal compared to gravity
GABA	γ-aminobutyric acid
GriE	Leucine dioxygenase from Streptomyces strain DSM 40835
HPLC	High performance liquid chromatography
HTSOS	High throughput screening or selection
IPTG	Isopropanyl β-D-1 thiogalactopyranoside
ISM	Iterative saturation mutagenesis
kDa	Kilo Dalton
Kan	Kanamycin
lac	Lactose promoter
LB	*Lysogeny* broth
LdoA	L-Leucine 5-hydroxylase from *Nostoc punctiforme*
L-DOPS	L-*threo*-3,4-dihydroxyphenylserine
L-Pip	L-pipecolic acid
MOPS	3-(*N*-morpholino)propanesulfonic acid
NSDOPS	*N*-succinyl-L-*threo*-3,4-dimethoxyphenylserine
NSDOPA	*N*-succinyl-3,4-dimethoxyphenylalanine
NpLDO	L-Leucine 5-hydroxylase from *Nostoc punctiforme*
OD_{600}	Optical density at 600 nm
ONC	Overnight culture
P4H	Proline-4-hydroxylase
PCR	Polymerase chain reaction
PutA	Proline dehydrogenase
rpm	Rotation per minute
rev	Reverse

SadA	*N*-succinyl amino acid hydroxylase from *Burkholderia ambifaria* AMMD
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SmPH	Proline hydroxylase from *Sinorhizobium meliloti*
SOC	Super optimal broth with catabolite repression
SucA	α-KG dehydrogenase E1 subunit
SucC	sucCoA synthetase β subunit
SucCoA	Succinyl-coenzyme A
SyrB1	Syringomycin biosynthesis enzyme 1
SyrB2	Syringomycin biosynthesis enzyme 2
TAE	TRIS-acetate-EDTA
TauD	α-ketoglutarate-dependent taurine dioxygenase
TB	Terrific broth
TCA	Tricarboxylic acid cycle
tet	Tetracycline promoter
TLC	Thin layer chromatography
Tris	Tris(hydroxymethyl)aminomethane
TY	Tryptone, yeast extract
U	Enzymatic units
α-KG	α-ketoglutarate
(αMe)Phe	α-methylphenylalanine
μ	specific growth rate

Table of contents

1 Introduction

1.1 Oxidoreductases and their biocatalytic application

The activation of inert carbon-hydrogen (C-H) bonds in a regio- and stereoselective manner represents a major challenge in synthetic organic chemistry (Wu *et al*., 2016; Turner and Humphreys, 2018).

In contrast, nature provides several highly selective oxyfunctionalisation catalysts. Oxidoreductases comprising, among others, dehydrogenases, oxygenases, oxidases and peroxidases (Figure 1) are widely distributed in nature amongst microbes, plants and animals. Generally, an exchange of electrons or redox equivalents between a donor and an acceptor molecule is catalysed by this group of enzymes. Key steps of oxidoreductase-catalysed reactions for instance involve proton extraction, electron transfer, hydride transfer or oxygen insertion. For the accomplishment of their physiological function, oxidoreductases use a variety of redox active centres including amino acid residues, coenzymes and metal ions or complexes (Munro *et al*., 2000; Xu, 2005). Advantages of industrial oxidoreductase biocatalysts are their biodegradability, their specificity and their high energy-saving potential promoting the development of environmentally-friendly, sustainable and highly efficient industries (Xu, 2005).

From the synthetic point of view, oxygenases are attractive enzymes as they are generating new functional groups by the insertion of oxygen into C–H, C–C and C=C bonds (Blank *et al*., 2010). They catalyse the incorporation of molecular oxygen into the substrate by using different mechanisms. Whereas monooxygenases (EC 1.13.-) incorporate one oxygen atom into the substrate, which requires a cofactor like NADPH for the reduction of the second oxygen to form water, dioxygenases (EC 1.14.-) are capable of simultaneously incorporating both atoms by forming a peroxo species (Dawson, 1988; Gunsalus *et al*., 1975; Hou, 1986).

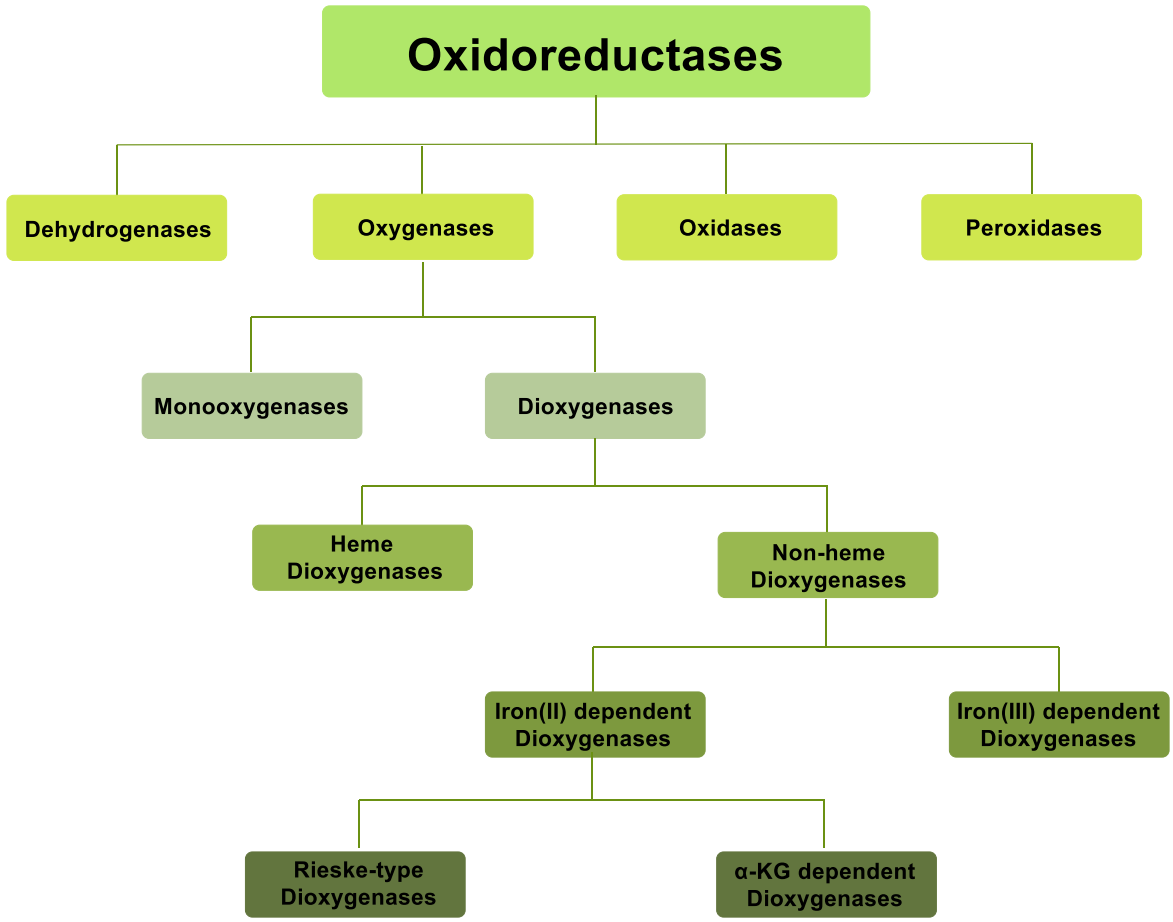

Figure 1: Overview of different types of oxidoreductases. Schematic illustration based on (Dong *et al.*, 2018; Faber, 1992; Gamenara *et al.*, 2012)

1.2 α-Ketoglutarate dependent dioxygenases

Two different classes of non-heme dioxygenases are known, including non-heme iron(II) dependent dioxygenases and iron(III) dependent dioxygenases (Figure 1). The former, including Rieske-type dioxygenases and α-ketoglutarate or 2-oxoglutarate dependent dioxygenases (α-KGDs or 2-OGs), activate O_2 directly, whereas the latter, comprising intradiol dioxygenases and lipoxygenases, activate the substrate (Gamenara *et al*., 2012).

Fe(II)/α-ketoglutarate dependent dioxygenases are widely distributed in various organisms including bacteria, fungi, plants and vertebrates (Wu *et al*., 2016). The first enzyme capable of hydroxylating the prolyl residue in collagen biosynthesis was reported in the 1960s demonstrating that ferrous iron and α-ketoglutarate are essential for the activity (Hutton *et al*., 1967).

Theoretically, α-ketoglutarate dependent oxygenases can act on various organic substrates. Besides amino acids, it is known that the enzyme class accepts DNA, RNA, proteins, oligosaccharides, lipids, non-ribosomal polypeptides, polyketides,

alkaloids and terpenoids (Walsh *et al*., 2013). They catalyse a variety of biochemical reactions and are involved in chromatin protein modification, chromatin DNA modification, control of transcription, RNA splicing control, translation and protein structure/stability (Herr and Hausinger, 2018). α-Ketoglutarate dependent oxygenases are involved in several biosynthetic pathways generating a high diversity of metabolites, more specifically intermediates of antibiotics and alkaloids (Hibi *et al*., 2016). Besides biosynthesis, it has been reported that α-ketoglutarate dependent oxygenases are further involved in biodegradation by decomposition of their substrates for metabolite recycling (Li *et al*., 2015).

1.2.1 Industrial application

Since their discovery in the 1960s, Fe(II)/α-ketoglutarate dependent dioxygenases have gained increasing attention for their ability of carbon-hydrogen oxyfunctionalisation (Herr and Hausinger, 2018; Islam *et al*., 2018). This strategy can either be exploited for the synthesis of early stage building blocks or for the functionalisation of an advanced intermediate at a later stage (Figure 2). The former requires a high catalytic efficiency and inexpensive cosubstrates and cofactors to produce an adequate quantity of a key intermediate, whereas for the production of the latter substrate promiscuity is more important rather than a high turnover number. Whereas late-stage functionalisation is mainly done by versatile cytochrome P450 monooxygenases, Fe(II)/α-ketoglutarate dependent dioxygenases are excellent candidates for building block synthesis as they do not require expensive cofactors or dedicated reductase partners (Li *et al*., 2019).

Amino acids belong to the most important natural products. In their hydroxylated form they serve as valuable building blocks for the pharmaceutical industry (Nájera and Sansano, 2007, Zwick and Renata, 2018a). Therefore, the main focus of biocatalytic applications was placed on the Fe(II)/α-ketoglutarate dependent amino acid hydroxylases for the microbial production of chiral building blocks as hydroxy-proline or -leucine for the use in chemical industry so far (Hüttel, 2013). The substrate hydroxylation at an unactivated carbon centre is the best studied and established reactivity of Fe(II)/α-KGDs. Besides hydroxylation, a variety of other oxidative transformations including halogenation, desaturation, epimerization, epoxidation and ring expansion/closure can be performed by this enzyme class (Hausinger, 2015).

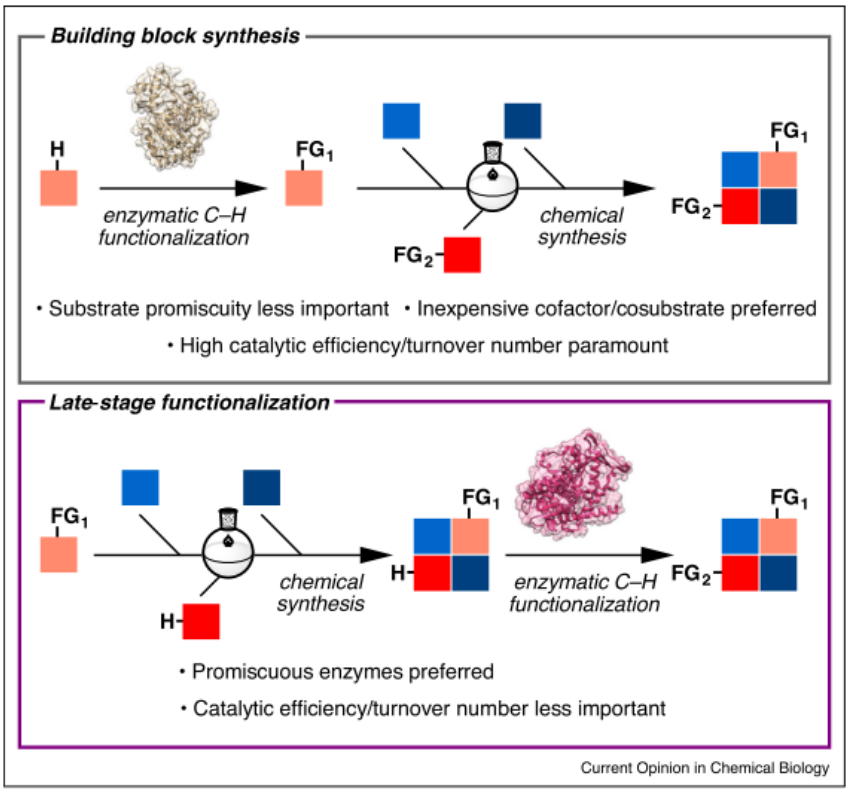

Figure 2: Biocatalytic C–H activation for building block synthesis and late-stage functionalisation in multi-step synthesis. Figure taken from (Li *et al*., 2019).

1.2.2 Reaction mechanism

According to crystallographic studies on various Fe(II)/α-ketoglutarate dependent oxygenases, two structural features are shared amongst their superfamily members. The Fe(II) in the active site is coordinated to the 2-His-1-carboxylate facial triad (Figure 3), which is a common feature in numerous non-heme iron(II) enzymes (Hegg and Que, 1997). The carboxylate can originate from either a Glu or an Asp residue. Further, the structural 2-His-1-carboxylate is surrounded by a characteristic double-stranded β-helix (DSBH) fold, also named jelly-roll (Aik *et al*., 2012; Aik *et al*., 2015; Martinez and Hausinger, 2015). The DSBH motif consists of two four-stranded antiparallel β-sheets forming a sandwich structure proposed to assist in selective binding of the primary substrate besides supporting the active site of the enzyme (Aik *et al*., 2012).

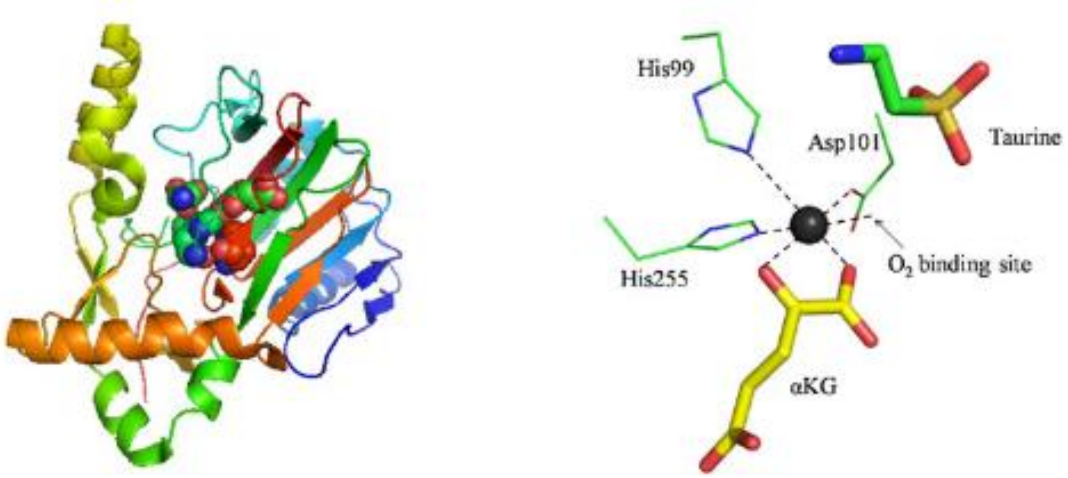

Figure 3: α-Ketoglutarate dependent dioxygenase TauD. Overall fold of the enzyme on the left and the His-Asp-His "facial triad" in TauD-α-KG-taurine complex structure (PDB ID: 1OS7) on the right. The iron centre is represented by the filled black circle. Figure taken from (Wu *et al*., 2016).

Initially, the Fe(II) is bound to the 2-His-1-carboxylate motif with three water molecules occupying the additional coordination sites (Figure 4 (a)). α-Ketoglutarate displaces two water molecules by chelating the Fe(II) with its keto group (Figure 4 (b)). By binding the "prime" substrate near the metallocentre, the third water molecule is displaced creating an O_2-binding site (Figure 4 (c)). The Fe(III)-superoxo species created by the binding of dioxygen (Figure 4 (d)) is then transformed into a Fe(IV)-oxo (ferryl) species with bound α-ketoglutarate (Figure 4 (e)).This common ferryl intermediate is utilised by all subsequent reactions *e.g.* hydroxylation, desaturation or cyclisation. The cleavage of the O-O bond leads to decarboxylation of α-ketoglutarate. The newly generated succinate is then bound to the Fe(IV)-oxo complex (Figure 4 (f)). The abstraction of a hydrogen atom from the primary substrate by the ferryl intermediate forms a Fe(III)-OH (Figure 4 (g)) and a substrate radical. By rebinding the OH-group to the substrate (Figure 4 (h)), the hydroxylated product and succinate is released concluding the catalytic cycle (Hüttel, 2013; Martinez and Hausinger, 2015; Mitchell *et al*., 2018).

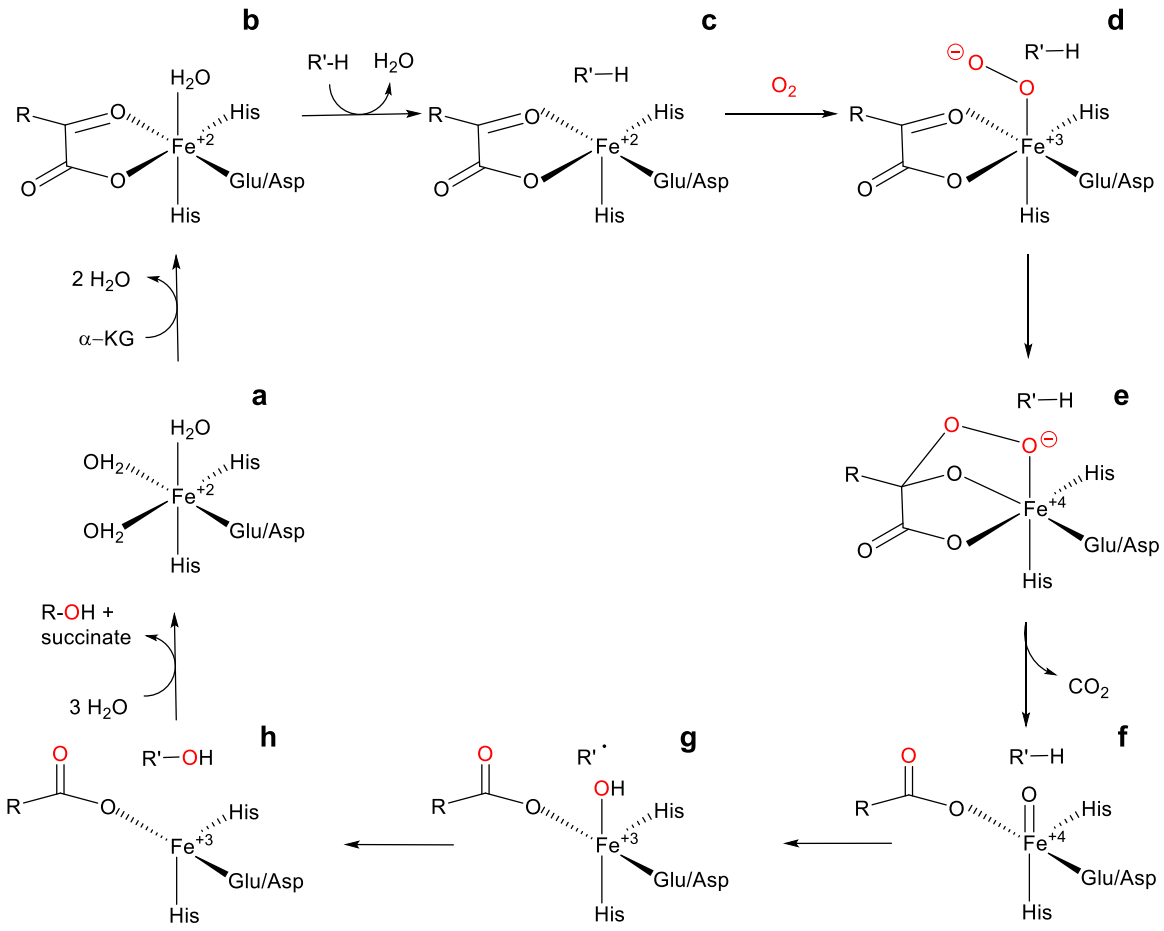

Figure 4: Common mechanism for hydroxylations with α-ketoglutarate dependent dioxygenases. Figure taken and adapted from (Martinez and Hausinger, 2015).

Beyond hydroxylation, a remarkable subgroup of Fe(II)/α-KG dependent enzymes is capable of catalysing halogenation reactions constructing a C–X bond (X = Cl− or Br−) by oxidative activation of the C–H bond (Neumann *et al*., 2008; Vaillancourt *et al*., 2006).

In 2005, the first Fe(II)/α-ketoglutarate dependent halogenase from *Pseudomonas syringae* B301D called SyrB2 was identified to be involved in the biosynthesis of syringomycin, a non-ribosomal peptide with antifungal activities (Iacobellis *et al*., 1992; Vaillancourt *et al*., 2005). Thereby, SyrB2 chlorinates the substrate L-tyrosine, which is tethered to SyrB1 at the methyl group. The enzyme requires the presence of Fe(II), oxygen, α-KG, and Cl^- to fulfil its function. Instead of the characteristic His^1-X-Asp/Glu-X_n-His^2 motif, structural studies on SyrB2 showed the iron to be coordinated by two histidine residues and a chloride atom (Blasiak *et al*., 2006). Further, the conserved aspartate or glutamate in the facial triad is exchanged by alanine. As alanine is smaller than aspartate or glutamate, more space is available for the chloride binding. The first three steps catalysed by α-ketoglutarate dependent

dioxygenases (Figure 4 (a,b,c)) are also performed by SyrB2. Then, the pathway diverges and instead of oxygen, a halogen is bound creating a haloferryl species (Figure 5 (d)). By abstraction of a hydrogen from the “prime” substrate, a *cis*-halohydroxo-ferric species and a substrate radical are formed (Figure 5 (e)). Then, the chloride atom is transferred to the substrate radical yielding a halogenated product (Figure 5 (f)) (Martinez and Hausinger, 2015; Pratter *et al*., 2014).

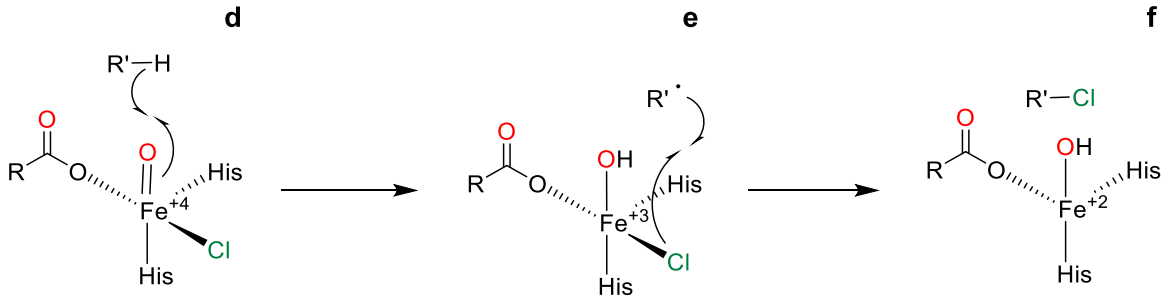

Figure 5: SyrB2 mediated halogenation reaction. Figure taken and adapted from (Martinez and Hausinger, 2015).

1.3 Different types of α-ketoglutarate dependent dioxygenases

Three selected dioxygenases investigated in this master thesis are presented in detail below.

1.3.1 L-Leucine dioxygenase from *Anabaena variabilis* (AvLDO)

The δ-specific L-leucine dioxygenase from the filamentous cyanobacterium *Anabaena variabilis* was recently cloned and characterised by Cordeiro *et al*. (Correia Cordeiro *et al*., 2018). The investigation of the substrate scope of AvLDO revealed that the enzyme is capable of hydroxylating L-leucine and L-norleucine with C5-specificity generating the respective intermediates (2*S*,4*S*)-5-hydroxyleucine and -norleucine (Figure 6). Whereas the sulfoxidation of L-methionine by AvLDO produces the corresponding (*R*)-enantiomer (Correia Cordeiro *et al*., 2018), BtIDO from *Bacillus thuringiensis* (Chapter 1.3.2) produces the (*S*)-sulfoxid (Hibi *et al*., 2011). Besides methionine, AvLDO also converts ethionine but is incapable of converting structural similar substrates as L-isoleucine, valine and norvaline.

Hence, the substrate promiscuity of AvLDO is rather small, compared to the sequentially quite different γ-specific Fe(II)/α-ketoglutarate dependent dioxygenase from *Bacillus thuringiensis* (BtIDO) (Peters and Buller, 2019). Highest sequence identity was reported for the previously characterised L-leucine 5-hydroxylase from

Nostoc punctiforme (LdoA or NpLDO). Subsequently, the two enzymes LdoA and AvLDO share many properties as the substrate scope or the preferred cultivation conditions (Correia Cordeiro *et al*., 2018; Hibi *et al*., 2013).

Further interesting observations obtained by Correia Cordero *et al.* are the preference for acidic conditions and the optimum temperature of 25 °C, where highest specific activities of 0.22 ± 0.024 U/mg towards L-leucine were observed (Correia Cordeiro *et al*., 2018).

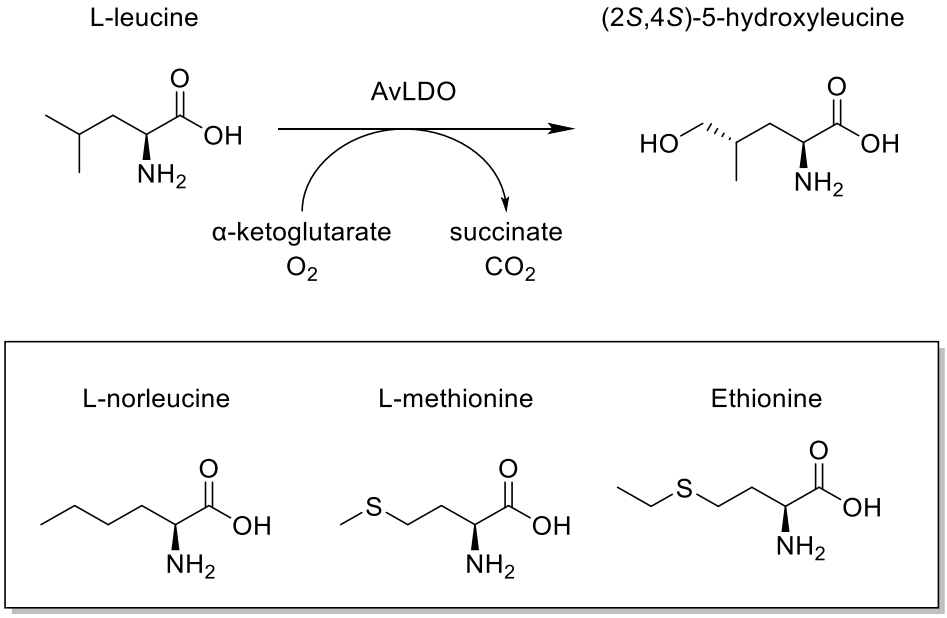

Figure 6: Oxyfunctionalisation of AvLDO towards L-leucine. Amino acids shown in the box are further accepted substrates. Figure taken and adapted from (Correia Cordeiro *et al*., 2018).

1.3.2 L-Isoleucine dioxygenase from *Bacillus thuringiensis* (BtIDO)

The first α-ketoglutarate-dependent dioxygenase acting on free aliphatic amino acids was discovered by Kodera *et al.* (Kodera *et al*., 2009). The L-Isoleucine hydroxylating enzyme was isolated from *B. thuringiensis* (BtIDO). Although there were already several known α-KG dependent dioxygenases hydroxylating free amino acids, BtIDO was the first enzyme with C4-specificity capable of generating the highly valuable (2*S*,3*R*,4*S*)-4-hydroxy-L-isoleucine (HIL) (Figure 7) (Hibi *et al*., 2011). HIL, first found in fenugreek seeds (Fowden *et al*., 1973), shows antidiabetic activity and is therefore a promising compound for drug development (Broca *et al*., 1999). The effective use of 4-HIL for the treatment of glycemia, insulinemia as well as body weight control was also reported (Jette *et al*., 2009). Further, by the

conversion of 4-HIL, the vitamin B12 antimetabolite (2*S*,3*R*)-2-amino-3-methyl-4-ketopentanoic acid (AMKP), showing antibiotic activity, can be produced by HIL (Hibi *et al*., 2011; Peters and Buller, 2019). This pathway from L-isoleucine via HIL to AMKP catalysed by L-isoleucine dioxygenase and HIL dehydrogenase is capable of shunting the incomplete tricarboxylic acid (TCA) cycle of *Bacillus thuringiensis*. The organism lacking α-KG conversion via the oxidative branch is only able to ensure succinate production via the γ-aminobutyrate pathway. This pathway converts α-KG into L-glutamate and further succinate being accompanied by the transamination of HIL and AMPK to form the corresponding keto-acids (Aronson *et al*., 1975; Hibi and Ogawa, 2014; Ogawa *et al*., 2011).

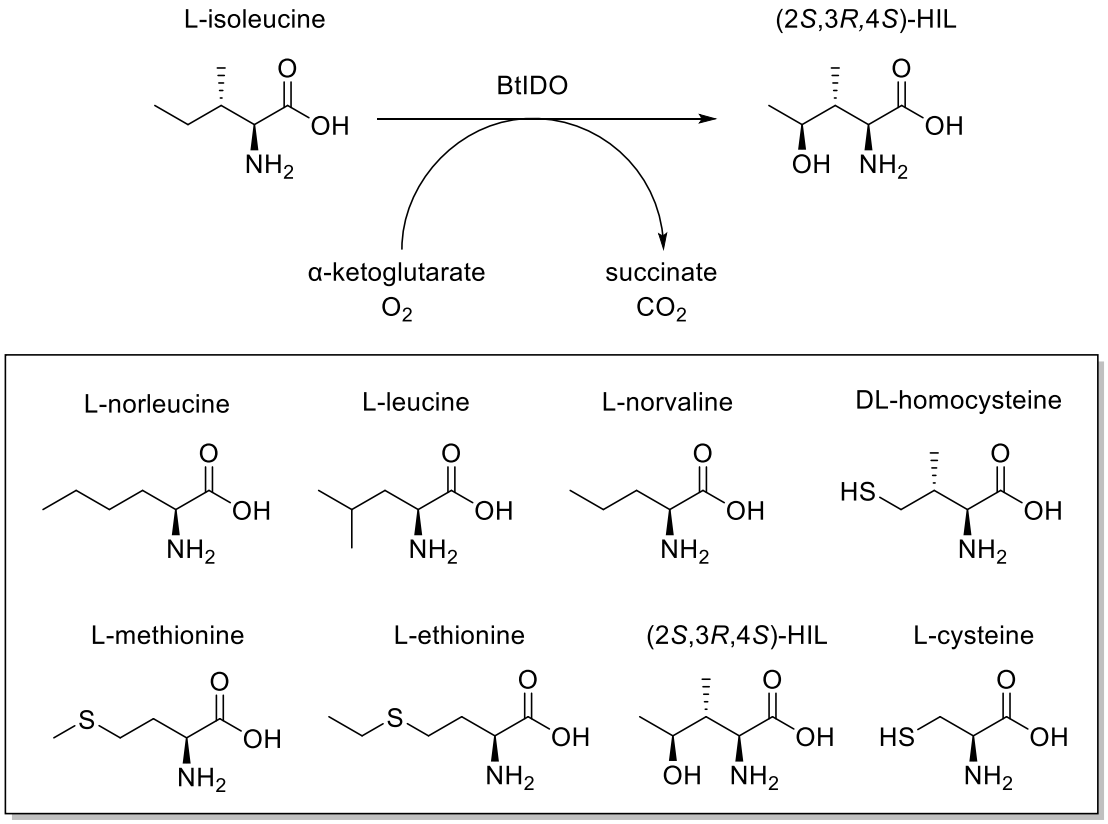

Figure 7: Oxyfunctionalisation of BtIDO towards L-isoleucine. Amino acids shown in the box are further accepted by BtIDO. Figure taken and adapted from (Hibi *et al*., 2011).

Detailed studies on the substrate specificity and identification of the corresponding products demonstrated that the enzyme catalyses hydroxylation, sulfoxidation and dehydrogenation depending on the substrate used for conversion. The enzyme also tolerates to some extent a variation in the L-isoleucine side chain (Hibi *et al*., 2011).

The primary reaction catalysed by IDO is the stereoselective hydroxylation at C-3, C-4 or C-5 position in several aliphatic L-amino acids bearing a main chain of at

least five carbon atoms. Methylene and methine carbons of the substrate are preferred over methyl carbons. Several L-amino acids are hydroxylated at γ-position including L-norleucine, L-leucine, L-isoleucine, L-allo-isoleucine and L-norvaline. Besides, δ-hydroxylation towards L-norleucine was reported (Enoki *et al*., 2016; Hibi *et al*., 2011). Thus, BtIDO exhibits a higher substrate promiscuity compared to other dioxygenases *like* AvLDO (Chapter 1.3.1) (Hibi *et al*., 2011). As the use of substrate analogues alters position and stereochemistry of the hydroxylation reaction, a certain degree of flexibility of the L-isoleucine side chain within the binding pocket is proposed (Gao *et al*., 2018).

1.3.3 L-Leucine hydroxylase GriE from *Streptomyces sp.* DSM40835

The α-ketoglutarate-dependent leucine hydroxylase GriE from *Streptomyces sp*. DSM40835, involved in the synthesis of the anti-tuberculosis molecules griselimycins (Kling *et al*., 2015), can be used for the chemo-enzymatic production of the antibiotic Manzacidin C and the antiviral natural product Cavinafungin B. GriE converts L-leucine into (2*S*,4*R*)-5-hydroxyleucine, the key building block for Manzacidin C. Further oxidation yields (2*S*,4*R*)-4-methylproline serving as building block for depsipeptide antibiotics and the antifungal substance Cavinafungin B (Figure 8) (Peters and Buller, 2019; Zwick and Renata, 2018a).

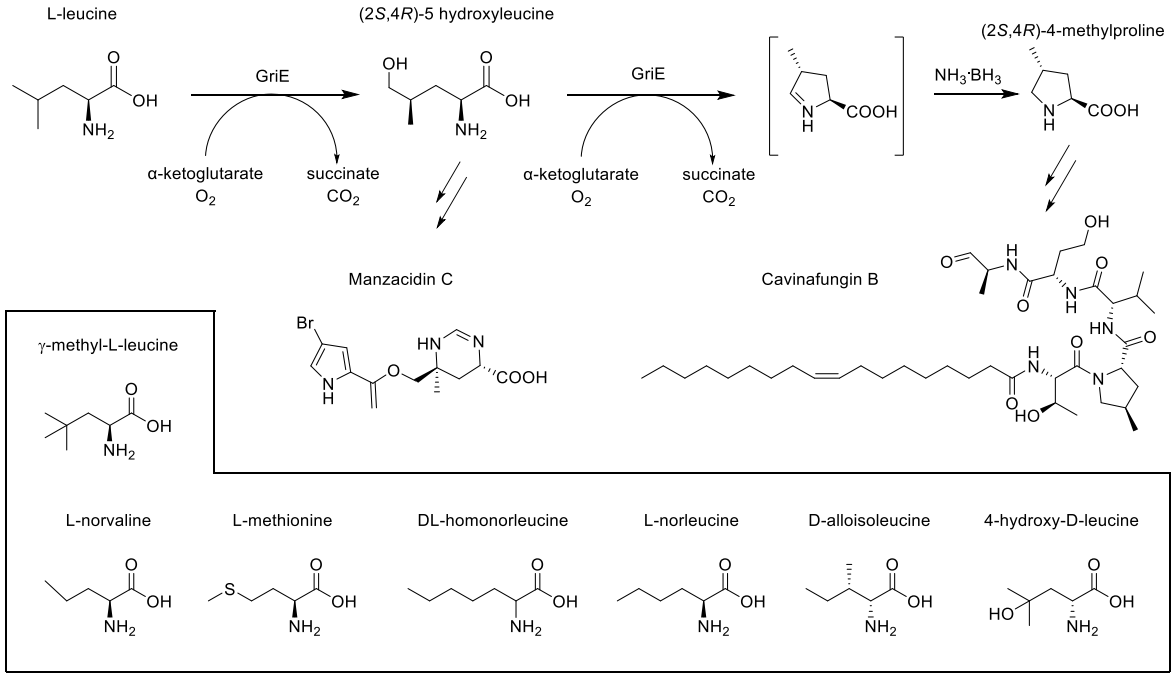

Figure 8: The 2-oxoglutarate-dependent oxygenase GriE was used to enzymatically prepare the key building blocks (2*S*, 4*R*)-5-hydroxyleucine and (2*S*, 4*R*)-4-methylproline utilised for the total chemo-enzymatic synthesis of Manzacidin C and Cavinafungin B. Amino acids shown in the box are further substrates accepted by GriE. Figure taken and adapted from (Peters and Buller, 2019; Zwick and Renata, 2018b).

Besides L-leucine, the work of Zwick and Renata (Zwick and Renata, 2018b) demonstrated, that a variety of amino acid derivatives including γ-methyl-L-leucine, L-norleucine, L-norvaline, D-alloisoleucine, L-methionine, 4-hydroxy-D-leucine and DL-homonorleucine are accepted as substrates to be hydroxylated at γ-position. However, for L-valine, D-leucine and L-isoleucine no enzyme activity was observed, which is also in accordance with previous studies (Lukat *et al*., 2017).

In contrast to substitution at the γ-position being well tolerated, GriE is sensitive to the substitutions at the β-position as the hydroxylation of L-allo-isoleucine was only hydroxylated with moderate turnover numbers. This observation suggests a steric effect of substituents located at the β-position on the enzyme (Zwick and Renata, 2018b). Except the sulfoxidation of L-methionine, GriE is selective for the γ-position and performs oxidations with excellent regio- and diastereoselectivities.

With its broad substrate specificity, the synthesis strategy focused on synergism of chemo- and biocatalytic C-H functionalisation for the preparation of various pyrrolidine building blocks (Zwick and Renata, 2018b).

1.4 Enzyme engineering strategies

Recent advances in the discovery of novel enzymes and biocatalytic engineering led to a notable increase in the global industrial enzyme market (Singh *et al*., 2016). Whereas in the past, the modification of process conditions was mainly applied, nowadays various state-of-the-art enzyme engineering strategies are utilised to overcome limitations in regio- and enantioselectivity, stability and specificity (Ibrahim and El-Dewany, 2018). Regarding the enzyme function enhancement, directed evolution, rational design, semi-rational design and computational *de novo* enzyme design are the four main state-of-the-art approaches (Denard *et al*., 2015).

1.4.1 Directed evolution

Directed evolution mimics the Darwinian evolution process by the *in vitro* generation of a large number of mutants of the gene encoding the enzyme to be improved. Screening and selection of desired variants is then performed with high-throughput screening methods. Several rounds of mutation and selection are applied to achieve an enhanced function of the biocatalyst (Arnold, 2015; Davids *et al*., 2013). The main strategies for creation of these mutant libraries are random mutagenesis, DNA shuffling and cassette mutagenesis. For random mutagenesis, the most common methods include error-prone PCR and site-saturation mutagenesis. Whereas error-

prone PCR uses a polymerase with low or no fidelity and conditions enhancing the mutation rate (*e.g.* unequal dNTP concentrations, Mn^{2+}), site-saturation mutagenesis aims at amino acid replacement for all possible amino acids within the enzymes. This random mutagenesis approach generates a large number of variants ($<10^6$) subjected to screening in order to detect desired ones (Adrio and Demain, 2014; Wu and Arnold, 2013). Directed evolution rates amongst the most powerful approaches in enzyme function improvement, not requiring any prior structural knowledge of the biocatalyst. One of the main drawbacks of this strategy is the requirement of high-throughput screening methods to effectively detect the improved enzyme variants (Jemli *et al*., 2016; Porter *et al*., 2016).

1.4.2 Rational design

In contrast, the rational design approach requires knowledge and understanding of the relationship between the enzyme structure and its function to achieve the changes desired. Therefore, an ample quantity of data including DNA/protein sequences, structural information with the residues responsible for catalytic activity, selectivity, specificity and stability is required to successfully apply this approach (Adrio and Demain, 2014; Denard *et al*., 2015).

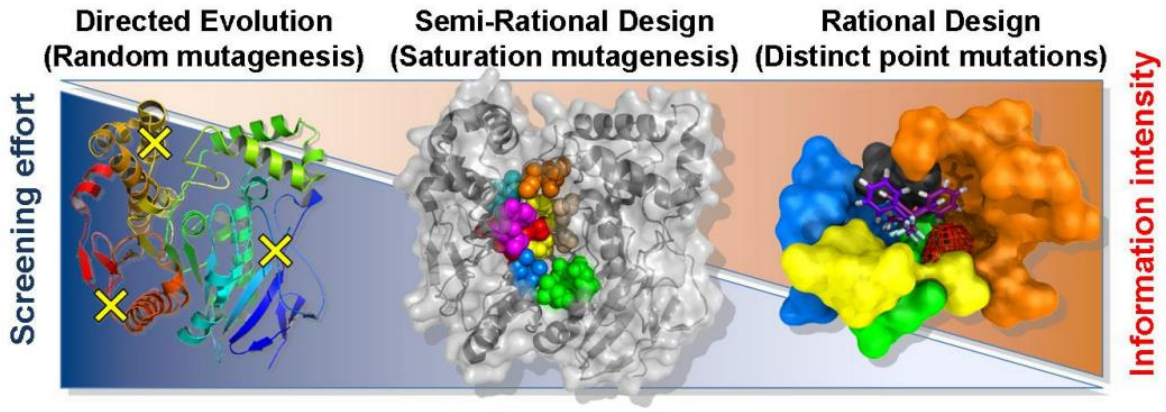

Figure 9: Approaches available for the generation of improved protein variants. The screening effort inversely correlates with the information intensity. Whereas directed evolution requires a tremendous screening effort, but low information intensity, for rational design it is the other way around. Semi-rational design requires moderate extent of information and an intermediate screening effort. Figure taken from (Schließmann, 2010).

1.4.3 Semi-rational design

The combination of rational design and directed evolution is called semi-rational design or focused directed evolution (Schließmann, 2010). It only requires preliminary information about structure and function of the enzyme in order to select a restricted number of amino acid residues to be targeted for mutation (*e.g.* residues

comprising the active site or involved in ligand binding). The aim is to create smaller and smarter libraries compared to random mutagenesis, which are highly enriched with mutations to be most likely effective for enzyme improvement (Zhuang *et al*., 2017). It is the method of choice if no high-throughput assay for screening is available. For semi-rational design, the protocols of combinatorial active-site saturation testing (CASTing) and iterative saturation mutagenesis (ISM) comprise the most effective approaches. The main advantage of CASTing is that saturation mutagenesis is applied at multiple sites within the active site simultaneously. It enables the investigation of synergistic effects, which multiple mutations possibly have on the function of the enzyme. However, careful assessment, if the available screening system is suitable for the number of created variants, should be performed (Parra *et al*., 2013; Shen *et al*., 2015). A considerably lower number of variants is generated by iterative saturation mutagenesis (ISM) evolved by Reetz *et al.* (Reetz *et al*., 2010b; Reetz *et al*., 2010a). Therefore, 2 – 4 groups of 1 – 3 amino acid residues important according to structural data are formed and individually subjected to saturation mutagenesis. The four different libraries created are in turn screened. Desired variants of each position are used for another round of saturation mutagenesis at another of the four positions. This process can be repeated until the desired modification is achieved (Reetz *et al*., 2010b; Reetz *et al*., 2010a). By using the semi-rational design strategy, many properties including catalytic activity, regio- and enantio-selectivity and substrate specificity of several enzymes were effectively improved (Li and Cirino, 2014; Wang *et al*., 2016).

1.4.4 Current strategies used for the engineering of Fe(II)/α-ketoglutarate dependent dioxygenases

Sun *et al.* (Sun *et al*., 2019) applied both random and site-directed mutagenesis for the engineering of leucine dioxygenase NpLDO from *Nostoc punctiforme* (also termed LdoA) for the production of 5-hydroxyl leucine, a promising intermediate for the synthesis anti-tuberculosis griselimycins. For the directed evolution approach, error-prone PCR was used for library generation and a succinate screening assay in microplates was carried out for screening desired mutants. By applying this strategy, a mutant with a 9.9-fold increase in the specific constant (k_{cat}/K_m) towards leucine could be detected. For the (semi)-rational approach, a three-dimensional (3D) homology model using GriE (Chapter 1.3.3) as a template was created.

Structural modelling and molecular dynamics simulation were used to identify residues to be targeted by site-directed mutagenesis. By combining mutations generated with the directed evolution and rational design approach, a variant with a 29.7-fold increase in the specific constant was created (Sun *et al*., 2019).

In contrast, Qin *et al.* (Qin *et al*., 2014) performed a structural optimization of Fe(II)- and α-ketoglutarate-dependent dioxygenase SadA from *Burkholderia ambifaria* AMMD to enhance the activity towards *N*-succinyl-L-*threo*-3,4-dimethoxyphenylserine (NSDOPS) to produce *N*-succinyl-3,4-dimethoxy-phenylalanine (NSDOPA), a precursor of the psychoactive drug L-*threo*-3,4-dihydroxyphenylserine (L-DOPS, Droxidopa®). As the crystal structure of the SadA-Zn(II)-α-KG complex (Qin *et al.*, 2013) was already available, a model of SadA with Zn(II), α-KG and NSDOPA was created to identify residues around the active site related with substrate recognition. Three rounds of site-directed saturation mutagenesis at two defined positions led to the generation of combinatorial mutants showing more than 6-fold increase in activity towards the desired substrate (Qin *et al*., 2014).

When no crystal structure of the enzyme, but also no high-throughput screening assays are available, a semi-rational approach can be applied (Ibrahim and El-Dewany, 2018). Information on the reaction mechanism and some knowledge of the protein family using the 3DM (3D molecular class specific information systems) database could assist in identifying hotspots within the genes being likely to influence the desired properties *e.g.* substrate promiscuity, stability, activity. The multiple sequence alignment created by 3DM is based on multiple structure alignments. Therefore, a large number of superfamily members with a known 3D-structure is crucial for the systemic analysis to succeed well in the discovery of protein functionalities (Kuipers *et al*., 2010). By targeting mutations close to the active site, possibly involved in substrate binding, the likelihood of creating mutations with a beneficial effect should be enhanced as according to Morley and Kazlauskas, closer mutations are more effective (Morley and Kazlauskas, 2005). For instance, Koketsu *et al.* and his coworkers used a homology model of the proline hydroxylase SmPH from *Sinorhizobium meliloti* guiding mutagenesis of residues in close proximity to the active site. SmPH catalyses the hydroxylation of L-pipecolic acid (L-Pip) generating the regioisomers *cis*-5-hydroxy-L-proline (*cis*-5-Hypip) and

cis-3-hydroxy-L-proline (*cis*-3-Hypip) in a 1:1 ratio. In this work, different strategies of enzyme engineering were combined to improve the enzyme activity and enhance selectivity towards the desired *cis*-5-Hypip. Three rounds of protein engineering were performed including amino acid substitutions at defined positions based on the knowledge obtained from the homology model, site saturation mutagenesis of various positions using degenerated primers and a random mutagenesis library created by error prone PCR. Libraries were screened by high performance liquid chromatography (HPLC) and thin layer chromatography (TLC). The adoption of a stepwise strategy first enhancing selectivity and then activity led to an increase in the regioselective hydroxylation of L-pip from 60 to 95 % (Koketsu *et al*., 2015).

1.4.5 High throughput screening and selection methods

Successful enzyme engineering is accompanied by the availability of an effective high throughput screening or selection (HTSOS) assay, not only to increase the probability of obtaining variants with desired properties but also to reduce time and costs (Xiao *et al*., 2015). In a directed evolution experiment, this stage is the most important aspect, but also the most challenging (Porter *et al*., 2016). Although advances in molecular biology enable the creation of various mutant libraries by the introduction of genetic diversity, the main limitation is the detection of the desired variants (Ruff *et al*., 2013; Yuan *et al*., 2005). While selection automatically eliminates undesired non-functional mutants by directly linking the acquired or improved enzyme function to cell growth, the objective of screening is to evaluate every protein variant for the desired characteristic (Leemhuis *et al*., 2009; Xiao *et al*., 2015).

As selection only carries positive mutants onto the next cycle of directed evolution, it enables the screening of much larger libraries of more than 10^{11} variants. Generally, high throughput selection methods can be divided into display and compartmentalisation. The former is based on the physical connection of the translated protein to its encoding gene or restricting the gene into a viral particle displaying the protein as plasmid or phage display. The latter constrains gene and protein spatially in a single compartment as *in vivo* compartments *e.g.* phage particles or bacterial/yeast cells or *in vitro* compartments *e.g.* water-in-oil emulsions (Xiao *et al*., 2015). Growth complementation is another high throughput selection strategy exclusively carried out in living cells, coupling enzyme property and fitness

of the host cell to an extent, that only cells containing the desired property are able to survive (Figure 10). As it is the case for most selection assays, the library size is only limited by the transformation efficiency of the cell. The main advantage is that the selection method has the potential to select enzymes with diverse properties. However, one major disadvantage is, that the selection strategy needs to be individualised and is limited to the selection of properties that can be coupled to the host cell metabolism (Xiao *et al*., 2015).

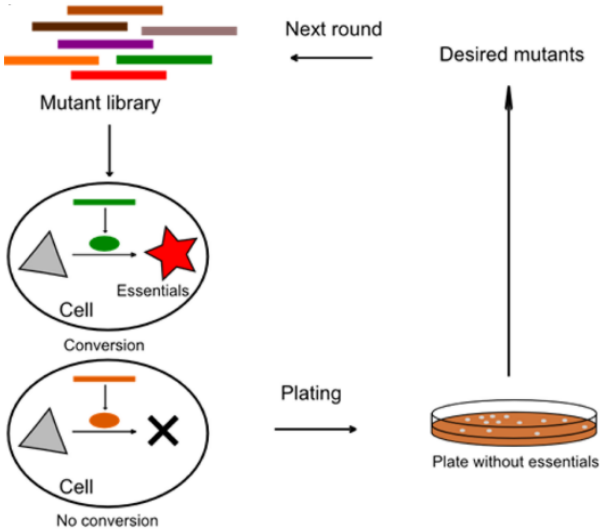

Figure 10: Growth complementation. Host cells were transformed with a mutant library. A precursor molecule (grey triangle) is converted into a vital compound (red star) by the active mutant protein (green oval). In contrast, the nonactive mutant protein (brown oval) does not convert the precursor and consequently the molecule essential for cell survival is not formed. After plating the cells onto plates lacking the compound essential for cell survival, only cells forming the mutant proteins will survive and be selected. Then, the active genes are recovered and targeted for the next round of mutagenesis. Figure taken from (Xiao *et al*., 2015).

1.5 Artificial TCA cycle for the selection of α-ketoglutarate dependent hydroxylase catalysts

The growing interest of hydroxylated amino acids as valuable building blocks for food and pharmaceutical industry promotes the engineering of metabolic pathways to improve the biosynthesis of these valuable products (Peters and Buller, 2019; Theodosiou *et al*., 2017). Hence, the construction of optimised cell factories to produce hydroxylated amino acids is a highly attractive field of investigation.

Since hydroxylases typically depend on the cosubstrate α-KG, recombinant *E. coli* harbouring a hydroxylase gene branch off α-KG from the TCA cycle. Whereas the enzyme is fuelled with α-KG, the coproduct succinate is in turn assimilated and recycled (Figure 11) (Hausinger, 2004; Theodosiou *et al*., 2017).

Generally, four pathways in *E. coli* lead to succinate synthesis: (1) under aerobic conditions from α-KG catalysed by the α-ketoglutarate dehydrogenase via the TCA cycle, (2) from isocitrate catalysed by isocitrate lyase via the glyoxylate pathway, (3) under anaerobic conditions from oxaloacetate via the reductive branch of the TCA-cycle and (4) via the γ-aminobutyric acid (GABA) pathway from L-glutamate and L-arginine (Samsonova *et al*., 2005).

In recent years, Smirnov *et al.* (Smirnov *et al*., 2010) and Theodosiou *et al.* (Theodosiou *et al*., 2017) engineered the TCA cycle of *E. coli* to optimize the production of (2*S*,3*R*,4*S*)-HIL and *trans*-4-hydroxy-L-proline. The enhanced α-KG levels in engineered *E. coli* led to an increased flux through the hydroxylating activity of the corresponding enzymes IDO and P4H (Peters and Buller, 2019; Smirnov *et al*., 2010; Theodosiou *et al*., 2017).

The engineering of *E. coli* MG1655 for instance yielded *trans*-4-hydroxy-L-isoleucine from L-isoleucine. The production of this insulinotropic, nonproteinogenic amino acid could be achieved using α-ketoglutarate-dependent L-isoleucine-4-hydroxylase (IDO) and the knockout of the genes *sucAB* (encoding α-KG dehydrogenase) as well as *aceA* (encoding isocitrate lyase). The aim of these modifications was to fuel carbon through IDO, restoring the conversion of α-ketoglutarate to succinate, which prevents the breakdown of the TCA cycle. Further, the gene *aceK* (isocitrate kinase/phosphatase) was deleted to prevent the deactivation of the isocitrate dehydrogenase and therefore ensure the α-ketoglutarate synthesis. Further, the Na^+ dependent branched-chain amino acid transporter (BrnQ) was overexpressed to avoid substrate uptake restrictions (Smirnov *et al*., 2010).

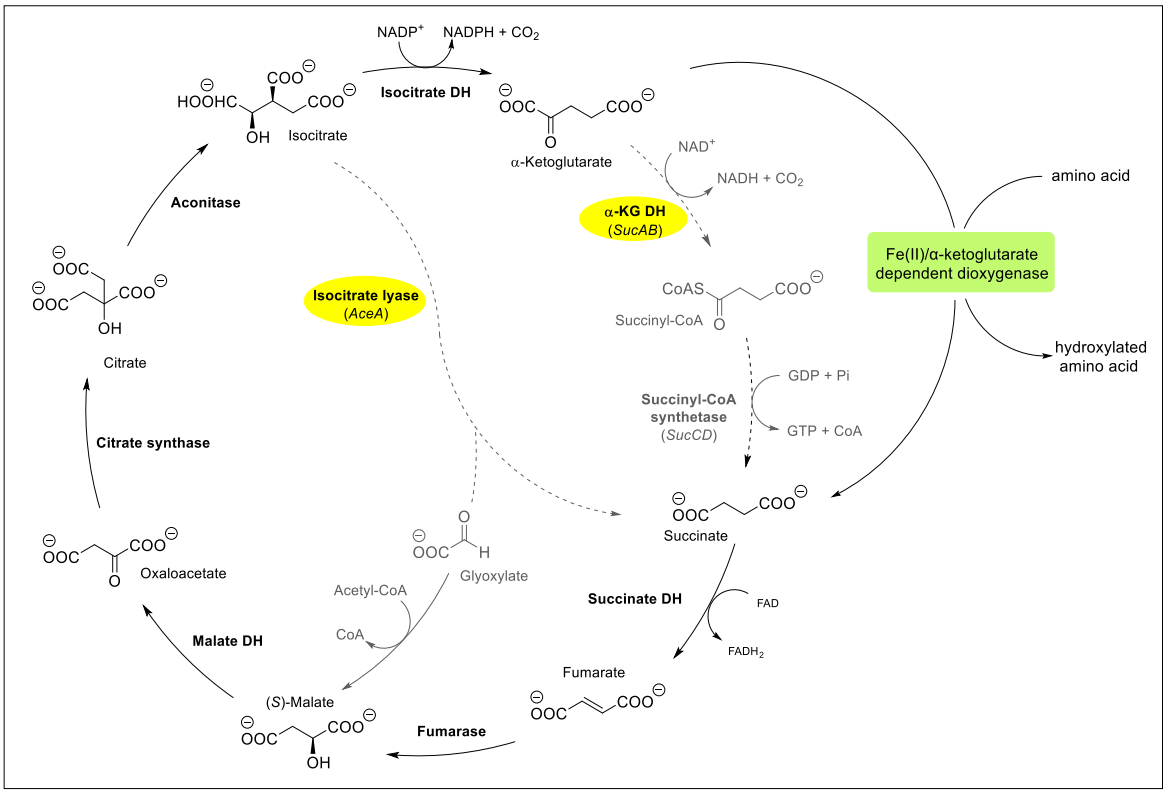

Figure 11: Restorage of the TCA cycle of engineered *E. coli* BL21(DE3) 3ΔsucA cells via dioxygenase catalysed L-amino acid hydroxylation. Figure taken and adapted from (Theodosiou *et al*., 2017).

Theodosiou *et al.* (Theodosiou *et al*., 2017) further engineered the TCA cycle in *E. coli* BL21(DE3) (pLysS) to produce *trans*-4-hydroxy-L-proline (hyp). It is a valuable building block for the production of pharmaceuticals (Hara and Kino, 2009) and can be synthesized by using an α-ketoglutarate dependent L-proline-4-hydroxylase (P4H). By separately deleting *sucA* (encoding α-ketoglutarate dehydrogenase E1 subunit) and *sucC* (encoding sucCoA synthetase β subunit) in combination with the isocitrate lyase enconding gene *aceA* (glyoxylate shunt pathway inactivation) the effect of each subunit on host physiology could be assessed. By the deletion of *putA* (encoding proline dehydrogenase), a yield of 100% for *trans*-4-hydroxy-L-proline was achieved by the avoidance of proline degradation (Theodosiou *et al*., 2017). In contrast to the knockout of *sucC*, the deletion of *sucA* combined with *aceA*, *putA* and the additional overexpression of the Na^+/L-proline transporter PutP led to a significant increase of the proline hydroxylation rate. Further, a decrease of the average cell size was observed. It could be explained by the fact that the intermediate sucCoA, not formed in this strain, is needed for diaminopimelate biosynthesis being important for cell wall formation (Theodosiou *et al*., 2017). However, it has been observed that *E. coli* cells in minimal medium excrete small

amounts glutamate, which can in principal be used to produce succinate via the GABA pathway (Li *et al*., 2006). As only the *sucAB*, but not the *sucCD* genes are knocked out in the 3Δ*sucA* strain (Figure 11), succinate can either be used to restore the TCA or for the production of sucCoA to assist cell wall formation. Since the impact of the *sucA* deletion on cell morphology is considerably high (Theodosiou *et al*., 2017), the contribution of this side pathway for sucCoA formation and cell wall synthesis is considered to be rather small.

E. coli Δ*sucA*Δ*aceA* not only shows great potential as a selection tool for Fe(II)/α-ketoglutarate dependent dioxygenases, but can also be used for the screening of mutant enzyme libraries generated through random mutagenesis (Smirnov *et al*., 2010; Theodosiou *et al*., 2017; Zhang *et al*., 2017).

For instance, Zhang *et al.*, 2017 (Zhang *et al*., 2017) used the double-knockout strain *E. coli* K-12 MG1655 Δ*sucA*Δ*aceA* to screen a mutant library of isoleucine dioxygenase variants generated by error prone PCR. The selection strain was transformed with the library and grown on M9 minimal medium supplemented with L-isoleucine and α-ketoglutarate. Five clones showing larger colonies on the plate were selected and four variants showed higher 4-HIL production indicating an increased enzyme activity towards L-isoleucine. One variant showed two mutated amino acids located in the prominent His^{159}-X-Asp^{161}-X_{50}-His^{212} motif. The location of the mutation is supposed to be responsible for the increased catalytic activity and for the improved thermal properties also detected by further investigation (Zhang *et al*., 2017). To overcome the limitation of substrate uptake by the cell membrane and avoid increased substrate consumption for cell growth, resting cells were used by Zhang *et al.* By the use of cells, frozen at -80°C prior to biotransformation, 151.9 mmol of 4-HIL/L (22.4 g/L) could be obtained in 12 hours and higher yields of 4-HIL for L-isoleucine and α-ketoglutarate were achieved compared to the biotransformation method used previously (Zhang *et al*., 2017).

In this study, the powerful selection strain Δ*sucA*Δ*aceA*Δ*putA* (also termed 3Δ*sucA*) (Theodosiou *et al*., 2017) was used to implement a selection assay for the discovery improved dioxygenase variants generated by site-directed saturation mutagenesis.

1.6 Applications of α-methyl amino acids

The stereoselective synthesis of quaternary stereocenters belongs to the most challenging aspirations in organic chemistry (Quasdorf and Overman, 2014). As several natural products, including lactacystin and sphingofungine E and F as well as bioactive molecules for pharmaceutical applications, comprise α-quaternary amino acids, these moieties are attractive targets for evolving reliable synthetic strategies (Kang *et al*., 2005). Lately, α,α-disubstituted amino acids were investigated to serve as building blocks for peptidomimetics. These molecules are used to design new short chain peptides to treat or prevent diseases. In comparison to α-hydrogen analogues, α,α-disubstituted amino acids in peptides tend to show a more defined conformation. The presence of an additional fourth stereocenter leads to the generation of specific conformational changes (Brun *et al*., 2008; Crisma *et al*., 2014).

Among all α,α-disubstituted amino acids, α-methyl-substituted amino acids are the most prominent candidates in the field of peptide research. Especially α-methylphenylalanine ((αMe)Phe) is a widely used peptide residue in sweeteners and chemoattractants (Parmee *et al*., 1991; De Zotti *et al*., 2015). As the generation of α-quaternary amino acids is a laborious task in organic synthesis, biocatalytic methods have been developed, but are still scarce. The most promising enzymes used for the generation of these valuable products show a strict substrate specificity and consequently a limited product range (Blesl *et al*., 2018).

The extension of the substrate scope of α-ketoglutarate-dependent dioxygenases towards α-methyl-substituted amino acids would enable the generation of a greater diversity of these appealing chemical building blocks and is therefore targeted in this thesis.

1.7 Aim

The aim of this master thesis was to implement a growth-based selection assay for the detection of improved dioxygenase variants generated by simultaneous site-directed saturation mutagenesis. A library of dioxygenase variants modified at residues located in close proximity to the substrate binding site was created and an engineered *E. coli* strain was employed for the selection of improved mutants. The goal was to find variants with an extended substrate scope, especially towards α-methyl amino acids for a broader industrial application.

Cultivation in minimal medium was performed to select mutants showing higher activity towards the substrate supplemented in the medium. After each cultivation cycle, part of the liquid enrichment culture was plated.

As conversion of the substrate is essential for cell survival, higher conversion resulted in faster growth and consequently clones formed on plates with larger size were picked for further analysis.

Selected variants were sequenced to figure out the mutations responsible for enrichment. Re-screening in test tubes was performed to compare the growth behaviour of the wild type and selected clones to detect false positive variants. Then improved enzyme variants were expressed under assay conditions and compared to the wild type enzyme.

Finally, an in-vitro activity assay was performed with the selected AvLDO variants to confirm the conversion of L-isoleucine via a TLC-based screening.

2 Materials

2.1 Chemicals and equipment

All chemicals and equipment for the conduction of this study are listed in Appendix 8.1.

2.2 Bacterial strains and plasmids

All strains and plasmids used in this study are shown in the table below. Vector maps of the different plasmids are shown in Appendix 8.2.

Table 1: List of strains and plasmids used in this study.

Strain or plasmid	Relevant characteristics	Source or reference
E. coli TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Thermo Fisher Scientific
E. coli BL21(DE3)	F– *omp*T *hsdS*$_B$ (r_B–, m_B–) *gal dcm* (DE3)	Thermo Fisher Scientific
E. coli BL21(DE3) 3Δ*sucA*ΔaceAΔputA (DE3) (pLyS) (3Δ*sucA*)	Deletion of *sucA*, *aceA* and *putA* encoding α-KG dehydrogenase subunit E1, isocitrate lyase and proline dehydrogenase; Cm^r	Helmholtz centre for environmental research – UFZ, Leipzig
E. coli BL21(DE3) 3Δ*sucA*[pET28a_AvLDO]	*E. coli* BL21(DE3) 3Δ*sucA* harbouring pET28a_AvLDO for growth-based selection assays; Kan^r	Hanreich, 2019
E. coli BL21(DE3) 3Δ*sucA*[pET28a_BtIDO]	*E. coli* BL21(DE3) *3ΔsucA* harbouring pET28a_BtIDO for growth-based selection assays; Kan^r	This study
E. coli BL21(DE3) 3Δ*sucA*[pET28a_GriE]	*E. coli* BL21(DE3) 3Δ*sucA* harbouring pET28a_GriE for growth-based selection assays; Kan^r	Hanreich, 2019
E. coli BL21(DE3) 3Δ*sucA*[pET28a(+)]	*E. coli* BL21(DE3) 3Δ*sucA* harbouring the empty vector pET28a(+); Kan^r	Hanreich, 2019
E. coli BL21(DE3) 3Δ*sucA*[pET22b(+)]	*E. coli* BL21(DE3) 3Δ*sucA* harbouring the empty vector pET22b(+); Amp^r	Hanreich, 2019
pET28a(+)	Vector containing a cloning/expression region transcribed by T7 RNA polymerase; N/C-term His-tag; Kan^r	in-house collection
pET28a_AvLDO	pET28a containing the L-leucine dioxygenase gene *AvLDO* (6091 bp) from *A. variabilis*; N-term His-tag; Kan^r	Ruhr Universität Bochum
pET28a_BtIDO	pET28a containing the L-isoleucine dioxygenase gene *BtIDO* (6007 bp) from *B. thuringiensis*; N-term His-tag; Kan^r	This study

pET22b_BtIDO	pET28a containing the L-isoleucine dioxygenase gene *BtIDO* (6007 bp) from *B. thuringiensis*; N-term His-tag; Amp^r	Ruhr Universität Bochum
pET28a_GriE	pET28a containing the L-leucine dioxygenase gene *GriE* (6121 bp) from *Streptomyces* strain DSM 40835; N-term His-tag; Kan^r	GenScript Biotech B.V. (Netherlands)

2.3 Enzymes

All enzymes relevant for this study are shown in the table below.

Table 2: List of enzymes used in this study.

Enzyme	Short name	Literature	Source
Leucine dioxygenase from *A. variabilis*	AvLDO	Correia Cordeiro *et al.*, 2018	Ruhr Universität Bochum
Isoleucine dioxygenase from *B. thuringiensis*	BtIDO	Kodera 2009, Hibi 2011, Enoki 2016	Ruhr Universität Bochum
Leucine dioxygenase from *Streptomyces* strain DSM 40835	GriE	Lukat and Katsuyama 2017	GenScript Biotech B.V. (Netherlands)

2.4 Primers

All general primers are listed in the table below. Mutagenesis primer used in this study are listed separately in the respective section.

Table 3: List of primers used in this study.

Primer	Sequence (5'-3')	Length [nt]	Target gene or region
mPCR_AvLDO_rev	GAAAGTTATACGTTGTCCATCCATC	25	*AvLDO*
mPCR_AvLDO_fwd	CTTCTCAACAGCTTAAGTCCAAAG	24	*AvLDO*
mPCR_BtIDO_rev	GTACTACATTCTGTTCCAAGTGGC	24	*BtIDO*
mPCR_BtIDO_fwd	CTATGATGATGGCGGGAAAGTTAG	24	*BtIDO*
mPCR_GriE_rev	CATTGGATGCTTTTCCACCATCTG	24	*GriE*
mPCR_GrieE_fwd	GAAAAGTACAAGAGTGACGGTTAC	24	*GriE*
T7 pET-mod	CTAGTTATTGCTCAGCGGT	19	T7 promoter
T7 term	CCCGCGAAATTAATACGACTCAC	23	T7 terminator
AvLDO_upstr_seq	TAGCAGTCATATGGCTGCCG	20	upstream region of *AvLDO*
AvLDO_downstr_seq	TTCGGCGATGGGATTACTTC	20	downstream region of *AvLDO*
AvLDO_upstr_1	CATTAAGTTCTGTCTCGGCG	20	inside *LacI* upstream

AvLDO_downstr_1	TTCAACAGGCCAGCCATTACG	21	inside *Kan*^{r} downstream

2.5 Cultivation media, buffers, and solutions

2.5.1 LB-medium

Lysogeny broth (LB) medium was prepared by dissolving 20 g of a ready-to-use mixture [5 g Yeast extract, 10 g NaCl, 10 g Tryptone] in 1 L double distilled water (ddH_2O) to a final concentration of 20 g/L. The medium was autoclaved for sterilisation.

2.5.2 LB-SOC medium

10x super optimal broth (SOC) was prepared by dissolving 0.1 g KCl, 1 g $MgCl_2$, 1 g $MgSO_4$ and 2 g Glucose monohydrate in 50 mL and autoclaved for sterilisation. In order to create a 1x LB-SOC mixture the 10x SOC is diluted with LB medium.

2.5.3 TB medium

Terrific broth (TB) medium was prepared by dissolving 12 g tryptone, 24 g yeast extract and 5 g glycerol in 900 mL ddH_2O. Further, 100 mL KH_2PO_4/K_2HPO_4 buffer (10x) was prepared by dissolving 2.31 g KH_2PO_4 and 12.54 g K_2HPO_4 in 100 mL ddH_2O. Both solutions were autoclaved separately and mixed afterwards to obtain the final 2x TB medium used for expression.

2.5.4 2x TY medium

2x TY medium was prepared by dissolving 15 g tryptone, 10 g yeast extract and 5 g NaCl in 1 L ddH_2O. The medium was autoclaved for sterilisation.

2.5.5 M9 medium

All stock solutions listed in Table 4 were prepared and sterilised separately. Thiamine hydrochloride, Biotin and Us^{Fe} trace element solution were stored in small aliquots at -20 °C. The medium was then prepared freshly if required.

Table 4: Composition of M9 minimal medium used for the cultivation of *E. coli* 3Δ*sucA*.

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	1 M	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
Thiamine x HCl	1 mg/L	20,000x	1 mg/mL	Sterilise over 0.22 µm filter and store aliquots at -20 °C.	1 mL
Biotin	5 g/L	1,000x	0.1 mg/mL	Sterilise over 0.22 µm filter. Store aliquots at -20 °C.	50 µL
Us^{Fe} trace element solution	1 mL/L	1,000x		Sterilise over 0.22 µm filter. Store aliquots at -20 °C.	1 mL
Glucose	5 g/L	20 %	200 g/L	Sterilised by autoclaving and store at 4 °C.	25 mL
Substrate	5 mM	10x	50 mM	Sterilise over 0.22 µm filter.	100 mL
ddH_2O				Sterilise by autoclaving	670 mL (w/o substrate: 770 mL)

Table 5: Composition of the Us^{Fe} trace element solution (1,000x) used for preparation of M9 minimal medium.

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 1 M HCl (82.8 mL 37 % HCl/L)	

2.5.6 LB-agar antibiotic plates

LB-agar plates were prepared by dissolving 16 g of a ready-to-use mixture LB [4 g Yeast extract, 8 g NaCl, 8 g Tryptone] and 12 g agar-agar in 800 mL ddH_2O. Before pouring the plates, 800 µL Ampicillin^{100} [Stock solution: 100 mg Ampicillin sodium salt/mL ddH_2O] or Kanamycin^{40} [Stock solution: 40 mg Kanamycin sulfate/mL ddH_2O] stock solution was added, respectively.

2.5.7 M9-agar antibiotic plates

In order to prepare 500 mL M9-agar 7.5 g agar-agar was dissolved in 285 mL ddH_2O and autoclaved. After cooling to - 80 °C, all solutions listed in Table 4 (except water and substrate) were added under sterile conditions. Respective stock volumes listed for 1 L medium were adapted for 0.5 L, respectively.

Instead of 5 mM L-leucine or L-isoleucine, 1 g/L L-leucine or L-isoleucine was used [Stock solution: 0.5 g/50 mL]. Further, the medium was supplemented with 1 g/L α-ketoglutaric acid [Stock solution: 0.5 g/ 50 mL]. 500 µL Kanamycin^{40} and 25 µL Isopropyl-β-D-thiogalactopyranosid (IPTG) [Stock solution (1 M): 2.383 g IPTG in 10 mL ddH_2O] were added additionally.

2.5.8 SDS sample buffer (4x)

For the preparation of 9.5 mL sample buffer (4x) the following components were mixed gently: 1.25 mL TRIS-HCl buffer [0.5 M, pH 6.8], 2 mL 10 % (w/v) SDS solution, 2 mL 0.5 % (w/v) bromophenol blue, 2.5 mL glycerol and 3.5 mL ddH_2O. Aliquots of 950 µL were stored at –20 °C and before use 50 µL β-mercaptoethanol was added per aliquot.

2.5.9 Staining solution for SDS-gel

1 g Coomassie Brilliant Blue R-250, 100 mL Glacial acetic acid (conc.), 300 mL Ethanol and 600 mL dH_2O were used for the preparation of 1 L staining solution.

2.5.10 Destaining solution for SDS-gel

100 mL Glacial acetic acid (conc.), 300 mL ethanol and 600 mL ddH_2O were used for the preparation of 1 L destaining solution

3 Methods

3.1 Overnight cultures

For plasmid preparation and storage of glycerol stocks, fresh colonies or 10 µL glycerol stock were used to inoculate 5 mL LB-antibiotic (Amp^{100} or Kan^{40}) and for expression studies under optimised conditions 10 mL LB-Kan^{40} were used. Overnight cultures (ONCs) for expression studies under assay conditions, growth studies and for the growth-based selection assay were prepared by inoculating 50 mL LB-Kan^{40} (300 mL shake flasks) with fresh colonies. All overnight cultures were cultivated at 37 °C and 120 rpm for 16 – 18 hours.

3.2 Storage of strains

For the long-term storage of strains, glycerol stocks were prepared by mixing 1 mL ONC (Chapter 3.1) and 1 mL sterile glycerol (60 %) in cryotubes. Stocks were stored at –20°C or –80 °C until further use.

3.3 Preparation of electrocompetent cells

The protocol was applied for the preparation of electrocompetent *E. coli* TOP10 cells, BL21(DE3) cells and for BL21(DE3) 3Δ*sucA* cells. All equipment required for the preparation of competent cells (centrifuge bottles, reaction tubes) was pre-cooled at -20 °C. Water and glycerol needed for resuspension were pre-cooled at 4 °C. ONCs of the desired strain were prepared using 50 mL 2x TY medium. After incubation at 37 °C at 200 rpm for 14 – 16 hours, 5 mL of the ONC were used to inoculate the main culture (2 L 2x TY medium). Cells were harvested at an OD_{600} of 0.5 – 0.6 after 2 – 5 hours.

Flasks were placed on ice for one hour before the culture was transferred to pre-cooled centrifuge bottles. All centrifugation steps were performed with the Avanti® J-20 XP using the JA-10 rotor (Beckmann coulter, USA). The first centrifugation step was performed at 3,000 g for 10 min at 4 °C. After removing the supernatant, the cell suspension was resuspended in ddH_2O and afterwards filled up to ¼ of the bottle's total volume with ddH_2O. After the second centrifugation step (4,000 g, 10 min, 4 °C) the supernatant was removed and the pellet was again resuspended in ddH_2O, but filled up with 10 % glycerol. The third centrifugation step was performed at 4,500 g for 20 min at 4 °C and then the supernatant was removed and the whole bottle was filled up with 10 % glycerol.

After centrifugation at 5,000 g at 4 °C and removal of the supernatant, 2 mL 10 % glycerol were added per litre culture medium. Aliquots of 80 µl were prepared and pre-cooled reaction tubes were immediately placed on ice. Aliquoted electrocompetent cells were stored at - 80 °C.

3.4 Electroporation

Electrocompetent *E. coli* cells were thawed on ice before 1 µL of purified plasmid DNA was added. The reaction was then transferred into pre-cooled electroporation cuvettes with 0.2 cm gap width (Cell Projects, UK). Electroporation was performed with a resistivity of 200 Ω, an electrical capacity of 25 µF and an electric voltage of 2.5 kV for approximately 4 – 6 ms using the GenePulser II electroporation system (Bio-Rad, USA). Then, 700 µL pre-warmed SOC medium was added and the liquid was transferred into 1.5 mL reaction tubes and incubated at 37 °C at 350 rpm. *E. coli* BL21(DE3) cells and *E. coli* TOP10 cells were regenerated for an hour (Amp^r) and for 1.5 hours (Kan^r), whereas the knockout strain *E. coli* BL21(DE3) 3Δ*sucA* was regenerated for 2 hours due to its impaired metabolism. After regeneration, 100 µL of the supernatant were plated on LB-Amp^{100} or LB-Kan^{40} plates (Chapter 2.5.6), depending on the plasmid type used. Then, the tube was centrifuged at 13,000 rpm for 0.5 min, the supernatant was removed, and the pellet was resuspended in the remaining medium and plated as well.

3.5 Determination of the transformation efficiency of electrocompetent *E. coli* cells

For the determination of the transformation efficiency, 80 µL electrocompetent *E. coli* (TOP10, BL21(DE3), BL21(DE3) 3Δ*sucA*) cells were transformed with 10 ng of the plasmid pET28a(+). Electroporation was performed as described in the section before (3.3) except the plating of the regenerated cell suspension.

After regeneration, the cells were diluted 1:10 and 1:100 with LB-SOC medium and 100 µL were plated on LB-Kan^{40} plates

After cultivating the plates overnight, colony forming units per µg plasmid DNA (CFU/µg) were determined by using Formula (1) for the calculation.

$$\frac{CFU}{\mu g} = colonies * dilution\ factor * 10\ ng * \frac{100}{\mu g} \qquad (1)$$

3.6 Cloning of *BtIDO* in pET28a(+)

In general, all molecular cloning procedures were performed according to standard techniques (Ausubel *et al*., 2003; Sambrook and Russel, 2001).

3.6.1 Restriction digest

Both vectors pET28a(+) and pET22b_BtIDO were digested with the restriction enzymes *Xho*I and *Nde*I (FastDigest™, Thermo Fisher Scientific, USA) in separate reaction tubes. The pipetting scheme for the restriction digest (20 µL final volume) included following components: 2 µL10X Fast Digest Green Buffer (Thermo Fisher Scientific, USA), 1 µL *Xho*I and 1 µL *Nde*I, 0.5–1 µg purified plasmid (maximum volume 10 µl) and ddH_2O, respectively. The reactions were incubated at 37 °C for 30 min. Afterwards, the enzymes were heat inactivated by incubation at 80 °C for 20 min.

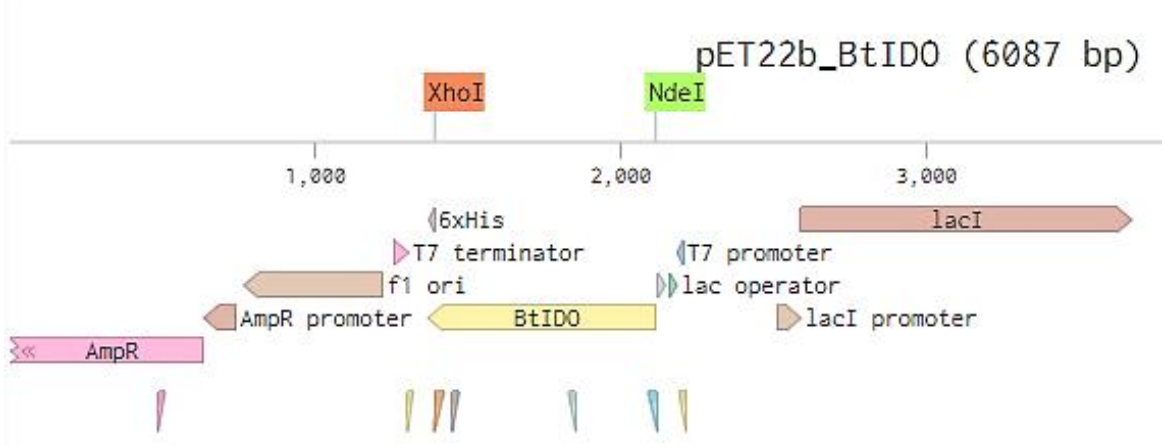

Figure 12: The restriction sites *Xho*I (red) and *Nde*I (green) were used to cut out the gene *BtIDO* (747 bp) from the vector pET22b_BtIDO (6087 bp) (adapted plasmid annotation detail from Benchling).

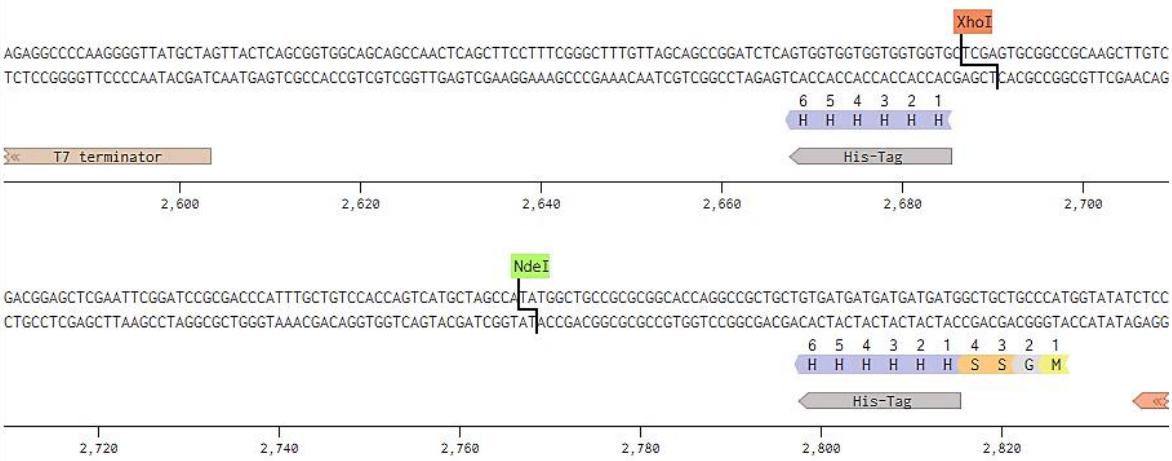

Figure 13: The stuffer fragment (81 bp) was cut-out of pET28a(+) using the restriction enzymes *Xho*I (red) and *Nde*I (green) and replaced with *BtIDO,* (adapted plasmid annotation detail from Benchling).

3.6.2 Dephosphorylation

The digested vector backbone and the corresponding cut-out stuffer fragment of pET28a(+) (Figure 12) were dephosphorylated by using FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific, USA). 1 µL phosphatase was added to the heat-inactivated restriction digest and incubated for 30 min. Then, heat inactivation of the phosphatase was performed for 5 min at 75 °C.

3.6.3 Purification

The GeneJET™ Gel Extraction Kit (Thermo Fisher Scientific, USA) was used to purify the digested fragments of pET28a(+) (vector backbone and cut-out stuffer) and pET22b_BtIDO (vector backbone and *BtIDO*). It was performed according to the manufacturer's instructions, except the elution. 20 µL prewarmed (65 °C) ddH_2O were used to elute the purified fragments instead of 50 µL elution buffer suggested in the instructions.

3.6.4 Ligation

A 20 µL ligation reaction was prepared using 100 ng digested, dephosphorylated and purified vector backbone of pET28a(+) and the corresponding amount of the insert *BtIDO* (vector-insert ratio 1:2), 1 µL ligase (T4 DNA Ligase, Thermo Fisher Scientific, USA) and 2 µL buffer (T4 DNA Ligase Buffer, Thermo Fisher Scientific, USA). The remaining volume was filled up with ddH_2O, respectively. Ligation was performed for 2 hours at room temperature. Afterwards, the enzyme was heat inactivated by incubation at 65 °C for 10 min.

3.6.5 De-salting

Before the ligation reaction could be used for electroporation, 8 µL were pipetted on a 0.025 µm filter (Type VSWP, Merck Millipore, USA) floating in ddH_2O and de-salted for 30 min.

3.6.6 Electroporation

Electrocompetent *E. coli* TOP10 cells (Table 1) were thawed on ice before 5 µL desalted ligation reaction was added. Electroporation was performed as described in Chapter 3.4. After 1.5 hours of regeneration 100 µL supernatant and the resuspended pellet were plated on LB-Kan[40] plates and incubated overnight at 37 °C.

3.6.7 Colony PCR

In order to prove the presence of cloned inserts, cell material was picked with sterile tips and resuspended in 30 µL ddH_2O. Cells were lysed at 95 °C for 10 min. After centrifugation at 13,000 rpm for 2 min, 5 µL supernatant were used as template for PCR

A Master Mix of 20 µL per reaction was prepared including 10 µL Biozym Red HS Taq Master Mix (Biozym, Germany), 0.8 µL of the forward primer mPCR_BtIDO_fwd (10 µM), 0.8 µL of the reverse primer mPCR_BtIDO_rev (10 µM) (Table 3), and 3.4 µL ddH_2O. The cycling program used is shown in Table 6 below.

Table 6: Cycling program for colony PCR using the Biozym Red HS Taq Master Mix.

PCR step	Temperature	Time	
Initial denaturation	95 °C	1 min	
Denaturation	95 °C	15 s	
Annealing	60 °C	15 s	30x
Extension	72 °C	1 min	
Final extension	72 °C	5 min	
Storage	4 °C	∞	

Afterwards, 10 µL PCR product were directly loaded onto a 1 % agarose gel and separated as described in Chapter 3.8.

3.7 Plasmid purification and sequencing

Overnight cultures were prepared as described in Chapter 3.1.

After harvesting, plasmid purification was performed with the GeneJET™ Plasmid Miniprep Kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions, except the elution, which was performed with 50 µL ddH_2O instead of Elution buffer.

Sanger Sequencing (Microsynth AG, Switzerland) was performed for the validation of created vector constructs. Therefore, 12 µL purified DNA was premixed with 3 µL in-house sequencing primer (Table 3) or primer provided by the sequencing company.

3.8 Agarose gel

For analytical separation of DNA fragments, agarose gel electrophoresis was performed using 1 %, 1.5 % or 2 % agarose gels [0.8 g, 1.2 g or 1.6 g Agarose dissolved in 80 mL 1x TAE-buffer [50x TAE stock solution: 2 M Tris, 1 M acidic acid, 0.05 M Na_2EDTA], (Sambrook *et al.*, 1989)] with the nanaPAC-300P Power Supply (Cleaver Scientific, UK). 4 µL Green Safe 20,000x (LabConsulting, Austria) was added to the gel in order to stain the separated PCR products and 5 µL of DNA ladder (GeneRuler^{TM} DNA Ladder Mix, Thermo Fisher Scientific, USA) was used as a marker. 1 µL TriTrack DNA Loading Dye (Thermo Fisher Scientific, USA) was mixed with 2 µL ddH_2O and 3 µL PCR product and loaded onto a polymerised gel. The gel was run in 1x TAE buffer at ± 115 V and 400 mA for ±45 min.

3.9 Mixed cultivation of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA* [pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE]

A mixed cultivation of the three different dioxygenases BtIDO (L-isoleucine dioxygenase from *B. thuringiensis*), AvLDO (L-leucine dioxygenase from *A. variabilis*) or GriE (leucine hydroxylase from *Streptomyces sp*. DSM40835) in the triple-knockout selection strain *E. coli* BL21(DE3) *3ΔsucA* was carried out to study the enrichment of different variants during various different cultivation cycles in an iterative approach. Further, a multiplex PCR was implemented for fast and reliable detection of the plasmids present after each cycle of cultivation. Information about the enrichment should assist in the implementation of a growth-based selection assay for the detection of improved dioxygenase variants of a mutant library.

The final setup used is depicted in the scheme below (Figure 14). Detailed information about each step is given in the sub-chapters below (3.9.1 – 3.9.4) Besides, variations tested over the course of implementation are also described in the corresponding sub-sections.

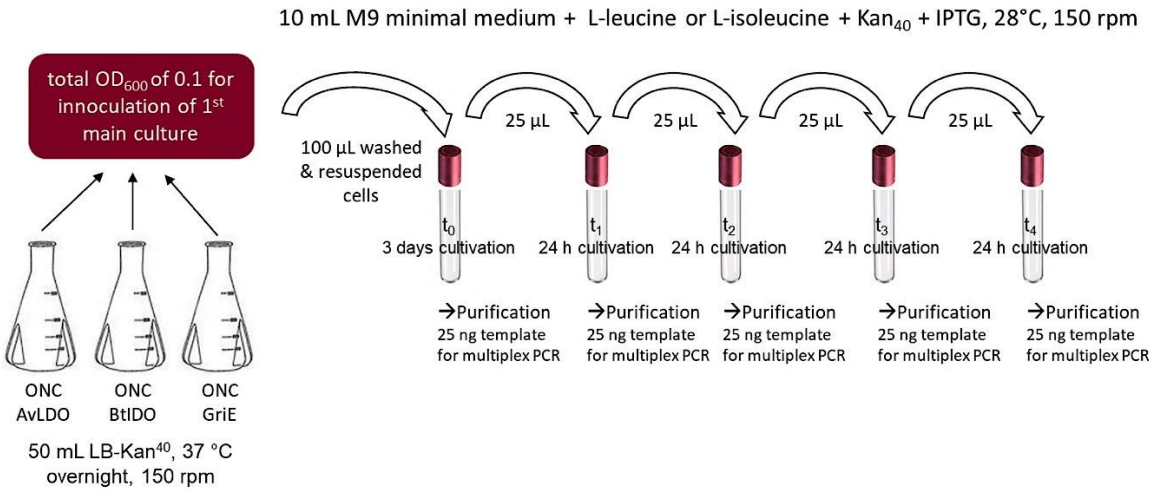

Figure 14: Schematic illustration of the mixed cultivation experiment for the implementation of a growth-based selection assay.

3.9.1 Preparation of overnight cultures

Three separate overnight cultures (ONCs) were prepared by using freshly electroporated (Chapter 3.3) *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA* [pET28a_BtIDO] or 3Δ*sucA*[pET28a_GriE] (Table 1) for inoculation (Chapter 3.4).

A second approach was carried out by directly transforming electrocompetent *E. coli* BL21(DE3) 3Δ*sucA* cells (Chapter 3.3) with 1 µL plasmid mix containing pET28a_AvLDO, pET28a_BtIDO and pET28a_GriE (19 ng/µL of each plasmid type) as described in Chapter 3.4 except the plating after regeneration. 100 µL of the supernatant were plated on a LB-Kan^{40} plate, but the remaining suspension was not pelleted and plated. It was transferred into 50 mL LB-Kan^{40} and cultivated for 72 hours at 30 °C and 120 rpm in a baffled flask.

3.9.2 Preparation of main cultures

After the determination of the optical density of each ONC, the culture volume required for inoculation of the main culture was calculated according to Formula (2).

$$V_{innoculation}[mL] = \frac{V_{main\ culture}\ [mL] * OD_{main\ culture}}{OD_{overnight\ culture}} \quad (2)$$

Two different media including M9 supplemented with L-leucine and M9 supplemented with L-isoleucine, were used for cultivation. The main culture containing 10 mL medium with Kan^{40} and 10 µL IPTG (1 M) for induction, was inoculated to an OD_{600} of 0.1. For the approach with the single ONC (mixed electroporation) Formula (2) could be used directly.

For the second approach preparing the three different ONCs, 0.034 was inserted in Formula (2) for '$OD_{main\ culture}$' to obtain the desired volume for each ONC to be harvested in order to start with equal cell densities. Further, an approach with 90 % *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] and only 5 % of 3Δ*sucA*[pET28a_BtIDO] and 5 % of 3Δ*sucA*[pET28a_GriE] was used. Since the former was lost immediately in previous mixed cultivation experiments an approach with an excess of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] was carried out.

The appropriate amounts of ONC of each strain were pooled and harvested by centrifugation for 5 min at 4 °C and 4,000 rpm. Afterwards, the supernatant was removed, and the pellet was resuspended with 500 µL M9 medium with the corresponding substrate. The centrifugation step was repeated as described before, but the pellet was resuspended in 100 µL medium. After washing twice, the suspension could be used for the inoculation of the main culture. Cultures were grown in 17 mL glass test tubes (test tube Duran® straight cut, DKW Life Sciences, Germany) for 72 hours at 28 °C and 150 rpm.

3.9.3 Iterative cultivation and plasmid purification

After 24 hours of cultivation, 25 µL culture volume were used to inoculate 10 mL of fresh minimal medium with Kan^{40} and 10 µL IPTG (1 M) supplemented with L-leucine or L-isoleucine. The remaining volume was used for plasmid preparation with the GeneJET™ Plasmid Miniprep Kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions, except the elution, which was performed with 30 µL ddH_2O instead of Elution buffer.

3.9.4 Detection of the dioxygenase genes via multiplex PCR

In order to identify the enrichment of specific plasmid harbouring clones, a multiplex PCR was implemented for the detection of pET28a_AvLDO, pET28a_BtIDO and pET28a_GriE. DreamTaq DNA Polymerase (5 U/µL) and the corresponding buffer (Thermo Fisher Scientific, USA) were used for amplification. The following Master Mix (Table 7) was prepared for a 25 µL approach per tube. All components were pipetted on ice. Generally, 25 ng plasmid (concentration determined with the NanoDrop™ 2000, (Thermo Fisher Scientific, USA)) per reaction were used as a template. The corresponding cycling program is shown in Table 8.

Table 7: Master Mix for multiplex PCR

Component	Amount
Buffer (incl. 2 mM MgCl) 10x	2.5 µL
dNTPs (2 µM)	2.5 µL
Primer Mix (1 mM each)	3.75 µL
Template	25 ng
DreamTaq (5 U/µL)	0.125 µL
ddH_2O	To 25 µL

Table 8: Cycling program for multiplex PCR

PCR step	Temperature	Time	
Initial denaturation	95 °C	1 min	
Denaturation	95 °C	30 s	30 cycles
Annealing	Gradient (55 – 70 °C)	30 s	
Extension	72 °C	1 min	
Final extension	72 °C	10 min	
Storage	4 °C	∞	

Primers used for amplification can be obtained from Table 3 and further characteristics are shown in the table below.

Table 9: Primers designed for multiplex PCR.

Primer	Length [nt]	GC-content [%]	Melting Temperature* [°C]	Length of amplified sequence [bp]
mPCR_AvLDO_fwd	24	40.00	53.98	737 bp fragment
mPCR_AvLDO_rev	24	45.83	55.60	
mPCR_BtlDO_fwd	25	45.83	55.28	405 bp fragment
mPCR_BtlDO_rev	24	41.67	54.01	
mPCR_GriE_fwd	24	41.67	53.84	549 bp fragment
mPCR_GriE_rev	24	45.83	56.64	

*Melting temperatures were calculated with Benchling using the SantaLucia 1999 algorithm

For analytical separation of DNA fragments, agarose gel electrophoresis was performed as described in Chapter 3.8.

3.10 Separate cultivation of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE]

In order to relate the enrichment of the different dioxygenase harbouring strains during the mixed cultivation experiment (Chapter 3.9) to the growth behaviour of the individual strains, the whole experiment was carried out in a separate cultivation approach. The experimental setup was the same as described above in Chapter 3.9, except the preparation of the main culture as no plasmid mixture was used for electroporation. *E. coli* BL21(DE3) 3Δ*sucA* cells were either transformed with pET28a_AvLDO, pET28a_BtIDO or pET28a_GriE. Three separate ONCs were prepared by using a single colony of each strain to inoculate 50 mL LB-Kan[40]. In the main culture, each strain was grown in both M9 medium with L-leucine and M9 medium with L-isoleucine. After each cycle of cultivation, the optical density at 600 nm was determined with the Eppendorf BioPhotometer 6131 (Eppendorf, Germany) to compare the growth behaviour of the three strains in the two different media also used in the mixed cultivation experiment.

3.11 Growth studies on dioxygenase harbouring *E. coli* BL21(DE3) 3Δ*sucA* under assay conditions

Single colonies of freshly electroporated *E. coli* BL21(DE3) 3Δ*sucA* cells were used to prepare overnight cultures as described in Chapter 3.1. For the preparation of the main culture the respective volume was harvested (Formula (2)) and washed as described in Chapter 3.9.2 to inoculate 50 mL medium supplemented with Kan[40] and 50 µL IPTG (1 M). In total, four different media were used for the cultivation of each strain including LB medium, M9 medium supplemented with L-leucine, M9 medium with L-isoleucine and M9 medium without any substrate. Main cultures were cultivated in 300 mL shake flasks.

Additionally, *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] was used as a negative control and *E. coli* BL21(DE3)[pET28a(+)] was used as a positive control to monitor non-impaired growth. Therefore, the dioxygenase harbouring strain as well as the positive and the negative control were cultivated in the four different media mentioned above resulting in the preparation of twelve cultures in total. The experiment was conducted with the dioxygenase bearing strains *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_BtIDO].

3.12 Expression studies under assay conditions

3.12.1 Culture preparation

For the preparation of overnight cultures, colony material of the required *E. coli* BL21(DE3) 3Δ*sucA* strains was used to inoculate 50 mL LB-Kan[40] according to Chapter 3.1.

The main culture containing 50 mL M9 medium with Kan[40] was inoculated to an OD_{600} of 0.1. Therefore, the appropriate amount of ONC had to be harvested and washed as described in Chapter 3.9.2. Besides the 100 µL washed cell suspension, 50 µL IPTG (1 M) was added for induction.

3.12.2 Cell harvest

In order to harvest the desired amount of cell suspension the optical density of the culture had to be determined. Formula (3) was used to calculate the appropriate culture volume for the use of 25 µL Bugbuster® Master Mix (Merck Millipore, USA).

$$\frac{0.025\ mL\ Bugbuster * 30}{OD_{600}} = culture\ volume\ [mL] \quad (3)$$

Cells were harvested by centrifugation in 15 mL tubes at 4 °C for 15 min and 4000 rpm. The supernatant was removed completely before the pellet was frozen at -20 °C to store it for the subsequent analysis via SDS-PAGE.

3.12.3 SDS-Gel

The pellet was thawed on ice and gently resuspended with 25 µL BugBuster® (Merck Millipore, USA) and incubated for 20 min at 300 rpm in an Eppendorf ThermoMixer® (Eppendorf, Germany). Equation (3) was used for the calculation of the specific amount of cell culture harvested in order to use 25 µL BugBuster®. For denaturation, 5 µL of the sample were added to 10 µL distilled water and 5 µL SDS sample buffer (4x) and heated at 95 °C for 10 min.

Besides whole cell analysis, another approach where the soluble and the insoluble fraction were separated after the treatment with BugBuster® via centrifugation (20 min, 4 °C, 13,000 rpm) was performed. The pellet was then resuspended in Tris-HCl buffer [20 mM, pH 7.4] using the same amount as BugBuster®. Both supernatant and pellet were prepared with sample buffer and distilled water as described above.

The mixture was loaded onto an ExpressPlus™ PAGE gel (12 %) (GenScript, USA). 10 µL PageRuler™ protein ladder (Thermo Fisher Scientific, USA) were used as a standard. Gels were run at 120 V and 200 mA for 1 h 20 min in Tris-MOPS running buffer (1x) (GenScript, USA) and stained with Coomassie Brilliant Blue staining solution overnight. Afterwards, gels were destained with destaining solution.

3.13 Generation of a mutant library of pET28a_AvLDO

For the creation of a mutant library of pET28a_AvLDO two positions in close proximity to the substrate binding site of L-leucine dioxygenase from *Anabaena variabilis* (AvLDO) were saturated simultaneously. Therefore, a modified QuikChange protocol including amplification of the whole plasmid by the use of oligonucleotide primers and a *Dpn*I digest (Miyazaki K, 2002) was carried out. The reaction mixture is shown in Table 10 and the cycling program with a gradient from 54 °C to 64 °C is depicted in Table 11. The aim was to saturate the isoleucine at position 70 with a codon for small or mostly hydrophobic amino acids (codon DYW) and the phenylalanine at position 78 with hydrophobic amino acids (codon NTT) and with tyrosine (TAT). Therefore, specially created primers (Table 12) were used for the saturation of the two desired positions to create a library containing 60 different variants with 35 different amino acid combinations. A master mix of seven reactions was prepared according to Table 10 using the Pfu Plus! DNA polymerase and the corresponding buffer (Roboklon GmbH, Germany). To obtain equimolar concentrations for all primers a mix for the generation of the forward primer was prepared using AvLDO_I70DYW_F78Y_fw and AvLDO_I70DYW_F87NTT_fw in a ratio of 1:5. The same mix was created for the reverse primers AvLDO_I70DYW_F78Y_rv and AvLDO_I70BYW_F78NTT_rv.

Table 10: Master mix for QuikChange reaction followed by the protocol suggested by the supplier of the polymerase (Roboklon GmbH, Germany).

Component	Volume (7 reactions)	Final concentration
ddH_2O	To 140 µL	
10x Pfu buffer	14 µL	1x
Primer fwd (10 µM)	2.8 µL	0.2 µM
Primer rev (10 µM)	2.8 µL	0.2 µM
Template (110 ng)	0.764 µL	12 ng
dNTP's (10 mM)	3.5 µL	0.25 µM
Pfu Plus! polymerase	1.4 µL	1 U

Table 11: Cycling program for QuikChange reaction.

PCR step	Temperature	Time	
Initial denaturation	95 °C	1 min	
Denaturation	95 °C	30 s	30 cycles
Annealing	Gradient (57 – 67 °C)	30 s	
Extension	68 °C	6.5 min	
Final extension	68 °C	10 min	
Storage	4 °C	∞	

Table 12: Degenerated oligonucleotide primers for QuikChange reaction (Schweiger, unpubl.) including the codons DYW (encoding for A, F, I, L, S, V, T), NTT (encoding for F, I, L, V) and TAT (encoding for Y).

Name	Mutation	Sequence (5' – 3')	T_m [°C]
AvLDO_I70DYW_F78Y_fw	I70DYW_F78Y	tgtgtatcctaaa**DYW**gagcgaattggcatcacagtg**TAT**gaatataaccaattagc	64.5
AvLDO_I70DYW_F78Y_rv	I70DYW_F78Y	ggttatattc**ATA**cactgtgatgccaattcgctc**WRH**tttaggatacacatcttgataac	63.2
AvLDO_I70DYW_F87NTT_fw	I70DYW_F78NTT	tgtgtatcctaa**DYW**gagcgaattggcatcacagtg**NTT**gaatataaccaattagc	63.7
AvLDO_I70DYW_F78NTT_rv	I70DYW_F78NTT	ggttatattc**AAN**cactgtgatgccaattcgctc**WRH**tttaggatacacatcttgataac	63.1

Table 13: Possible variants generated with the degenerated primers

Codon	Variants	A	F	I	L	S	T	V	Y
DYW	12	GCA	TTT	ATA	TTA	TCA	ACA	GTA	
		GCT		ATT		TCT	ACT	GTT	
NTT	4		TTT	ATT	CTT			GTT	
TAT	1								TAT

After the reaction had finished, 3 µL were taken for gel analysis (1 % agarose gel) according to Chapter 3.8. To the remaining 17 µL of PCR reaction 0.5 µL *Dpn*I was added for digestion of the methylated parental as well as hemi-methylated parental/mutant hybrid DNAs (2.5 h at 37 °C, 20 min at 80 °C).

All positive reactions were pooled and used for transformation of electrocompetent *E. coli* TOP10 cells. After de-salting (Chapter 3.6.5), 5 µL of the mixture were used for electroporation, which was carried out as described in Chapter 3.6.6. After regeneration for 1.5 hours, 100 µL of a 1:10 dilution was prepared, plated on LB-

Kan^{40} plates incubated overnight at 37 °C. The remaining volume was used to inoculate 10 mL LB-Kan^{40} and was cultivated as described in Chapter 3.1.

10 clones were picked from the plate to prepare overnight cultures (Chapter 3.1) for plasmid purification in order to verify the success of mutation by Sanger sequencing as described in Chapter 3.7. Further, the liquid culture was purified to obtain the plasmid library. Electrocompetent *E. coli* TOP10 were then transformed with 1 µL library as described in Chapter 3.4 and the regenerated cell suspension was then again used to inoculate 5 mL LB-Kan^{40} in order to prepare glycerol stocks for the storage of the library (Chapter 3.2).

3.14 Selection assay for the enrichment of improved AvLDO variants

In order to detect improved variants of a library generated with site-directed saturation mutagenesis, a selection assay (Figure 15) employing the triple knockout strain *E. coli* BL21(DE3) 3Δ*sucA* (Chapter 1.5) was carried out.

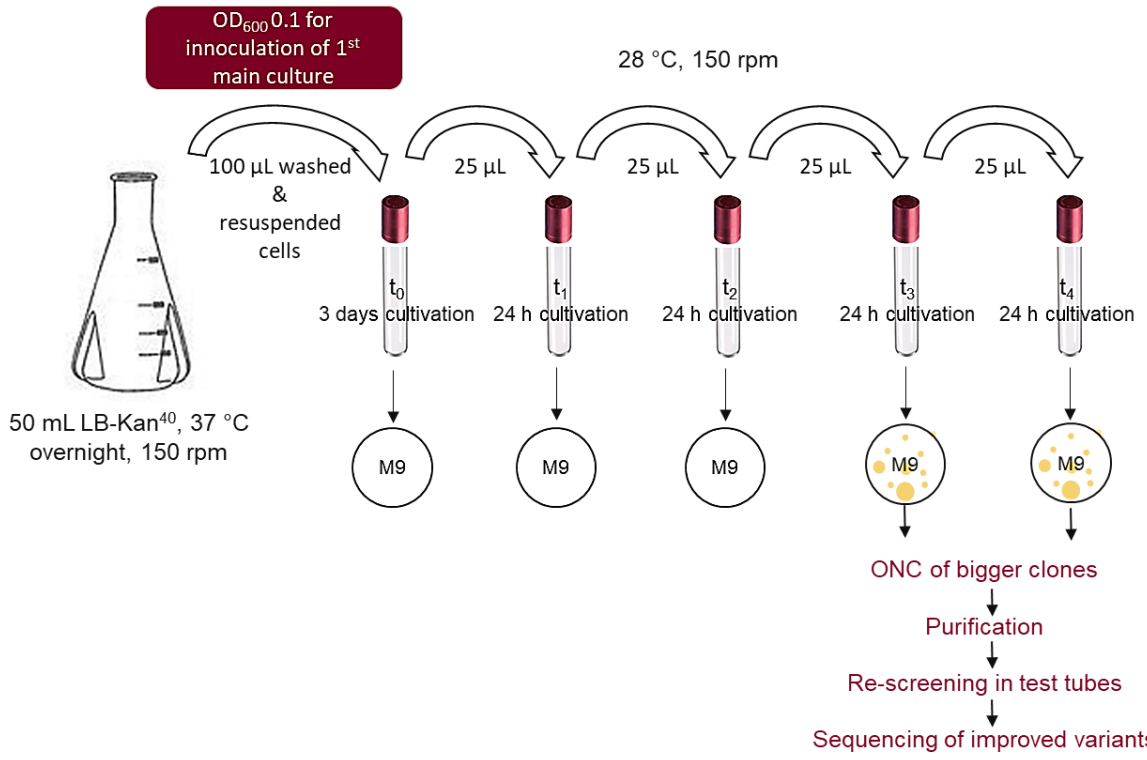

Figure 15: Schematic illustration of the growth-based selection assay implemented for the detection of improved dioxygenase variants.

Therefore, 1 µL (~100 ng/µL) of library was used to electroporate the electrocompetent 3Δ*sucA* cells (Chapter 3.4). Afterwards, 10 µL of the regenerated cell suspension was plated on LB-Kan^{40} plates to evaluate the transformation

efficiency. The remaining volume was used to prepare an ONC as described in Chapter 3.1. Cells were harvested to inoculate 10 mL M9 supplemented with L-isoleucine, Kan[40] and IPTG to an OD_{600} of 0.1 as described in Chapter 3.9.2. The first main culture was grown for 72 hours at 28 °C and 150 rpm to enable adaption to the minimal medium. Then, 25 µL were used to inoculate fresh medium. All further cultures were then grown for only 24 hours under the same conditions as the first main culture to promote selection of improved variants. After a few cycles, 100 µL of each liquid culture were plated on M9-agar antibiotic plates (Chapter 2.5.7) additionally. Therefore, different dilutions in a range of $1:(4 \times 10^6)$ to $1:(4 \times 10^8)$ were prepared to achieve the growth of single colonies facilitating the detection of variants larger in size. Plates were cultivated for 48 to 72 hours at 30 °C, depending on the growth cycle. Clones showing increased size were picked and plated on LB-Kan[40] plates (incubation for 37 °C overnight) for storage at 4 °C. Afterwards, selected clones were cultured and purified as described in Chapter 3.7.

3.15 Re-screening of selected variants

Re-screening was performed by transforming *E. coli* BL21(DE3) 3Δ*sucA* cells with 1 µL of purified plasmid of the selected clones (Chapter 3.4). Additionally, the selection strain was also transformed with pET28a(+) as a negative control and with the wild type pET28a_AvLDO for comparison. Overnight cultures of the different strains were prepared by using a single colony of freshly electroporated cells (Chapter 3.4) to inoculate 50 mL LB-Kan[40]. Main cultures were prepared by inoculating 10 mL M9 medium supplemented with L-isoleucine and Kan[40] as stated in Chapter 3.9.2. Cultivation was performed in 17 mL glass test tubes (test tube Duran® straight cut, DKW Life Sciences, Germany) and 10 µL IPTG (1 M) were used for induction. Sampling of 100 µL culture volume, which was diluted 1:10 with sterile medium was performed every few hours and the OD_{600} was determined with the Eppendorf BioPhotometer 6131 (Eppendorf, Germany) to obtain growth curves of each strain. For comparison of the different mutants, OD values of the exponential growth phase of each strain were consulted to calculate specific growth rates (μ) as described below (Formula (4),(Lindqvist and Barmark, 2014)).

$$\mu\,[h^{-1}] = \frac{\Delta ln OD_{600}}{\Delta t[h]} \qquad (4)$$

Additionally, the same variants were also re-screened in an iterative approach as performed in the separate cultivation of three different dioxygenase harbouring strains (Chapter 3.10). The setup was the same as used for the growth-based selection assay (Chapter 4.7). Before each new inoculation step, the optical density of the culture was determined.

Purified plasmids of the mutant strains picked for re-screening were then analysed via Sanger sequencing as described in Chapter 3.7.

3.16 Verification of L-isoleucine conversion of selected variants via thin layer chromatography

Variants selected in the re-screening, which showed faster growth than the wild type in minimal medium supplemented with L-isoleucine, were analysed via thin layer chromatography (TLC) to verify the conversion of the respective substrate.

3.16.1 Expression

Overnights cultures were prepared using freshly electroporated *E. coli* BL21(DE3) cells bearing the respective variant to inoculate 10 mL LB-Kan[40]. Cultivation was performed as described in Chapter 3.1. 10 mL ONC was used to inoculate 200 mL TB-Kan[40]. Cells were grown in 1 L baffled flasks at 37 °C and 120 rpm to an OD_{600} of 0.5 – 0.8 before 0.5 mM IPTG was used for induction. Expression was performed at 20 °C and 120 rpm overnight for 18 – 22 hours.

3.16.2 Harvest and cell disruption

After expression overnight, cells were harvested by centrifugation for 20 min at 4 °C and 8,000 rpm. 1 mL Tris-HCl buffer [20 mM, pH 7.4, 400 mM NaCl] per 100 mg cells was used to resuspend the cell pellet. Disruption was performed by sonication with the Branson sonifier (Thermo Fisher Scientific, USA), (Duty cycle 50 %, Output control 5 - 6, 2x2 min (cooling in between)). The lysate was cleared by centrifugation (15,000 g, 30 min, 4°C) and the obtained supernatant was sterilised by filtration (pore size: 0.22 µm).

3.16.3 Determination of protein concentration

The protein concentration was determined performing a bicinchoninic acid (BCA) assay (Pierce™ BCA Protein Assay Kit, Thermo Fisher Scientific, USA) in microplates according to the manufacturer's instructions. Bovine serum albumin was used as a standard to obtain a calibration curve in the range of 0.1 – 1 mg/mL.

All samples were prepared in triplicates and the absorption was measured in with a plate reader at 562 nm (FLUOstar® Omega, BMG Labtech, Germany).

3.16.4 SDS-Gel

Before cell harvest, a sample of each expression culture was taken for analysis of the whole cell profile as described in Chapter 3.12.3 using the formula below.

$$V_{sample}\,[mL] = \frac{7}{OD_{600}} \tag{5}$$

The amount of BugBuster® was calculated according to Formula (7). Cell disruption and sample preparation was performed as described in Chapter 3.12.3.

$$V_{Bugbuster}\,[mL] = \frac{OD_{600} \times V_{sample}\,[mL]}{30} \tag{6}$$

Additionally, samples of the CFE were also prepared by using 5 µg of sample. The volume was calculated according to protein concentrations obtained from the BCA assay. The remaining volume was filled up with ddH_2O and mixed with 5 µL SDS sample buffer (4x). Conditions for denaturation, loading and running the gel were the same as already stated in Chapter 3.12.3.

3.16.5 Activity Assay

The composition of the standard reaction mixture is shown in Table 14. All stock solutions were prepared freshly to prevent oxidation or decomposition of $FeSO_4$ and ascorbate. L-isoleucine was used as a substrate to test the activity of all variants and compare it to the activity of the AvLDO wild type. Further, L-leucine was applied for the wild type. As the latter substrate is known to be converted by AvLDO from previous studies (Hanreich, 2019), the reaction with L-leucine should serve as a positive control. An additional reaction using the same amount of Tris-HCl buffer instead of CFE was set up as a blank to enable the detection of unspecifically degraded reaction compounds. Biocatalytic reactions were performed in 1.5 mL glass vials with a total reaction volume of 0.5 mL at 25 °C and 750 rpm. Samples of 50 µL were taken after 0, 1, 2, 18, 24 and 92 hours. After adding 25 µL acetic acid (50 % v/v) to quench the reaction, the sample was vortexed and centrifuged (1 min, 4 °C, 16,000 rpm). The supernatant was transferred into a fresh reaction tube and stored at – 20°C for further analysis via TLC.

Table 14: All reaction components of the activity assay.

Components	Concentration of stock solution	Final concentration	Final volume [µL]
Substrate	50 mM	5 mM	50 µL
α-Ketoglutarate	100 mM	10 mM	50 µL
Ascorbate	100 mM	10 mM	50 µL
$FeSO_4x7H_2O$	50 mM	0.5 mM	5 µL
Sodium acetate buffer (pH 4)	1 M	50 mM	25 µL
CFE			320 µL

3.16.6 Thin layer chromatography

The aim was to verify the conversion of L-isoleucine to L-hydroxyisoleucine by selected variants. Therefore, 2 µL of reaction mixture were pipetted onto a silica gel TLC plate (TLC Silica gel 60 F_{254}, Merck Millipore, USA). Chromatograms were developed by using 1-butanol, acetic acid and ddH_2O in a ratio of 22:3:5 as a mobile phase. After drying the plates, spots were detected with staining solution of ninhydrin (0.2 % w/v dissolved in acetone). L-leucine, L-isoleucine and L-hydroxyisoleucine served as standards and were prepared by mixing 10 µL of the standard mixture (5 mM amino acid) with 5 µL acetic acid. Formula (7) was used to calculate the retention factor (R_f).

$$R_f = \frac{distance\ spot\ moved}{distance\ solvent\ moved} \quad (7)$$

4 Results

4.1 Growth studies with dioxygenase harbouring *E. coli* BL21(DE3) 3Δ*sucA*

4.1.1 *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO]

The ampicillin resistance gene expressed by *E. coli* BL21(DE3) 3Δ*sucA*[pET22b_BtIDO] underlies a different mechanism than the kanamycin resistance gene expressed by *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO]. Therefore, growth studies previously performed with *E. coli* BL21(DE3) 3Δ*sucA*[pET22b_BtIDO] by (Hanreich, 2019) had to be repeated with the newly constructed 3Δ*sucA*[pET28a_BtIDO] (Chapter 3.6). The triple-knockout strain 3Δ*sucA* harbouring the dioxygenase containing plasmid pET28a_BtIDO was cultivated in four different media (LB with antibiotic, M9 with antibiotic, M9 with antibiotic and L-leucine, M9 with antibiotic and L-isoleucine) to monitor the growth behaviour under assay conditions (Chapter 3.11). *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] was used a negative control and *E. coli* BL21(DE3)[pET28a(+)] as a positive control.

Figure 16 (a) shows the growth in LB medium. However, significantly lower growth would be expected due to results of previous growth studies in LB using *E. coli* BL21(DE3) 3Δ*sucA*[pET22b_BtIDO] (Hanreich, 2019). Since the cultures showed an intense yellow colour, it was assumed that contamination with bacterial strains only growing in LB, but not in minimal medium occurred, as the other media used did not show that prominent dark yellow colour.

In Figure 16 (b) the growth of the positive control *E. coli* BL21(DE3)[pET28a_BtIDO], the negative control *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] as well as the knockout strain 3Δ*sucA*[pET28a_BtIDO] were cultivated in M9 minimal medium with antibiotic, but without any substrate. As expected, the positive control showed highest growth with the exponential phase starting after 20 hours and the negative control did not show significant growth throughout the experiment. *E. coli* BL21(DE3) 3Δ*sucA* [pET28a_BtIDO] started to grow after 90 hours with a specific growth rate of 0.035 h^{-1} .

The growth in M9-antibiotic containing L-leucine is shown in Figure 16 (c) and containing L-isoleucine in Figure 16 (d). In accordance with previous studies using

E. coli BL21(DE3) 3Δ*sucA*[pET22b_BtIDO] , the strain showed earlier growth by the use of L-isoleucine as external substrate source with an exponential growth after 40 hours, a specific growth rate of 0.166 h^{-1} and a maximal OD_{600} of 3.21 compared to L-leucine, where exponential growth starts after about 80 hours with a lower specific growth rate of 0.138 h^{-1}.

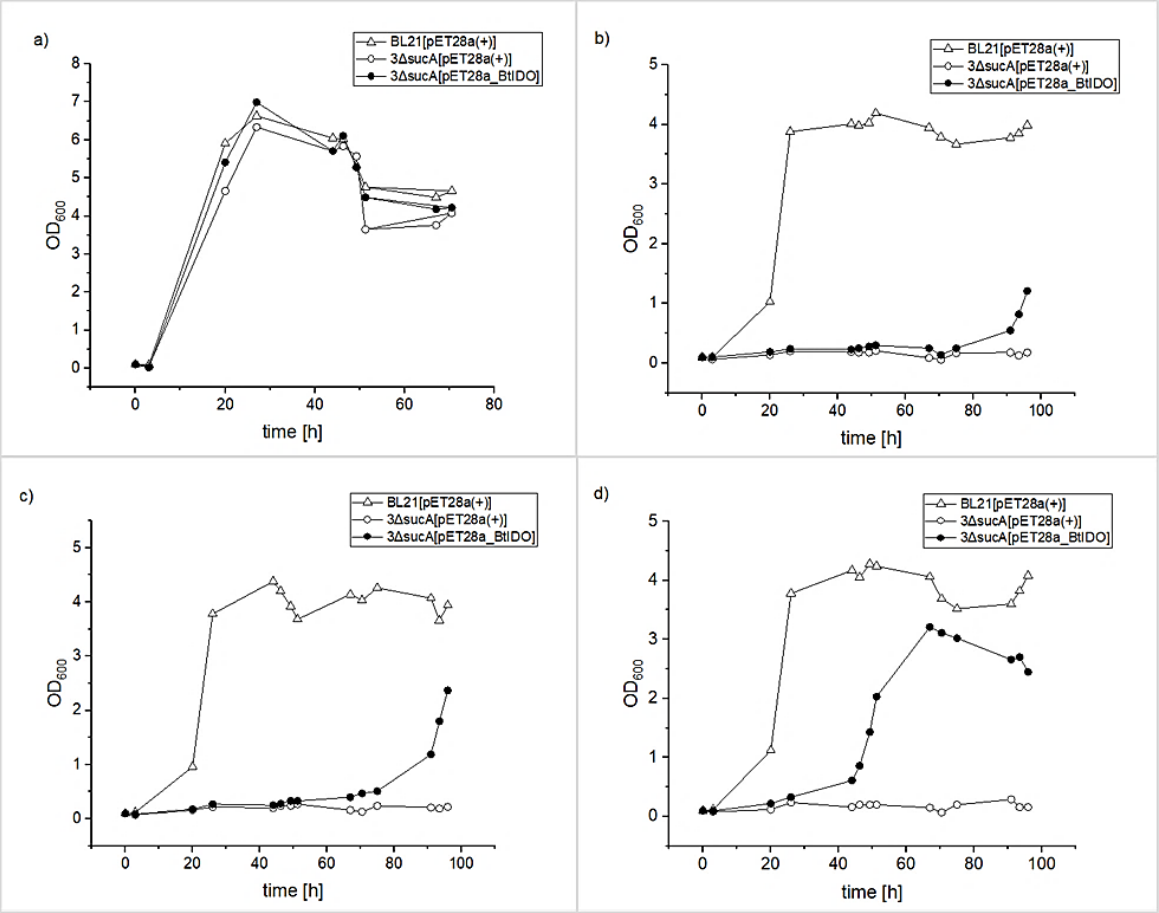

Figure 16: Growth studies with BtIDO. Comparison of the growth behaviour of *E. coli* BL21(DE3) harbouring pET28a(+) (open triangles) and *E. coli* BL21(DE3) 3Δ*sucA* carrying pET28a_BtIDO (closed circles) or the empty vector pET28a(+) (open circles). Recombinant strains were cultivated in four different media: **(a)** LB medium supplemented with kanamycin **(b)** M9 minimal medium supplemented with kanamycin but without any substrate **(c)** M9 minimal medium supplemented with kanamycin and L-leucine **(d)** M9 minimal medium containing kanamycin and L-isoleucine. As expected, higher growth rates were observed for *E. coli* BL21(DE3) cultivated in minimal medium (b,c,d) compared to the mutant strain 3Δ*sucA*. *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO] reached the exponential growth phase after 40 hours in L-isoleucine (d) and after 80 hours in L-leucine (c).

4.1.2 *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO]

As the backbone of pET28a_AvLDO (Ruhr Universität Bochum) was exchanged by the backbone of pET28a_GriE (GenScript Biotech B.V. (Netherlands)) in this study due to point mutations in the backbone, the growth behaviour of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] harbouring the new construct was studied in minimal medium supplemented with L-leucine, L-isoleucine or without any substrate and in LB medium to gain insight into the substrate utilisation and the connected growth behaviour of the strain under assay conditions (Chapter 3.11). *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] was used a negative control and *E. coli* BL21(DE3)[pET28a(+)] as a positive control. As expected, the positive control *E. coli* BL21(DE3)[pET28a(+)] started its exponential growth earlier than the knockout strain in all four types of media used (Figure 17). In LB medium the positive control started its growth after approximately five hours and reached a final OD_{600} of 6 after 25 hours of growth (Figure 17 (a)), whereas in M9 medium strong exponential growth started after 20 hours of cultivation and final ODs between 4 and 5 were reached for all three types of media, M9 without substrate (Figure 17 (b)), M9 supplemented with L-leucine (Figure 17 (c)) and M9 supplemented with L-isoleucine (Figure 17 (d)). Regarding the dioxygenase harbouring knockout strain *E. coli* BL21(DE3) 3Δ*sucA*[pET28_AvLDO], exponential growth started after 120 hours in minimal medium with L-leucine (Figure 17 (c)) and after 140 hours in minimal medium supplemented with L-isoleucine (Figure 17 (d)) reaching final ODs of 3.95 and 4.28 after about 20 – 25 hours of exponential growth displaying specific growth rates of 0.117 h^{-1} and 0.087 h^{-1}. Further, *E. coli* BL21(DE3) 3Δ*sucA*[pET28_AvLDO] also started to grow in M9 medium without any substrate (Figure 17 (b)) after 100 hours reaching a final OD of 3.7 after 143 hours (0.063 h^{-1}), which can be explained by hydroxylation of endogenously formed L-leucine as it was reported for L-proline in previous studies (Theodosiou *et al*., 2017).

However, the growth of the negative control *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] in M9 medium with L-isoleucine (Figure 17 (d)) after 100 hours of cultivation with a specific growth rate of 0.090 h^{-1}, before the dioxygenase harbouring strain 3Δ*sucA*[pET28a_AvLDO] started to grow was an unexpected observation. As samples of sterile controls of all four media were taken regularly and no cell growth could be detected, a contamination-induced growth is rather unlikely.

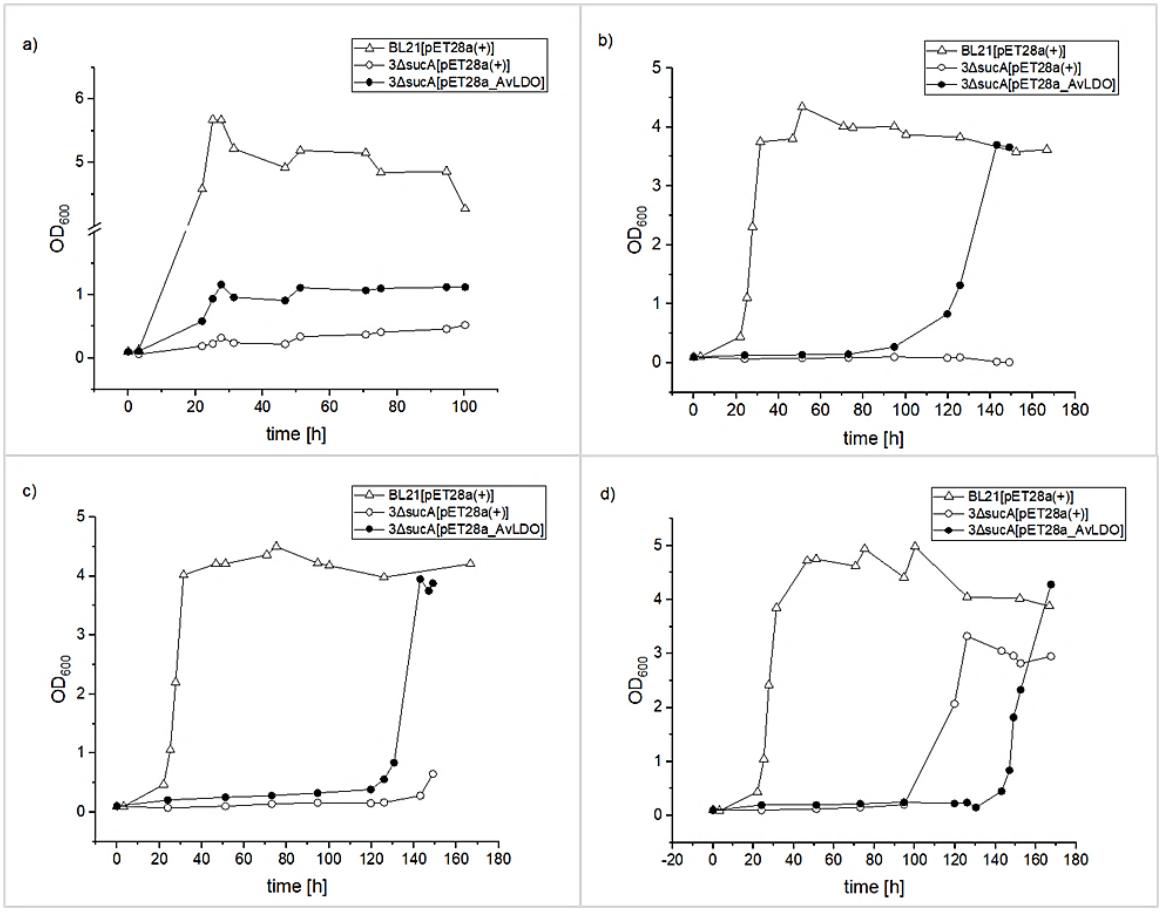

Figure 17: Growth studies with AvLDO. Comparison of the growth behaviour of *E. coli* BL21(DE3) harbouring pET28a(+) (open triangles) and *E. coli* BL21(DE3) 3Δ*sucA* carrying pET28a_AvLDO (closed circles) or the empty vector pET28a(+) (open circles). Recombinant strains were cultivated in four different media: **(a)** LB medium supplemented with kanamycin **(b)** M9 minimal medium supplemented with kanamycin but without any substrate **(c)** M9 minimal medium supplemented with kanamycin and L-leucine **(d)** M9 minimal medium containing kanamycin and L-isoleucine. As expected, earlier growth and higher growth rates were observed for *E. coli* BL21(DE3) with unimpaired metabolism. *E. coli* BL21(DE3) 3Δ*sucA* [pET28a_AvLDO] (closed circles) started exponential growth after 120 hours of cultivation in L-leucine **(c)** and after 140 hours in L-isoleucine (d) reaching similar final ODs of ~4.

4.2 Expression studies under assay conditions

For the performance of expression studies using the triple-knockout strain *E. coli* BL21(DE3) 3Δ*sucA* the same conditions as used in the mixed cultivation of dioxygenases (Chapter 3.9) were applied. The aim was to study the enzyme formation in the course of cultivation. The expression of the recombinant dioxygenases AvLDO (L-leucine dioxygenase from *A. variabilis*), BtIDO (L-isoleucine dioxygenase from *B. thuringiensis*) and GriE (L-leucine dioxygenase from *Streptomyces sp*. DSM40835) was analysed by SDS-gel. The samples harvested at t_0 should serve as negative controls, as no proteins, except endogenous *E. coli* proteins should be present at that time. Figure 18 shows the expression of BtIDO under assay conditions. As expected, at t_0 the gene was not expressed, and only low amounts of endogenous *E. coli* proteins can be observed in cultivation of both M9 medium with L-isoleucine and M9 medium with L-leucine. For minimal medium with L-isoleucine the highest amount of BtIDO was formed after 24 hours (t_1). Protein formation declined after 48 and 72 hours and no specific bands for BtIDO could be observed at all afterwards (96 hours (t_4) and 165 hours (t_5)). In return, the amount of unspecific proteins formed increased after 24 hours. For the cultivation in minimal medium supplemented with L-leucine a faint band can be observed at t_1. Afterwards, the protein formation increased and bands after 48 hours and 72 hours (t_3 and t_4) showed a higher intensity. At t_4 and t_5 no formation of BtIDO can be observed, but a high amount of unspecific protein is formed.

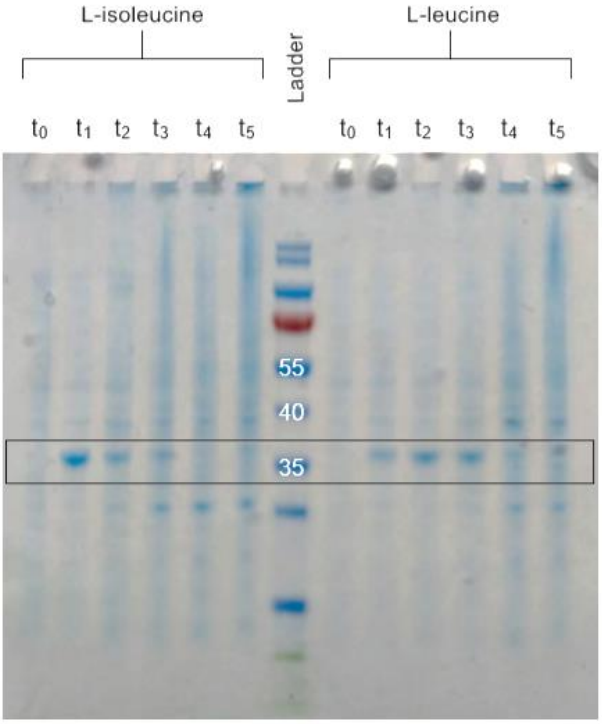

Figure 18: Expression of BtIDO by *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO] analysed by SDS-PAGE. BtIDO (29 kDa) is presumed to be formed 24 hours (t_1), 48 hours (t_2) and 72 hours (t_3) after induction in both L-isoleucine and L-leucine. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

The expression of AvLDO and GriE is shown in Figure 19 and Figure 20. Equally to the expression of BtIDO, AvLDO and GriE are not expressed at t_0, except for GriE in M9 medium with L-isoleucine. The small amount of protein formed at t_0 in minimal medium with L-isoleucine (Figure 20) could be the result of a weak basal expression (Novy and Morris, 2001) of the pET system used. In contrast to the expression of BtIDO, for both AvLDO and GriE highest amounts of enzyme were formed after 24 hours of induction (t_1) in minimal medium supplemented with L-leucine. At t_4 and t_5 no enzyme formation could be observed anymore, which could be linked to the fact that the bacterial cells already reached stationary or death phase. Regarding the cultivation in minimum medium supplemented with L-isoleucine, same amounts of AvLDO (Figure 19) were expressed from 24 hours after induction (t_1) to 72 hours of induction (t_3). For GriE (Figure 20) expression was higher from 48 to 72 hours (t_2, t_3) compared to the sample taken after 24 hours of cultivation (t_1). Equally to BtIDO, in the stationery and death phase (144 hours (t_4) and 168 hours (t_5)) no clear bands were observed anymore in both M9 medium supplemented with L-leucine and L-isoleucine correlating with the growth progression of the strains. When *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and

3Δ*sucA*[pET28a_GriE] reached their stationary phase, no protein formation could be detected via SDS-gel.

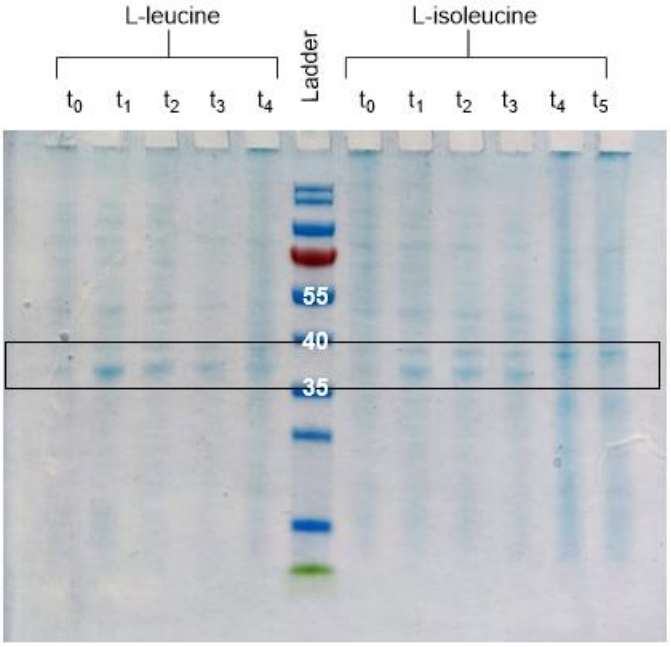

Figure 19: Expression of AvLDO by *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] analysed by SDS-PAGE. AvLDO (33 kDa) is presumed to be formed 24 hours (t_1), 48 hours (t_2) and 72 hours (t_3) after induction in both L-isoleucine and L-leucine. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

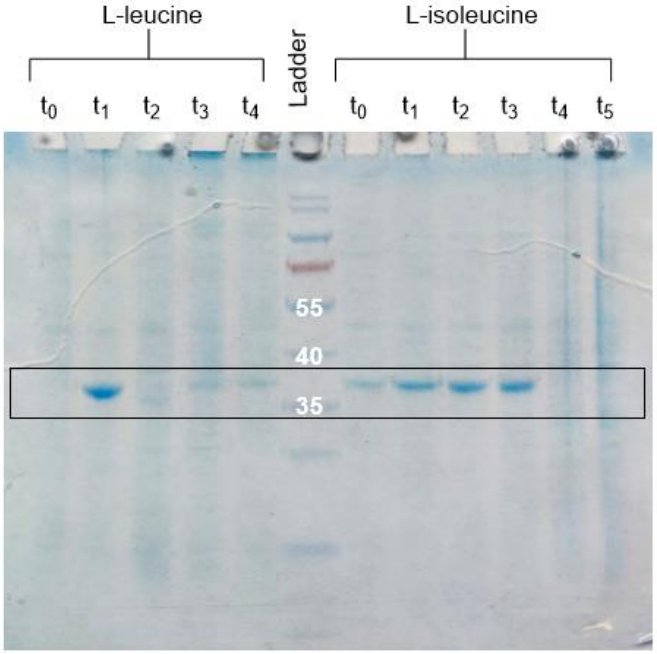

Figure 20: Expression of GriE by *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_GriE] analysed by SDS-PAGE. GriE (32 kDa) is presumed to be formed 24 hours (t_1), 48 hours (t_2) and 72 hours (t_3) after induction in both L-isoleucine and L-leucine. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

4.3 Multiplex PCR

A multiplex PCR was established to detect dioxygenase enrichment of AvLDO, BtIDO and GriE in a mixed growth using the selection strain *E. coli* BL21(DE3) 3Δ*sucA*. Therefore, primer pairs designed (Table 9) were tested in separate reactions using the respective template pET28a_AvLDO, pET28a_BtIDO or pET28a_GriE. After amplification of the intended fragment by the respective primer pairs was verified via gel electrophoresis, the primer mix including all three oligo pairs was used for the amplification of each individual fragment. As unspecific binding of other primer pairs could be excluded as well under the conditions applied, all three templates pET28a_AvLDO, pET28a_BtIDO and pET28a_GriE were amplified in one PCR reaction including the six primers shown in Table 9. The result of this final reaction is shown in Figure 21. Similar intensities of the fragments AvLDO (737 bp), GriE (549 bp) and BtIDO (405 bp) can be observed under the PCR conditions applied (Table 7, Table 8) for all annealing temperatures from 60 to 65 °C. Since the same amount of template was used for pET28a_AvLDO, pET28a_BtIDO and pET28a_GriE and an equimolar primer mix was used, it can be stated that the multiplex PCR established enables a reliable detection of the three different dioxygenases in a mixed cultivation.

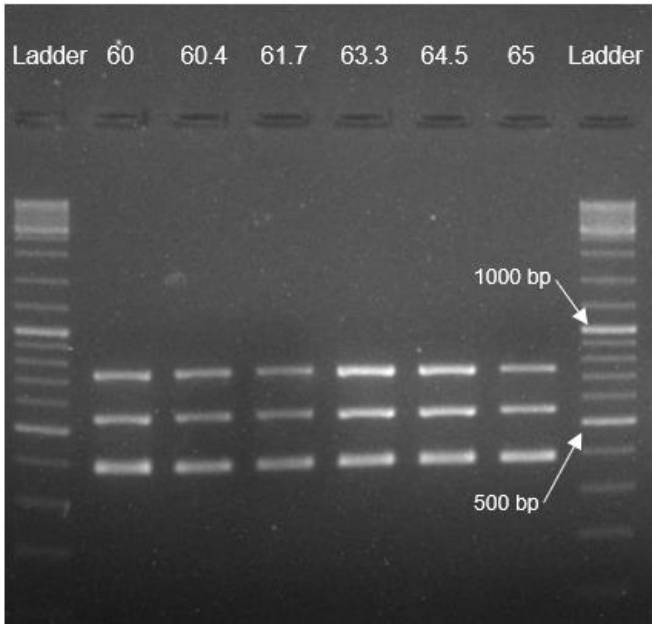

Figure 21: Gradient multiplex PCR (60-65 °C) using pET28a_BtIDO, pET28a_AvLDO and pET28a_GriE as a template. A primer mix including the primer pairs for all three templates (mPCR_BtIDO_fwd, mPCR_BtIDO_rev, mPCR_AvLDO fwd, mPCR AvLDO rev, mPCR_GriE_fwd, mPCR_GriE_rev) was used. The PCR works for all annealing temperatures used and no unspecific PCR product formation can be observed. GeneRuler™ DNA Ladder Mix (Thermo Fisher Scientific, USA) was used as a marker.

4.4 Mixed cultivation of *E. coli* BL21(DE3) 3Δ*sucA* [pET28a_AvLDO], 3Δ*sucA* [pET28a_BtIDO] and 3Δ*sucA* [pET28a_GriE]

The multiplex PCR implemented in the previous Chapter (4.3) was used to detect dioxygenase enrichment of a mixed culture of the triple-knockout selection strain *E. coli* BL21(DE3) *3ΔsucA* harbouring either BtIDO (L-isoleucine dioxygenase from *B. thuringiensis*), AvLDO (L-leucine dioxygenase from *A. variabilis*) or GriE (leucine hydroxylase from *Streptomyces sp*. DSM40835). The aim was to figure out which cultivation intervals promote the enrichment of one of the three dioxygenases. The information obtained from this experiment should be further used for the implementation of a growth-based selection assay to detect improved variants of a mutant library. During implementation, different cultivation periods including 24 hours and 48 hours as well as an approach based on the optical density of the cultures were carried out. Summarising all the information obtained from these experiments, it can be stated that *E. coli* BL21(DE3) *3ΔsucA*[pET28a_GriE] always enriched in M9 medium supplemented with L-leucine and *3ΔsucA*[pET28a_BtIDO] always enriched in M9 medium supplemented with L-isoleucine, whereas *3ΔsucA*[pET28a_AvLDO] got lost rapidly in both media. Further, it was observed that the strains need some time for adaption after switching from LB medium to minimal medium, otherwise they show poor growth during the following cultivation cycles. Since the growth of the strains changed from cycle to cycle an OD-based approach, where new medium is inoculated at a specific optical density, turned out to be inconvenient. The comparison of 24 hours cultivation and 48 hours of cultivation per growth cycle indicated that 24 hours per cycle have the potential to promote the enrichment of better variants *e.g. 3ΔsucA*[pET28a_BtIDO] got lost in L-leucine and *3ΔsucA*[pET28a_GriE] was enriched. On the contrary, both strains were detected after the final growth cycle using 48-hour intervals for growth, which indicates, that the selection pressure is rather weak in this case (Figure 22).

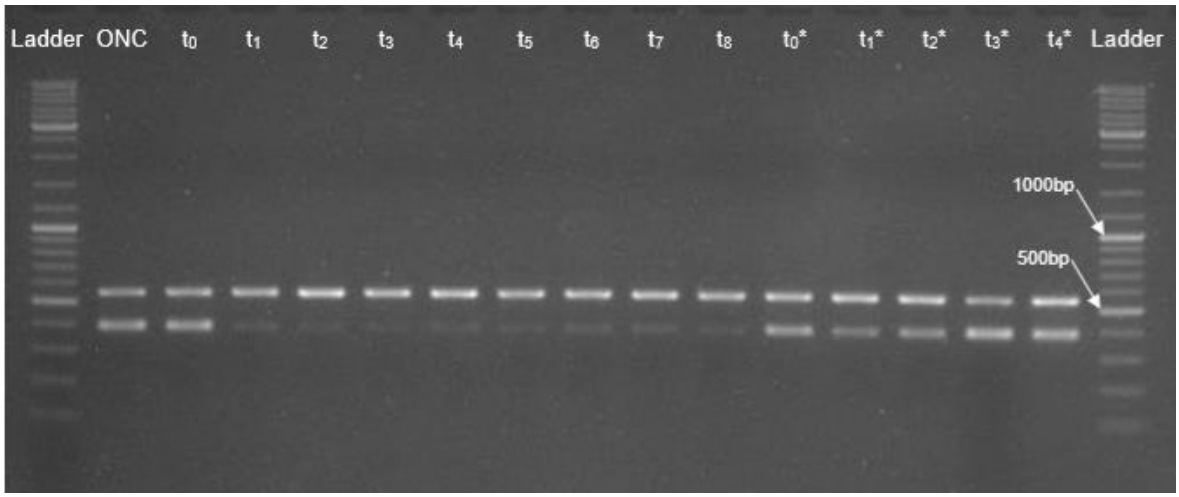

Figure 22: Multiplex PCR of mixed cultures containing *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE] harvested and purified after 24 hours (t_0 – t_8) and 48 hours (t_0* - t_4*). After the respective cultivation period the culture was used for inoculation of new M9 medium with L-leucine and a plasmid miniprep was performed with the remaining culture volume. The purification of the ONC (prepared by regeneration suspension of mixed electroporation) shows that the plasmids pET28a_BtIDO (405 bp), pET28a_GriE (549 bp) were present after the mixed transformation in *E. coli* BL21(DE3) *3ΔsucA*, whereas pET28a_AvLDO (737 bp) could not be detected in the purified plasmid mixture. Consequently, pET28a_AvLDO was not present in all subsequent samples as well. The cultivation shows, that the plasmid pET28a_BtIDO almost gets lost in the 24 hours-based approach at t_1, but remains in the 48 hours-based cultivation cycles, indicating a longer adaption time of the strain in *E. coli* BL21(DE3) *3ΔsucA*[pET28a_BtIDO] compared to 3Δ*sucA*[pET28a_GriE].

Due to obvious different transformation efficiencies of the plasmids pET28a_AvLDO, pET28a_BtIDO and pET28a_GriE, the preparation of three individual overnight cultures was preferred for the inoculation of the main culture. For instance, in Figure 22 pET28a_AvLDO could not be detected in the plasmid mixture isolated from the overnight culture.

Figure 23 shows the results of the final setup used for mixed cultivation as illustrated schematically in Figure 14.

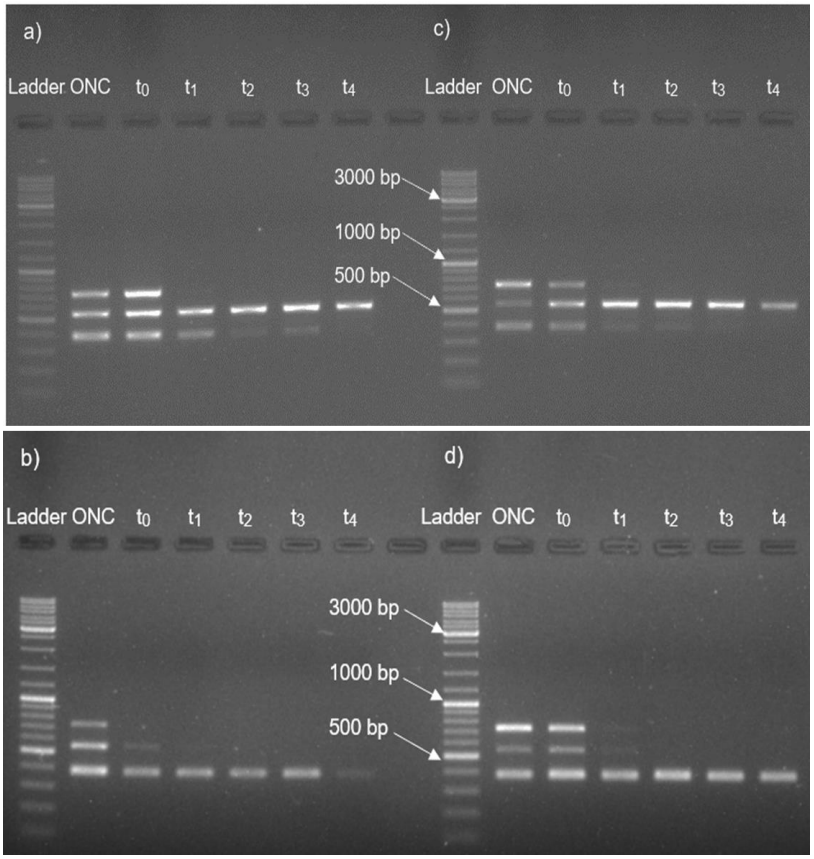

Figure 23: Multiplex PCR of mixed cultures containing *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] (737 bp), 3Δ*sucA*[pET28a_BtIDO] (405 bp) and 3Δ*sucA*[pET28a_GriE] (549 bp). The ONC sample is a mix of three separate ONCs of the respective strains in a ratio of 1:1:1 **(a,b)** 18:1:1 **(c,d)** for 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE]. Volumes harvested are adapted according to the OD of each culture to obtain a final OD of 0.1 for inoculation of the main culture containing M9 minimal medium with IPTG and Kan[40] supplemented with L-leucine **(a,c)** or L-isoleucine **(b,d)**. The sample t_0 was cultivated for 3 days to enable adaption to the minimal medium. The samples t_1-t_4 were cultivated for only 24 hours to enrich predominant strains. 25 ng purified plasmid were used per sample. GeneRuler™ DNA Ladder Mix (Thermo Fisher Scientific, USA) was used as a marker.

A separate ONC of each strain was prepared and the first main culture (t_0) was inoculated to an OD_{600} of 0.1 and grown for 72 hours to enable adaption, whereas 24 hour intervals were selected for all further cycles (t_1 – t_4). Figure 23 (a) and (c) show the cultivation in minimal medium with L-leucine and Figure 23 (b) and (d) in L-isoleucine. Whereas the same number of cells was used for inoculation of the first main culture (t_0) in Figure 23 (a) and (b) (ratio 1:1:1 for *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE]), an excess of 3Δ*sucA*[pET28a_AvLDO] was used in in Figure 23 (c) and (d). Since 3Δ*sucA*[pET28a_AvLDO] always got lost after one cycle of cultivation even in medium supplemented with its preferred substrate L-leucine (Correia Cordeiro *et al*., 2018) an 18-fold excess of the strain was applied (ratio 18:1:1 for *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and

3Δ*sucA*[pET28a_GriE]). 3Δ*sucA*[pET28a_AvLDO] got lost in the second cycle of cultivation in M9-L-leucine Figure 23 (a), whereas in M9-L-isoleucine (b) it got lost in the first cycle. Regarding the excess of 3Δ*sucA*[pET28a_AvLDO], the strain can be detected after the first cycle (72 hours cultivation, sample t_0), but not after the second cycle (24 hours cultivation, sample t_1) in both media Figure 23 (c) and (d). In minimal medium supplemented with L-leucine Figure 23 (a) and (c) 3Δ*sucA*[pET28a_GriE] was enriched and 3Δ*sucA*[pET28a_BtIDO] at least got lost in the last to cycles of cultivation. In contrast, minimal medium supplemented with L-isoleucine Figure 23 (b) and (d) 3Δ*sucA*[pET28a_GriE] could not be detected after the second cycle of cultivation (t_1), whereas 3Δ*sucA*[pET28a_BtIDO] was enriched.

4.5 Separate cultivation of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE]

The different optical densities obtained after each cycle of cultivation for *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE] are shown in Figure 24. The three strains were cultivated separately in M9 minimal medium supplemented with either L-leucine (Figure 24 (a)) or L-isoleucine (Figure 24 (b)). The sample t_0 was cultivated for 72 hours, whereas all other samples (t_1 – t_4) were only cultivated for 24 hours. Before 25 µL of the culture were used to inoculate fresh medium, the optical density was determined. After 72 hours of cultivation all three strains showed proper growth in both media and reached ODs of ~0.6 – 1, except 3Δ*sucA*[pET28a_AvLDO] in minimal medium supplemented with L-leucine (Figure 24 (a)). Sample t_1 was inoculated with 25 µL of t_0-culture and then only cultivated for 24 hours and therefore showed considerably lower growth. In M9 medium with L-leucine 3Δ*sucA*[pET28a_AvLDO] (t_1, dark grey bar), 3Δ*sucA*[pET28a_BtIDO] (t_1, medium grey bar) almost disappeared, whereas 3Δ*sucA*[pET28a_GriE] (t_1, light grey bar) showed significantly faster growth and reached an OD_{600} of 0.4 (Figure 24 (a)). Regarding minimal medium with L-isoleucine sample t_1 of 3Δ*sucA*[pET28a_BtIDO] obtained the highest OD_{600} of approximately 0.4. However, for 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_GriE] a weak OD_{600} of 0.1 was measured for sample t_1 as well (Figure 24 (b)). 3Δ*sucA*[pET28a_GriE] maintained its OD_{600} after the second cultivation of 24 hours (t_2) and 3Δ*sucA*[pET28a_AvLDO]

could not be detected anymore in M9 with L-leucine (Figure 24 (a)). Regarding the t_2 samples cultivated in M9 medium supplemented with L-isoleucine, the OD_{600} of 3Δ*sucA*[pET28a_BtIDO] slightly increased from 0.446 to 0.472, whereas the OD_{600} of 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_GriE] were cut in half from 0.129 to 0.068 and from 0.137 to 0.065 (Figure 24 (b)). In minimal medium supplemented with L-isoleucine the cell density of 3Δ*sucA*[pET28a_BtIDO] slightly increased from t_1 to t_4 after each cycle from 0.446 (t_1) to finally 0.543 (t_4) (Figure 24 (b)). The OD_{600} of 3Δ*sucA*[pET28a_GriE] however, was maintained or slightly decreased from cycle t_1 (0.405) to t_4 (0.346) in M9 medium with L-leucine (Figure 24 (a)). Whereas in minimal medium supplemented with L-isoleucine all three dioxygenase harbouring selection strains showed detectable growth after each cycle using the assay setup (Figure 24 (b)), in M9 medium with L-leucine it was not the case for 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_BtIDO] (Figure 24 (a)).

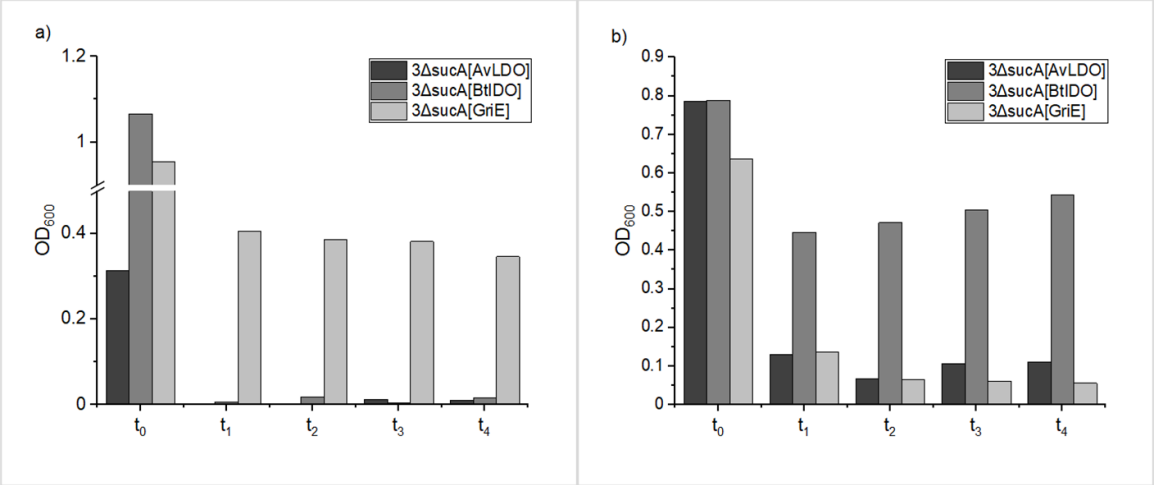

Figure 24: OD_{600} of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] (dark grey bar), 3Δ*sucA*[pET28a_BtIDO] (medium grey bar) and 3Δ*sucA*[pET28a_GriE] (light grey bar) measured after each cycle of iterative cultivation. The sample t_0 was cultivated for 72 hours, whereas all other samples (t_1 – t_4) were only cultivated for 24 hours. Cultivation was performed in M9 minimal medium with L-leucine **(a)** and L-isoleucine **(b)**. After 72 hours of cultivation 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE] obtained ODs of ~1 in M9 with L-leucine, whereas the OD of 3Δ*sucA*[pET28a_AvLDO] was considerably lower. In M9 with L-isoleucine all three strains obtained similar ODs. Switching to the 24-hour intervals of cultivation, it can be observed that 3Δ*sucA*[pET28a_GriE] grew fastest in minimal medium supplemented with L-leucine and 3Δ*sucA*[pET28a_BtIDO] in medium supplemented with L-isoleucine. Whereas in M9 with L-leucine only 3Δ*sucA*[pET28a_GriE] showed proper growth after all cycles, in L-isoleucine 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_GriE] also showed observable growth after each 24-hour cultivation interval (t_1-t_4).

4.6 Generation of a mutant library of pET28a_AvLDO

Two positions in close proximity to the substrate binding site, more specifically the isoleucine at position 70 and the phenylalanine at position 78, both pointing towards the substrate binding pocket (Figure 25), were saturated via QuikChange reaction. Therefore, a primer mix with degenerated codons was used. After digestion of the wild type template, electrocompetent *E. coli* TOP10 cells were transformed with the library. The regenerated cell suspension was split up and a part of it was plated on LB-Kan[40] plates, whereas the other part was used to inoculate a liquid culture. The number of clones formed on the plate were counted and according to the volume plated, the number of clones present in the liquid culture could be calculated.

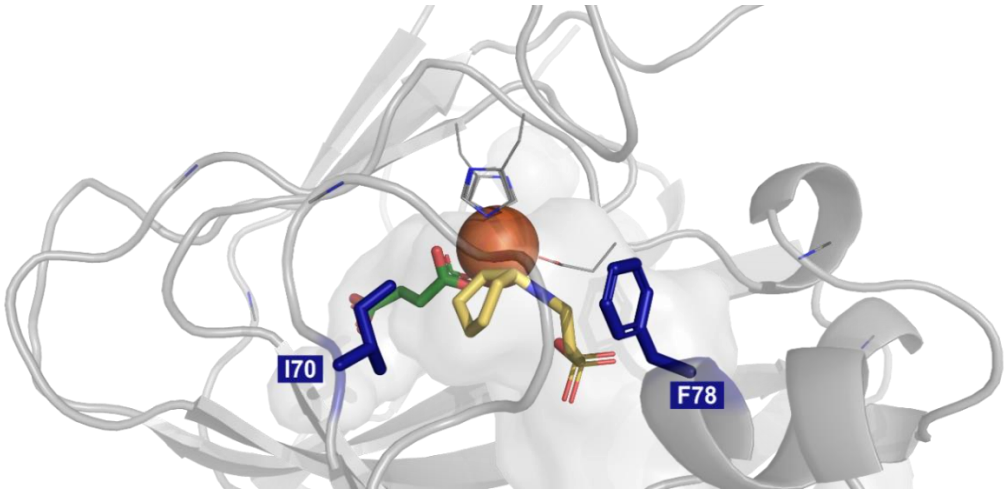

Figure 25: Crystal structure of AvLDO (unpubl.) displayed in PyMOL shows position Ile70 and Phe78 (both labeled in blue), which were targeted for simultaneous saturation mutagenesis via QuikChange reaction. Both redidues are pointing towards the substrate binding site and are therefore considered to influence the substrate scope of the enzyme. The iron centre is represented by the red sphere and succinate is shown in green. The buffer component CHES (*N*-Cyclohexyl-2-aminoethanesulfonic acid), shown in yellow, is assumed to take a position similar to what would be expected for a substrate molecule. Figure taken from (Schweiger, upubl.).

Table 15 shows how it was verified, that the library generated via QuikChange reaction contained enough variants for screening to obtain a coverage of 95 % as it is recommended for enzyme engineering (Schließmann, 2010). As 20,100 variants would be present in the liquid culture, screening enough variants for 95 % coverage is feasible. However, not only the transformation efficiency of electrocompetent *E. coli* TOP10 cells for the initial generation of the library, but also the transformation efficiency of the electrocompetent *E. coli* BL21(DE3) 3Δ*sucA* cells used as selection host is crucial for the experiment. The transformation efficiency of the selection strain was determined according to Chapter 3.5. *E. coli* BL21(DE3) 3Δ*sucA* cells showed a transformation efficiency of 26,250,000 CFU/µg (determined with the

empty vector pET28a(+) according to Chapter 3.5). As approximately 100 ng plasmid DNA of the library are used for electroporation, which would result in 2,625,000 colony forming units the transformation efficiency is high enough to screen a library with 60 different variants.

Table 15: Verification of number of variants present in library generated. Electrocompetent *E. coli* TOP10 cells were used for transformation.

Variants generated via QuikChange	Volume plated [µL]	Colonies	Volume for inoculation of liquid culture [µL]	Variants present in liquid culture	Variants to be screened (95 % coverage)
60	10	300	670	20,100	180

Table 16 shows the sequencing results of the 10 clones picked to verify the success of saturation mutagenesis. In general, 35 variants with different amino acid combinations and 60 variants with different codon combinations can be created with the designed mutagenesis primer. By sequencing 10 clones, both 8 different amino acid combinations and 8 different variants were obtained. As ten clones were picked randomly, the success of the mutagenesis method can be verified with this result. In detail, five out of twelve mutations were obtained in position 70 and four out of five mutations in position 78. Although the wild type was observed twice (clone 2 and clone 10), there is no doubt the *Dpn*I digest did not work, because the degenerated primer mix used for the QuikChange reaction also contained primers exactly generating the wild type.

Table 16: Testing 10 clones to verify the success of the QuikChange reaction.

Clone	Mutations	
1	I70A (GCA)	F78I (ATT)
2	I70I (ATT)	F78F (TTT)
3	I70A (GCA)	F78L (CTT)
4	I70F (TTT)	F78F (GTT)
5	I70S (TCT)	F78V (GTT)
6	I70A (GCA)	F78I (ATT)
7	I70F (TTT)	F78V (GTT)
8	I70T (ACT)	F78V (GTT)
9	I70T (ACT)	F78F (TTT)
10	I70I (ATT)	F78F (TTT)

4.7 Selection of AvLDO variants

Iterative cultivation (Figure 15) of *E. coli* BL21(DE3) 3Δ*sucA* cells harbouring plasmids with different AvLDO variants generated via a QuikChange reaction was performed in minimal medium with L-isoleucine in order to select strains bearing enzyme variants with improved substrate utilisation. Before a small volume of liquid culture was used to inoculate fresh minimal medium after each cycle of cultivation, the optical density of the culture was determined. In Figure 26 the optical density of the triple-knockout selection strain *E. coli* BL21(DE3) 3Δ*sucA* transformed with the AvLDO library (black bar) is compared to the wild type 3Δ*sucA*[pET28a_AvLDO] (striped bar). After the initial 72 hours of cultivation, which are required for adaptation to the conditions in minimal medium, 3Δ*sucA*[pET28a_AvLDO library] showed an almost 2-fold higher optical density (1.53) than the wild type 3Δ*sucA*[pET28a_AvLDO] (0.79). For the measurements conducted after the first 24-hour cultivation interval the differences were even higher as 3Δ*sucA*[pET28a_AvLDO library] showed half the OD of t_0 (0.87), whereas the OD of 3Δ*sucA*[pET28a_AvLDO] showed a 6-fold decrease to 0.129. After the following cycles t_2 to t_4 a slight decrease in the OD can be observed for 3Δ*sucA*[pET28a_AvLDO library] from cycle to cycle with a final OD of 0.393 after t_4. On the contrary, the wild type 3Δ*sucA*[pET28a_AvLDO] declined after t_2 (0.068), but then slightly went up again for the last two cycles reaching a final OD of 0.111.

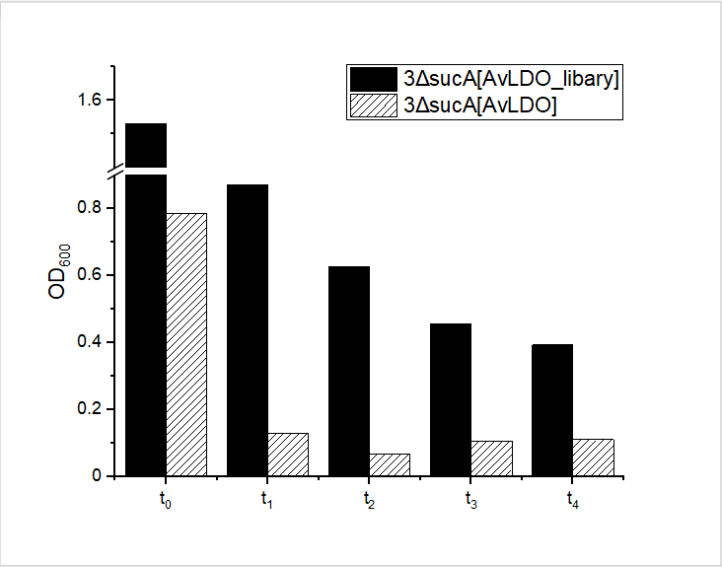

Figure 26: OD_{600} of the AvLDO library harbouring *E. coli* BL21(DE3) 3Δ*sucA* cells (black bar) reported after each cycle of cultivation compared to the wild type *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] (striped bar). After 72 hours of cultivation (t_0) 3Δ*sucA*[pET28a_AvLDO library] (black bar) showed a 2-fold higher OD_{600} than the wild type 3Δ*sucA*[pET28a_AvLDO] (striped bar).

After each cycle of cultivation, part of the enrichment culture was plated on M9-agar antibiotic plates for the selection of improved variants. In Figure 27 one of the plates is shown for illustration. Larger sized clones, marked with a black arrow, were picked, and preserved for further analysis.

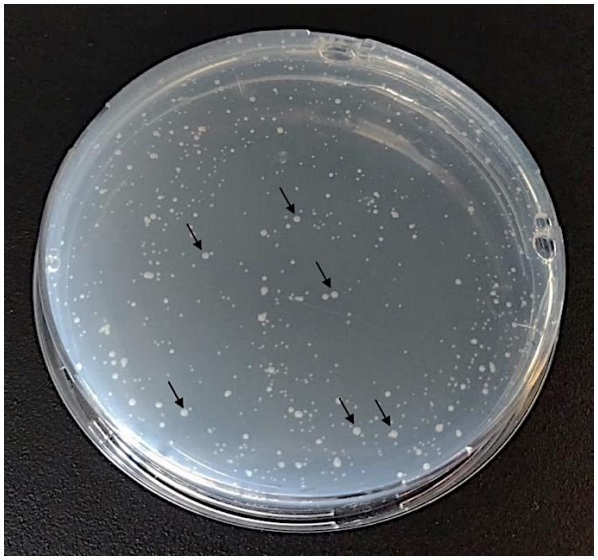

Figure 27: Enrichment culture of the AvLDO library plated on M9-agar antibiotic plates after five cycles of cultivation. Larger sized clones were picked for further analysis (black arrows).

4.8 Sequencing of selected variants

Table 17 shows the sequencing results of all 14 clones picked from M9-agar plates. Eight out of 14 mutants (3Δ*sucA*_varA, 3Δ*sucA*_varB, 3Δ*sucA*_varI - 3Δ*sucA*_varM) have the same genotype including an amino acid exchange from L-isoleucine to L-threonine at position 70 and no exchange at position 78. In variant C both positions 70 and 78 are exchanged by L-leucine and in variant D position 70 is also exchanged by L-leucine, whereas position 78 remains unchanged. A sequence alignment of these variants can be found in the Appendix (Chapter 8.3, Figure 41). The sequencing result of variant E and F could not be aligned to the AvLDO template and no sequencing result at all was obtained for variant G by the use of the in-house sequencing primers specific for the T7 promoter and terminator region.

Table 17: Genotype of variants used for re-screening.

Genotype of improved variants			
Variant	**Sequencing possible**	**Position 70 (ATT)**	**Position 78 (TTT)**
varA	yes	I70T (ACT)	F78F
varB	yes	I70T (ACT)	F78F
varC	yes	I70L (TTA)	F78L (CTT)
varD	yes	I70L (TTA)	F78F
varE	not alignable		
varF	not alignable		
varG	no sequencing result		
varH	yes	I70T (ACT)	F78I (ATT)
varI	yes	I70T (ACT)	F78F
varJ	yes	I70T (ACT)	F78F
varK	yes	I70T (ACT)	F78F
varL	yes	I70T (ACT)	F78F
varM	yes	I70T (ACT)	F78F
varN	yes	I70T (ACT)	F78F

4.9 Re-screening of selected variants

4.9.1 Growth curves in test tubes

Figure 28 shows the re-screening of selected variants in test tubes. Variants were re-transformed in 3Δ*sucA* cells to exclude false positives caused by mutations in the knockout strain itself. The growth of eight variants, *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a(+)] was monitored over 150 hours. Strains were grown in M9 minimal medium supplemented with L-isoleucine, kanamycin and IPTG. The exponential phase of the wild type 3Δ*sucA*[pET28a_AvLDO] started after 75 hours of cultivation. A final OD of 1 was obtained after 142 hours showing rather linear than exponential growth with a specific growth rate of only 0.025 h^{-1}.

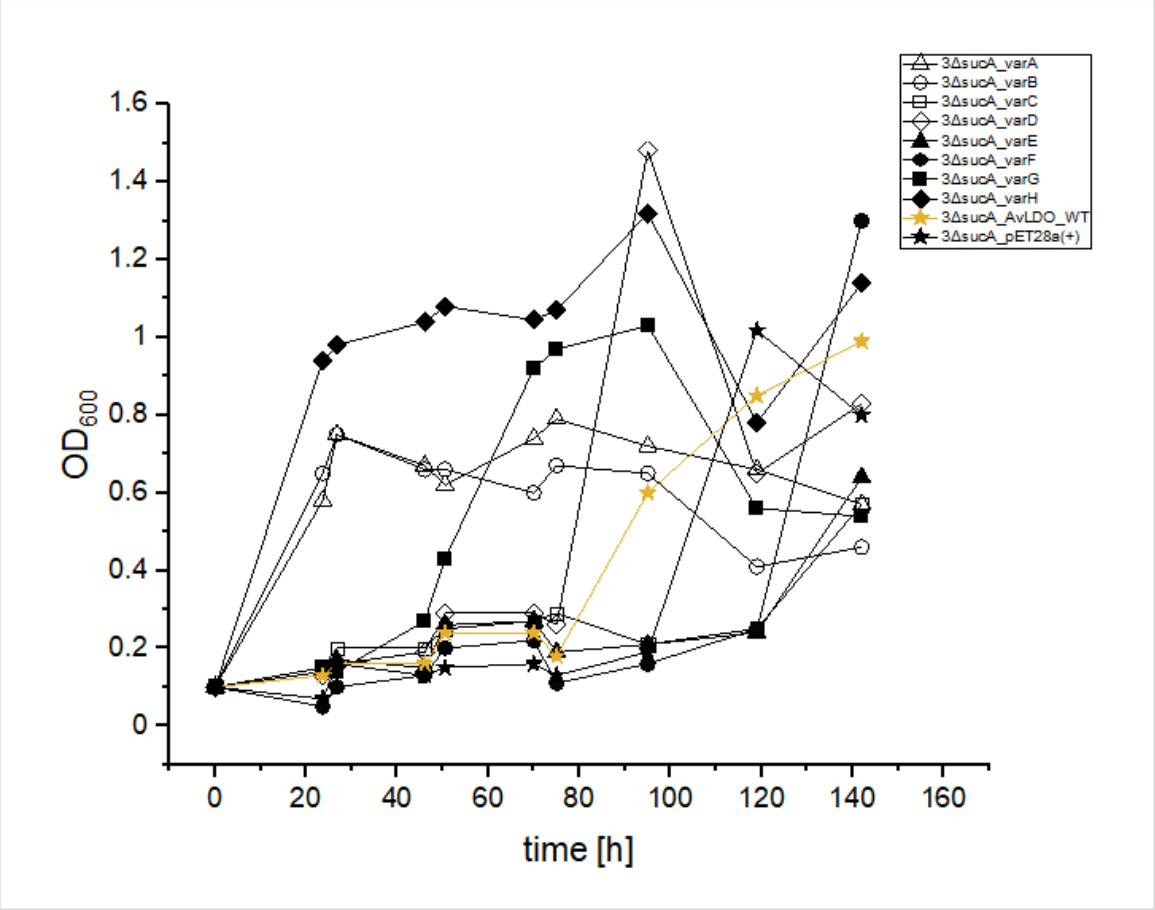

Figure 28: Growth curves of selected clones to detect false positive variants. *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] (black star) was used as a negative control. The wild type *E. coli* BL21(DE3)[pET28a_AvLDO] (yellow star) was used for comparison to detect strains bearing faster growing variants. 3Δ*sucA*_varH (black hash), 3Δ*sucA*_varA (white triangle), 3Δ*sucA*_varB (white circle), 3Δ*sucA*_varG (black square), and 3Δ*sucA*_varD (white hash) started its exponential growth before the wild type (yellow star) and are therefore considered to be improved variants. 3Δ*sucA*_varC (white square), 3Δ*sucA*_varE (black triangle) and 3Δ*sucA*_varF (black circle) starting their growth after the wild type were assumed to be false positive clones.

Five out of eight mutants showed an earlier exponential growth than the wild type including 3Δ*sucA*_varA, 3Δ*sucA*_varB, 3Δ*sucA*_varD, 3Δ*sucA*_varG and 3Δ*sucA*_varH. 3Δ*sucA*_varH was the fastest and showed its exponential growth phase during the first 24 hours after inoculation, followed by 3Δ*sucA*_varA and 3Δ*sucA*_varB (same genotype). Although 3Δ*sucA*_varA and 3Δ*sucA*_varB also started exponential growth in the first 24 hours, 3Δ*sucA*_varH showed a considerably higher OD (0.98) after 24 hours compared to 3Δ*sucA*_varA (0.58) and 3Δ*sucA*_varB (0.65) resulting in a higher specific growth rate of 0.086 h^{-1}. In contrast, 3Δ*sucA*_varA and 3Δ*sucA*_varB only showed a specific growth rate of 0.076 h^{-1}. Afterwards, 3Δ*sucA*_varG started its exponential growth after approximately 40 hours reaching a final OD of 1. The fourth mutant, which started its exponential growth before the wild type was 3Δ*sucA*_varD. It started exponential

growth after 75 hours and obtained a final OD of 1.48 after 95 hours being the mutant with the highest specific growth rate (0.087 h^{-1}) of all strains re-screened. After four days of lag-phase the negative control 3Δ*sucA*[pET28a(+)] started to grow as well with a specific growth rate of 0.070 h-1 reaching a maximum OD of 1 after 120 hours. As 3Δ*sucA*_varC, 3Δ*sucA*_varE and 3Δ*sucA*_varF started to grow 48 hours after the wild type 3Δ*sucA*[pET28a_AvLDO] and 24 hours after the negative control 3Δ*sucA*[pET28a(+)] they were considered to be false positive clones and cultivation was stopped.

4.9.2 Iterative cultivation in test tubes

The bar chart below (Figure 29) shows the re-screening of eight variants, *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a(+)] in test tubes. Minimal medium supplemented with L-isoleucine was inoculated to an OD_{600} of 0.1 to prepare the first main culture (t_0) and grown for 72 hours to enable adaption to minimal conditions. Whereas the wild type 3Δ*sucA*[pET28a_AvLDO] obtained an OD of 0.213 after 72 hours of cultivation, two mutants 3Δ*sucA*_varG and 3Δ*sucA*_varH showed 5-fold higher optical densities. Further, 3Δ*sucA*_varA and 3Δ*sucA*_varB showed more than 2-fold higher ODs as the wild type. After 72 hours, the negative control 3Δ*sucA*[pET28a(+)] only had an OD of 0.1, which was the OD at the beginning of cultivation. 3Δ*sucA*_varF showed even lower growth and only obtained an OD of 0.072 and almost disappeared after the following cultivation cycles (t_1 and t_2) indicating the selection of a false positive variant. After the first 24-hour cultivation (t_1) the OD_{600} of all strains significantly dropped with maximal ODs of 0.17 (3Δ*sucA*_varG). After the following cycles, the ODs remained low, but 3Δ*sucA*_varG and 3Δ*sucA*_varH always showed considerably higher ODs than the wild type. Whereas the ODs of 3Δ*sucA*_varA and 3Δ*sucA*_varB slightly declined from cycle to cycle or at least remained constant, the ODs of 3Δ*sucA*_varC, 3Δ*sucA*_varD, 3Δ*sucA*_varE, the wild type 3Δ*sucA*[pET28a_AvLDO] and the negative control 3Δ*sucA*[pET28a(+)] slightly increased.

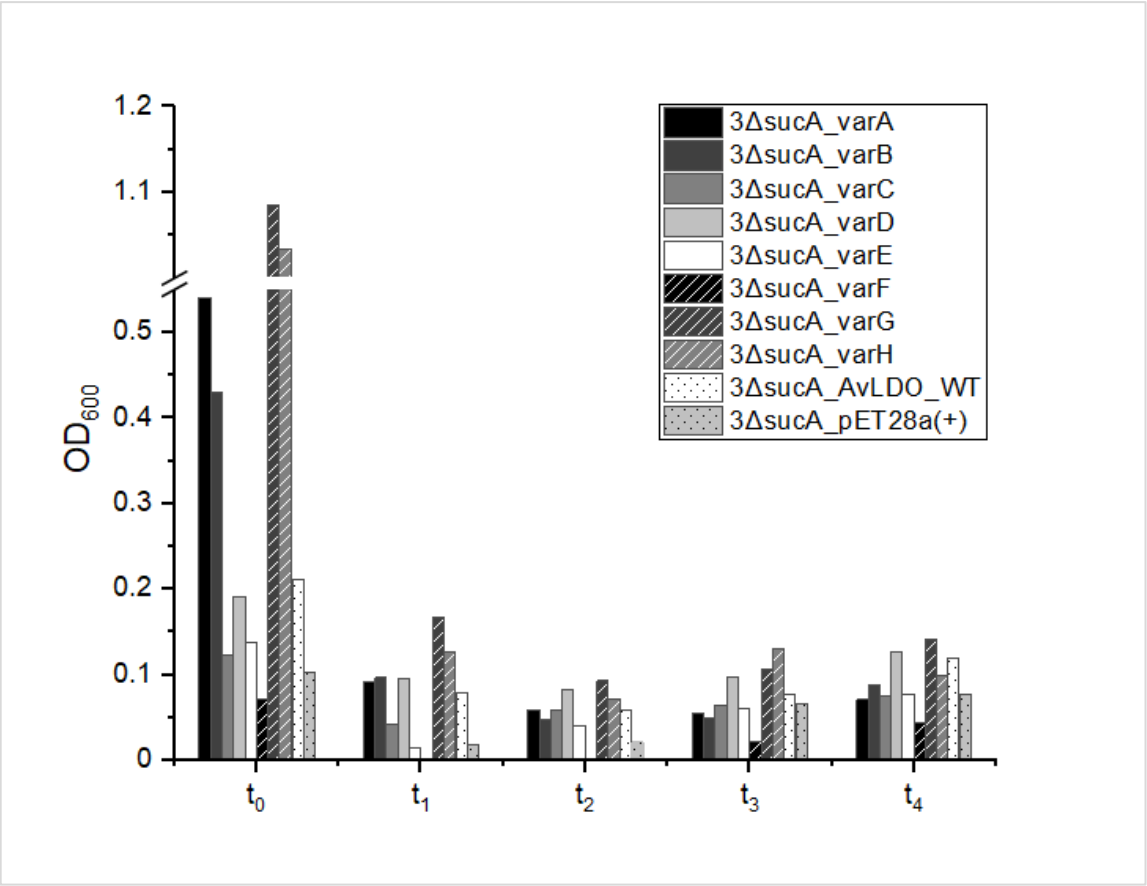

Figure 29: Iterative cultivation of selected clones in test tubes to detect false positive variants. *E. coli* BL21(DE3) 3Δ*sucA*[pET28a(+)] (grey dotted bar) was used as a negative control. The wild type *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] (white dotted bar) was used for comparison to detect faster growing variants. The first main culture (t_0) was inoculated to an OD_{600} of 0.1 and grown for 72 hours. 3Δ*sucA*_varA (black bar), 3Δ*sucA*_varB (dark grey bar), 3Δ*sucA*_varG (medium grey striped bar) and 3Δ*sucA*_varH (light grey striped bar) showed higher optical densities than the wild type 3Δ*sucA*[pET28a_AvLDO] (white dotted bar) after the adaption phase of 72 hours. Then, fresh medium was inoculated with the first main culture and cultivated for 24 hours. After the first 24-hour interval 3Δ*sucA*_varD (light grey bar) additionally reached a higher OD_{600} than the wild type. Poorest growth was observed by 3Δ*sucA*_varF (black striped bar), which obtained lower ODs than the negative control 3Δ*sucA*[pET28a(+)] (grey dotted bar).

4.9.3 Evaluation of re-screening results

Table 18 summarises the re-screening results obtained from the growth curves and the iterative cultivation in test tubes. All variants showing slower growth than the wild type strain 3Δ*sucA*[pET28a_AvLDO] are assumed to be false positive. Variants showing faster growth than the wild type were expected to be higher active towards L-isoleucine. Both re-screening approaches led to the identification of the same positive and false positive variants. However, for varG, which was assumed to be positive, no sequencing result was obtained. Further, for varE and varF, both negative, sequencing results could not be aligned to the template.

Table 18: Evaluation of the re-screening results. Positive variants showed faster growth than the AvLDO wild type harbouring *E. coli* BL21(DE3) 3Δ*sucA* strain, whereas negative variants displayed slower growth.

Variant	Genotype	Growth curves	Iterative cultivation
varA	I70T (ACT); F78F	positive	positive
varB	I70T (ACT); F78F	positive	positive
varC	I70L (TTA); F78L (CTT)	negative	negative
varD	I70L (TTA); F78F	positive	positive
varE	not alignable	negative	negative
varF	not alignable	negative	negative
varG	no sequencing result	positive	positive
varH	I70T (ACT); F78I (ATT)	positive	positive

4.10 Expression of selected variants under assay conditions

Expression of all variants showing faster growth in the re-screening (Chapter 4.9) was performed in minimal medium supplemented with L-isoleucine and samples were analysed via SDS-gel. For each variant the whole cell profile, but also a separation of soluble and insoluble fraction was analysed separately. In Figure 30, the expression of the AvLDO wild type enzyme and variant B (I70T, F70F), which was obtained several times by random selection of 14 larger clones, was investigated. Regarding the whole cell sample, the wild type enzyme as well as variant B were formed under assay conditions in minimal medium supplemented with L-isoleucine. However, the wild type enzyme was detected in slightly larger amounts. In the soluble fraction both enzymes were present in similar intensities, but more background can be observed for variant B. In the insoluble fraction it is difficult to compare as no clearly defined bands can be observed. However, both wild type enzyme and variant B seem to be present in the insoluble fraction as well.

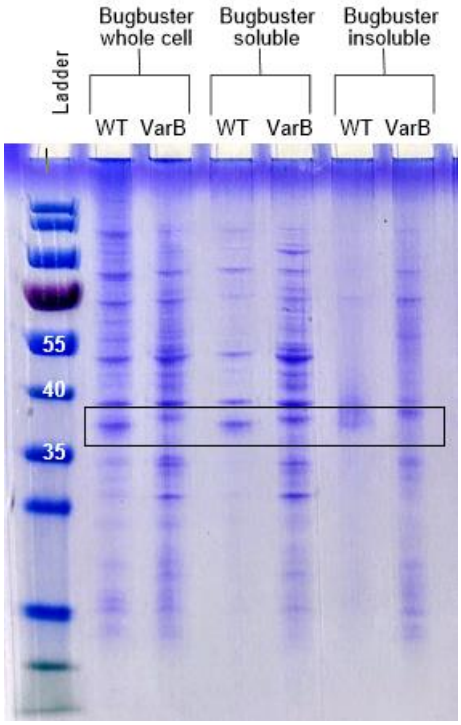

Figure 30: Expression of the AvLDO wild type (33 kDa) and variant B (I70T, F78F) by *E. coli* BL21(DE3) 3Δ*sucA* analysed by SDS-PAGE 24 hours after induction in minimal medium supplemented with L-isoleucine. The whole cell approach and a separation of soluble and insoluble fraction was investigated. Both AvLDO wild type and variant B are present in the whole cell sample as well as in the soluble and insoluble fraction. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

Besides the wild type enzyme AvLDO and variant B, the expression of further variants displaying faster growth in the re-screening (Chapter 4.9) was analysed (Figure 31). Regarding the whole cell samples, a clearly defined band can only be observed for the wild type enzyme. Variant B and variant G indicate small amounts of the respective enzyme, whereas no enzyme formation can be observed for variant D and variant H. For the separation of the soluble and insoluble fraction it looks rather similar as only for the wild type and variant D clearly defined bands with extremely low intensity can be seen in the respective height. For variant B, a faint band that is slightly higher than the bands of the WT and variant D is displayed on the gel. In contrast, for variant G and variant H no protein formation can be observed on the gel separating soluble and insoluble fractions.

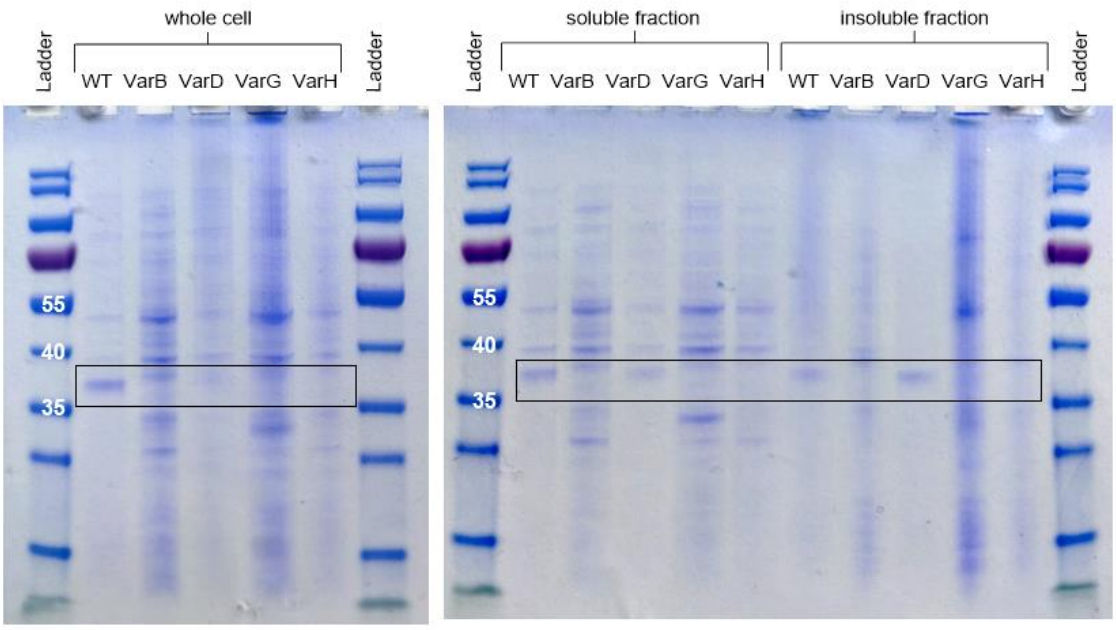

Figure 31: Expression of AvLDO (33 kDa) variants by *E. coli* BL21(DE3) 3Δ*sucA* analysed by SDS-PAGE 24 hours after induction in minimal medium supplemented with L-isoleucine. The whole cell approach and a separation of soluble and insoluble fraction analysed for the wild type and four variants obtained from the selection assay. Only the wild type enzyme can be detected in the whole cell sample and in both soluble and insoluble fraction. The expression of the different enzyme variants is even weaker than the expression of the wild type enzyme. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

4.11 Verification of L-isoleucine conversion of selected variants via thin layer chromatography

After expression under optimised conditions (Correia Cordeiro *et al*., 2018; Hanreich, 2019) and cell disruption via sonication, the CFE was used to set up biocatalytic reactions with the AvLDO wild type enzyme and the three positive variants detected in the re-screening. The aim was to verify the conversion of L-isoleucine to L-hydroxyisoleucine by the selected variants. The wild type enzyme was also tested with L-leucine. As the conversion of L-leucine to L-hydroxyleucine was observed by TLC for the AvLDO wild type enzyme using similar conditions (Hanreich, 2019), the reaction served as a positive control.

4.11.1 SDS-Gel

Samples of all expression cultures were taken before harvest to analyse the whole cell profile of the wild type and each AvLDO variant in order to prove the presence of the respective enzyme 18 – 22 hours after induction with IPTG. Therefore, Bugbuster® was used for cell disruption. Additionally, the CFE obtained after sonication was analysed via SDS-gel. The objective was to check if any soluble

protein is present in the CFE after cell disruption. All whole cell samples showed considerable amounts of AvLDO (33 kDa), whereas in the CFE no enzyme was detected at the respective height for all three variants and the wild type (Figure 32). In general, almost no protein was observed in the CFE samples, neither endogenous *E. coli* proteins, although a defined amount of protein was applied according to the BCA assay (Chapter 3.16.3). Considering the whole cell samples, WT and variant D show slightly larger amounts of enzyme than variant B and variant H.

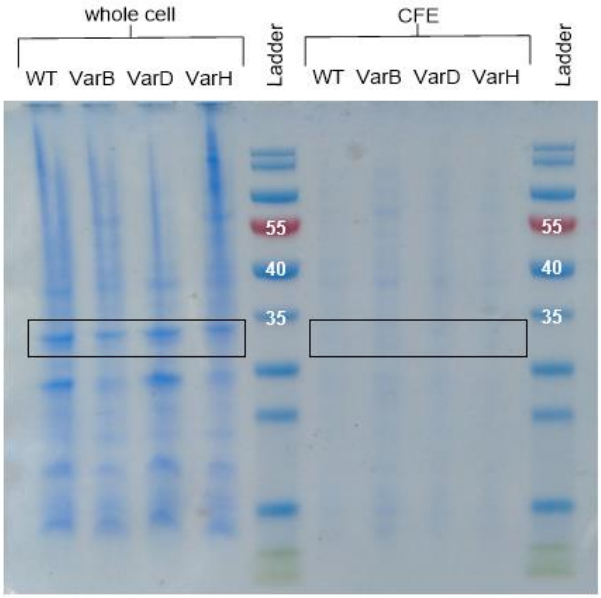

Figure 32: Expression of AvLDO (33 kDa) variants by *E. coli* BL21(DE3) in TB medium analysed by SDS-PAGE. Samples were taken after expression overnight to analyse the whole cell profile. Further, the CFE was analysed after cell disruption via sonication. AvLDO was only detected in the whole cell samples obtained from the chemical disruption with Bugbuster®, but not in the CFE samples. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific).

4.11.2 TLC

As the objective was to verify if the variants selected are capable of converting L-isoleucine to L-hydroxyisoleucine rather than to compare the extent of conversion of different variants, biocatalytic reactions were performed using the maximum volume of CFE possible for every reaction (320 µL CFE, total reaction volume of 500 µL). Nonetheless, the protein concentration was determined for each variant via BCA assay (Chapter 3.16.3). Concentrations are shown in Table 19. R_f-values of all amino acids and standards used in this study are shown in Table 20.

Table 19: Protein concentration of the CFE, absolute amount of total protein present in the biocatalytic reaction and evaluation of activity towards L-isoleucine according to TLC results.

Enzyme	Total protein concentration [mg/mL]	Absolute amount of total protein present in reaction [mg]	Detectable activity towards L-isoleucine
AvLDO WT	11.31	3.62	No
VarB	9.10	2.91	No
VarD	8.40	2.69	No
VarH	8.61	2.76	No

Table 20: R_f-values of all amino acids and standards used in this study.

Compound	**R_f-value**
L-leucine	0.46
L-isoleucine	0.40
L-hydroxyisoleucine	0.28

Figure 33 shows the activity assay with the AvLDO wild type using L-leucine as a substrate. The reaction should serve as a positive control. However, no product formation was observed in any samples taken from 0 – 92 hours. Whereas the blank reaction only shows the substrate and one spot above the substrate, which was identified to be ascorbate (Appendix, Figure 42), unspecific by-products were detected in the reaction samples applying the CFE.

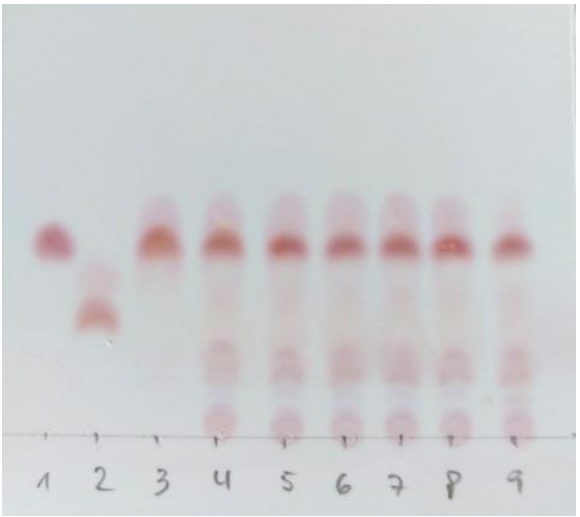

Figure 33: Biocatalytic reaction with the AvLDO wild type enzyme using L-leucine as a substrate serving as a positive control. Reactions were performed with the CFE. L-leucine and hydroxyisoleucine were used as a standard. Final reaction conditions: 5 mM substrate, 10 mM α-ketoglutarate, 10 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, 320 µL CFE.(11.32 mg/mL) in 500 µL total volume at 25 °C. [1] L-leucine; [2] L-hydroxyisoleucine; [3] Blank t_1 [4] t_0; [5] t_1; [6] t_2; [7] t_{18}; [8] t_{24}; [9] t_{92}.

Figure 34 shows the activity assay with the AvLDO wild type [A], variant B [B], variant D [C] and variant H [D] using L-isoleucine as a substrate. Chromatograms of all enzyme variants as well as the wild type do not show any formation of the desired product L-hydroxyisoleucine. However, extensive by-product formation was observed in all samples except the blank reaction indicating undesired activities of unspecific *E. coli* protein in the CFE. Further, it is assumed that a large extent of proteins in the CFE was inactive, as almost no protein can be found in the soluble fraction analysed via SDS-PAGE (Figure 32).

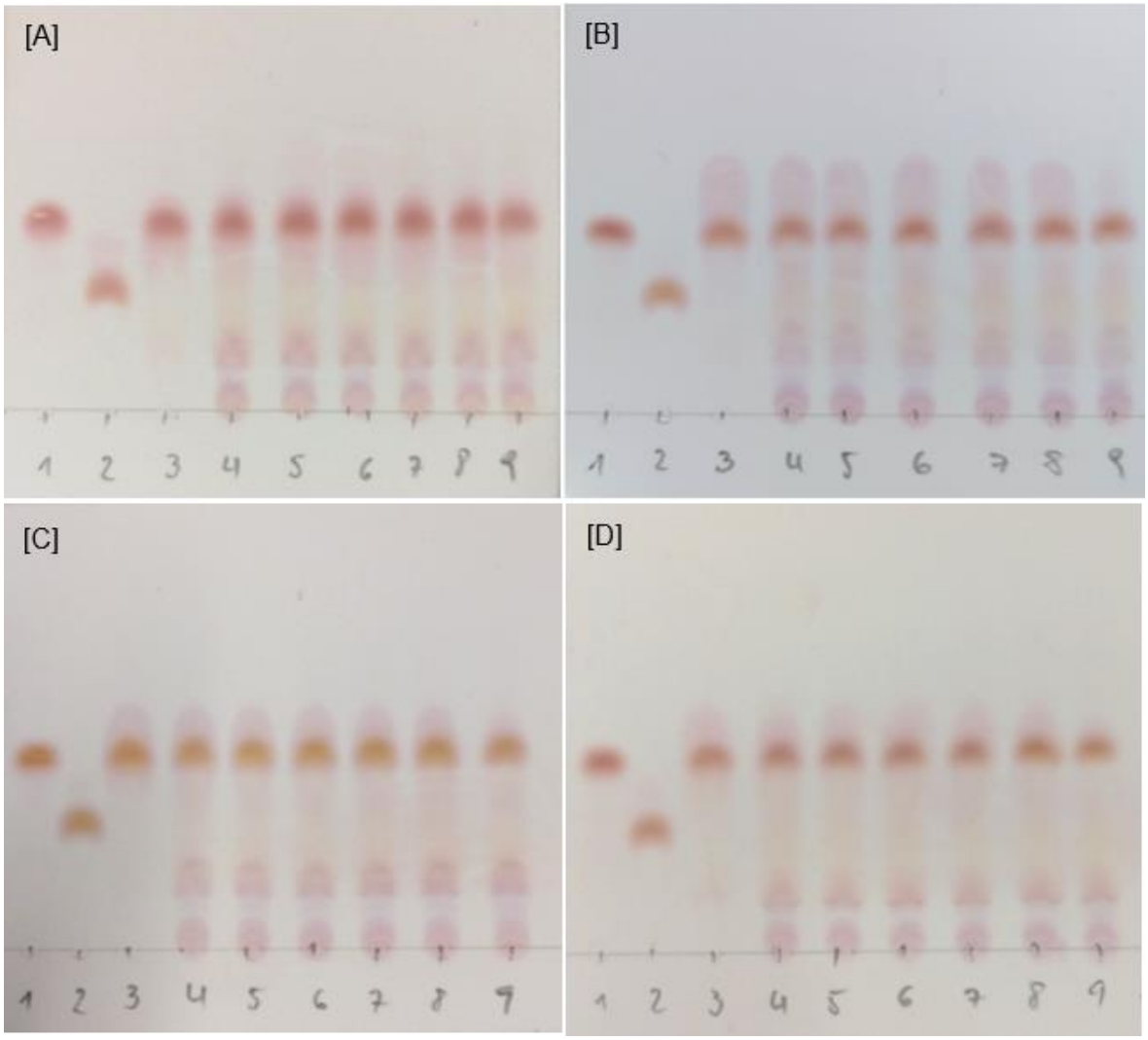

Figure 34: Biocatalytic reactions with the AvLDO wild type enzyme and three improved variants using L-isoleucine as a substrate. Reactions were performed with the CFE. L-isoleucine and hydroxyisoleucine were used as a standard. Final reaction conditions: 5 mM substrate, 10 mM α-ketoglutarate, 10 mM ascorbate, 0.5 mM $FeSO_4x7H_2O$, CFE 320 µL in 500 µL total volume at 25 °C. **[A]** Wild type: 11.32 mg/mL **[B]** Variant B (I70T, F78F): 9.10 mg/mL **[C]** Variant D (I70L, F78F): 8.40 mg/mL **[D]** Variant H (I70T, F78I): 8.61 mg/mL. [1] L-isoleucine; [2] L-hydroxyisoleucine; [3] Blank t_1 [4] t_0; [5] t_1; [6] t_2; [7] t_{18}; [8] t_{24}; [9] t_{92}.

5 Discussion

CH-functionalisation of small molecules as performed by Fe(II)/α-ketoglutarate dependent oxygenases is a valuable transformation for the biosynthesis of noncanonical amino acids. It is of industrial importance for the generation of pure hydroxy amino acids serving as precious building blocks for the pharmaceutical industry (Hüttel, 2013; Peters and Buller, 2019; Zwick and Renata, 2018a).

In the first part of this thesis, growth and expression studies in minimal medium were performed with three different Fe(II)/α-ketoglutarate dependent dioxygenases employing the triple knockout selection strain *E. coli* BL21(DE3) 3Δ*sucA*. Moreover, a mixed iterative cultivation of three dioxygenases followed by qualitative detection via multiplex PCR was performed to study their enrichment.

Valuable information gained from these preliminary experiments should assist in the implementation of a growth-based selection assay for the detection of improved dioxygenase variants. The assay was applied to select the mutant library of the L-leucine dioxygenase from *A. variablilis* for improved variants with an extended substrate scope.

5.1 Growth studies

Growth of the triple-knockout mutant *E. coli* BL21(DE3) 3Δ*sucA* can be restored by re-cyclisation of the impaired TCA via L-amino acid hydroxylation (Theodosiou *et al*., 2017). This principle was applied in this study to select active Fe(II)/α-ketoglutarate dependent dioxygenases. Before the setup of a growth-based assay for the selection of improved enzyme variants could be implemented, several studies on the three dioxygenases AvLDO, BtIDO and GriE were performed. The aim was to obtain valuable information about their growth behaviour and their course of expression that should support the implementation of the assay.

The growth behaviour of AvLDO, BtIDO and GriE was already investigated in previous studies by cultivation of dioxygenase harbouring 3Δ*sucA* strains in minimal medium supplemented with L-leucine or L-isoleucine or without any substrate and compared to growth in LB (Hanreich, 2019). As the backbone of the BtIDO-harbouring plasmid was exchanged from pET22b to pET28a to enable a mixed cultivation of all three plasmids, growth studies were performed with the new construct. As the ampicillin resistance gene of pET22b(+) underlies a different

mechanism than the kanamycin resistance gene of pET28a(+), the experiment was repeated to figure out if this change affects growth behaviour. Further, it is known that ampicillin is hydrolysed by β-lactamases secreted by *E. coli* (Bajaj *et al*., 2016). Stability tests on several antibiotics conducted by Peteranderl *et al.* under elevated temperatures (50 °C) for 120 hours, reported no observable decay of kanamycin, whereas ampicillin was amongst the antibiotics exhibiting lowest stability (Peteranderl *et al*., 1990). The knockout strain used in this study shows an extended lag-phase when cultivated in minimal medium ranging from 40 to 120 hours depending on the substrate supplemented (Hanreich, 2019). The slow growth can be explained by the scarcity of succinyl-CoA, important in cell wall formation and the absence of the glyoxylate shunt (Theodosiou *et al*., 2017). Hence, the change to a plasmid expressing a more stable antibiotic resistance gene was expected to be beneficial. Growth curves and final optical densities obtained with *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO] (Figure 16) were similar to those of 3Δ*sucA*[pET22b_BtIDO] reported in previous studies (Hanreich, 2019). However, it is notable that exponential growth in M9 medium with L-leucine started approximately 20 hours earlier in this study. In contrast, for cultivation in L-isoleucine containing minimal medium, only minor differences were observed. Of course, the different lag-phases and specific growth rates of cultivation in minimal medium with L-leucine (~80 hours, 0.138 h^{-1}) and L-isoleucine (~40 hours, 0.166 h^{-1}) in this study can be explained by the 2.3-fold higher specific activity towards the latter substrate (Hibi *et al*., 2011) combined with the high expression of BtIDO in isoleucine-supplemented medium at an earlier time than in leucine-supplemented medium (Chapter 4.1.2). Regarding the different lag-phases for 3Δ*sucA*[pET28a_BtIDO] (~80 hours, this study) and 3Δ*sucA*[pET22b_BtIDO] (~120 hours, Hanreich, 2019) in minimal medium with L-leucine, an explanation could be that the different antibiotic resistance mechanism affected the expression of BtIDO and in turn the re-cyclisation of the TCA and the growth behaviour was altered. However, expression studies of 3Δ*sucA*[pET22b_BtIDO] would need to be conducted to verify this assumption. Further, it must be mentioned that Hanreich (Hanreich, 2019) used a slightly different procedure for inoculation of the main cultures in M9 medium. Hence, it is unclear why only the growth behaviour in minimal medium with L-leucine, but not in minimal medium with L-isoleucine would be affected by these differences. Nevertheless, due to the benefits of using the pET28 vector, it was

decided to continue this study with the newly created construct pET28a_BtIDO and focus on library generation and assay development.

Growth studies were further performed with *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] as the backbone of pET28a_AvLDO (Ruhr Universität Bochum) was exchanged by the backbone of pET28a_GriE (GenScript Biotech B.V. (Netherlands)) in this study due to point mutations in the former backbone. As significantly higher transformation efficiencies were observed for *E. coli* BL21(DE3) 3Δ*sucA* using the new construct, it was assumed that the re-cloning would also have a beneficial effect on cell growth. However, growth curves of 3Δ*sucA*[pET28a_AvLDO] (Figure 17) were comparable to the results of Hanreich (Hanreich, 2019) with lag-phases of 120 hours for cultivation in minimal medium supplemented with L-leucine and 140 hours for minimal medium supplemented with L-isoleucine. Whereas finial ODs of 3Δ*sucA*[pET28a_AvLDO] obtained for cultivation in minimal medium with L-leucine were quite similar in both studies (~4), in this study an almost 3-fold higher final OD of 4.3 was achieved in minimal medium with L-isoleucine. Further, the negative control 3Δ*sucA*[pET28a(+)] showed a shorter lag-phase and stronger growth in this study compared to Hanreich (Hanreich, 2019).

As four pathways lead to succinate synthesis in *E. coli* (Samsonova *et al*., 2005) and in the applied selection strain 3Δ*sucA* only two of them are knocked out (aerobic from α-KG and glyoxylate pathway), there remain further possibilities for the re-cyclisation of the TCA.

Weak growth of the *E. coli* BL21(DE3) 3Δ*sucA* selection strain harbouring only the empty vector pET28a(+) can be explained by the formation of succinate via other side pathways as the aminobutyrate pathway, where L-glutamate leads to the formation of succinate (Samsonova *et al*., 2005). Li *et al.* reported the excretion of small amounts of L-glutamate by *E. coli* BW25113 Δ*sucA* cells cultivated in minimal medium with glucose (10 g/L) after 20 hours of batch cultivation (Li *et al*., 2006). The minimal medium used in this study also contained 5 g/L glucose and cultivation times were far beyond 20 hours.

The anaerobic pathway of *E. coli* producing mixed fermentation products also includes succinate (Thakker *et al*., 2012) and could also lead to growth restoration. However, it is unclear why the negative control only showed high growth in minimal

medium with L-isoleucine, but not in medium with L-leucine or without any substrate. As the respective sterile control remained sterile during the whole course of cultivation, the strong growth of the negative control only in minimal medium with L-isoleucine remains unexplained. However, the longer lag-phase of 3Δ*sucA*[pET28a_AvLDO] compared to the negative control 3Δ*sucA*[pET28a(+)] in M9 medium with L-isoleucine can be explained by the additional metabolic burden caused by the enzyme overexpression as carbon is withdrawn from the TCA cycle for AvLDO formation. Similar observations were reported in previous studies for P4H catalysis by Theodosiou *et al.* (Theodosiou *et al*., 2017).

The extensive growth of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] in minimal medium without any substrate could be explained by sufficient amounts of endogenous substrate hydroxylation. This phenomenon was previously reported by Theodosiou *et al.* (Theodosiou *et al*., 2017) where the endogenous proline hydroxylation rate of the enzyme was sufficient to restore the TCA-cycle. The finding of Li *et al.* (Li *et al*., 2014), who reported a significantly upregulated amino acid biosynthesis in *E. coli* cells grown in minimal medium compared to rich medium emphasises this observation. Theodosiou *et al.* obtained a 1.4-fold lower specific growth rate using endogenous substrates (0.10 h^{-1}) than with external substrate addition (0.14 h^{-1}) (Theodosiou *et al*., 2017). In this study, the specific growth rate was even 1.9-fold lower (0.063 h^{-1}) than in minimal medium with L-leucine (0.117 h^{-1}) and 1.4-fold lower than in minimal medium with L-isoleucine (0.087 h^{-}1). However, 3Δ*sucA*[pET28a_AvLDO] started exponential growth already after 90 hours of cultivation in minimal medium without any substrate compared to 120 hours in minimal medium with L-leucine and 140 hours with L-isoleucine. Interestingly, *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO] (Figure 16 (b)) started its exponential growth also after about 90 hours in medium without any substrate, but showing a slightly lower specific growth rate of 0.035 h^{-1}. However, with a specific growth rate of 0.166 h^{-1} for 3Δ*sucA*[pET28a_BtIDO] in presence of the preferred substrate L-isoleucine, growth was even 4.7-fold lower in medium without substrate. As L-isoleucine should not be converted by AvLDO according to Cordeiro *et al.* (Correia Cordeiro *et al*., 2018) a lower specific growth rate would be assumed in M9 medium with L-isoleucine than in M9 medium with L-leucine for 3Δ*sucA*[pET28a_AvLDO]. As expected, growth started about 20 hours later and the

specific growth rate of 0.087 h^{-1} was 1.3-fold lower than in minimal medium with L-leucine, but the final OD of 4.28 was higher than in L-leucine (3.88).

Although cultivation was performed in defined minimal medium and sterile controls were sampled regularly to monitor sterility and exclude handling related contamination, cultivation over one week involves a certain risk of contamination. Therefore, the growth experiment should at least be repeated to check if the phenomena discussed above can be reproduced in order to draw clear conclusions.

5.2 Expression studies under assay conditions

Expression studies under assay conditions were performed to find out after how many hours of cultivation highest amounts of dioxygenase are expressed. The findings of the experiment were considered for the development of the growth-based selection assay. Expression was investigated for AvLDO, BtIDO and GriE in minimal medium with L-leucine and minimal medium with L-isoleucine. All dioxygenases were expressed in both media.

However, highest amounts of GriE (Figure 20) were already expressed after 24 hours of cultivation in minimal medium supplemented with L-leucine. Afterwards only low amounts of GriE were expressed, whereas in L-isoleucine expression was rather similar from 24 to 72 hours. As L-leucine is the preferred substrate of GriE, whereas the enzyme is unreactive towards L-isoleucine (Zwick and Renata, 2018b), differences in the expression pattern might be related to substrate preferences of the enzyme. Further, growth studies revealed that *E.coli* BL21(DE3) 3Δ*sucA* cells harbouring GriE started their exponential growth much earlier (after 60 hours of cultivation) than cells harbouring AvLDO (120 hours) and BtIDO (120 hours) (Hanreich, 2019). It is assumed that the strong expression of GriE after 24 hours in L-leucine containing media promotes earlier exponential growth. Further, it coincides with observations of this study, where *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_GriE] always was enriched in minimal medium with L-leucine in a mixed cultivation of all three dioxygenase harbouring strains (Chapter 4.4, Figure 23 (a)). Moreover, 3Δ*sucA*[pET28a_GriE] showed highest growth in the separate cultivation of all three dioxygenase harbouring strains (Chapter 4.5, Figure 24(a)).

In contrast, highest amounts of BtIDO (Figure 18) were expressed in minimal medium with L-isoleucine after 24 hours of cultivation, whereas significantly lower amounts were formed afterwards. As L-isoleucine is its preferred substrate (Hibi *et*

al., 2011) with a specific activity of 0.71 ± 0.29 U/mg and the exponential growth of BtIDO harbouring 3Δ*sucA* cells cultivated in minimal medium with L-isoleucine is much earlier (after 40 hours of cultivation with a specific growth rate of 0.166 h^{-1}, Chapter 4.1.1) compared to AvLDO (140 hours with a specific growth rate of 0.087 h^{-1}, Chapter 4.1.2*)* and GriE (120 hours, (Hanreich, 2019)) results are in coincidence with previous findings. Besides, *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO] always was enriched in minimal medium with L-isoleucine in a mixed cultivation of all three dioxygenase harbouring strains (Chapter 4.4, Figure 23(b)) and showed highest growth of all three dioxygenase bearing strains in the separate cultivation experiment (Chapter 4.5, Figure 24(a)).

The highest expression of AvLDO (Figure 19) was also observed after 24 hours in its preferred substrate L-leucine (Correia Cordeiro *et al*., 2018) and then it declined after 72 hours, whereas after 144 hours no expression was observed anymore. As for BtIDO and GriE, the enzyme expression was linked to growth, as no expression could be detected after the stationary phase was reached. Generally, expression of AvLDO under assay conditions appeared to be lower than expression of the other two dioxygenases BtIDO and GriE. These findings support the result of the mixed cultivation from this study (Chapter 4.4) where *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO] was overgrown by 3Δ*sucA*[pET28a_GriE] in its preferred substrate L-leucine and by 3Δ*sucA*[pET28a_BtIDO] in its preferred substrate L-isoleucine. Considering the results of the separate cultivation of all three dioxygenase harbouring strains (Chapter 4.5), 3Δ*sucA*[pET28a_AvLDO] showed significantly lower growth in minimal medium with L-leucine compared to 3Δ*sucA*[pET28a_GriE] (Chapter 4.5, Figure 24(a)). Higher expression of BtIDO and GriE promotes the restoring of the broken TCA cycle resulting in faster growth. To conclude, the rapid loss of 3Δ*sucA*[pET28a_AvLDO] in mixed cultivation could be linked to slower growth caused by weaker expression of AvLDO compared to the other two dioxygenases BtIDO and GriE.

5.3 Multiplex PCR

A multiplex PCR for the detection of the different dioxygenase genes encoding AvLDO, BtIDO and GriE was established to study the enrichment of the different enzyme-harbouring selection strains in a mixed iterative cultivation. For the implementation of a growth-based selection assay, it was important to develop a

reliable method for fast detection to find out which conditions are beneficial for enrichment. In contrast to singleplex PCR, several primer pairs are acting in a multiplex system requiring some additional factors to be considered. While buffers, dNTPs and DNA polymerases of commercial 'ready-to-use' multiplex kits often do not need any further optimization, a carful assessment of the primer concentrations and the thermocycling reaction are still essential for the achievement of stable and balanced reactions. Primer design is crucial as their internal stability, melting temperature and mutual interference can strongly influence their performance (Apte and Daniel, 2009).

Therefore, the aim was to design primer pairs with similar length, melting temperatures and GC-content. Further, amplicons need to show difference in size allowing unambiguous differentiation of the generated fragments in gel electrophoresis (Sint *et al*., 2012). As the same vector backbone was used for all three genes, a limited region on the specific dioxygenase gene had to be selected to design fragments of different length. As Cerna *et al.* (Cerna *et al*., 2003) used a size difference of at 150 – 170 bp, which enabled a clear discrimination, similar size differences were aimed in this study.

During implementation it was further important to check for unbalanced amplification strength. As it is known that fragments targeted by more efficient primer pairs are amplified preferentially (Markoulatos *et al*., 2002), amplification of the desired fragments in a single- and multiplex approach was compared. As expected, in the singleplex reaction for the longest fragment AvLDO (737 bp) a more intensive band was observed than for GriE (549 bp) and BtIDO (405 bp). Bands of the two smaller fragments showed rather similar intensities, which would mean, that slightly more fragment of BtIDO is present than of GriE as the fragment obtained from BtIDO is smaller. In contrast, under multiplex conditions, intensities of the fragments were rather similar for each annealing temperature applied (Figure 21). Provided that the same amount of template was used for each primer pair, it would mean that multiplex conditions have a slightly negative effect on the efficiency of the AvLDO primers. However, it must be kept in mind that the results of the spectrophotometer applied to determine the concentration of the templates used is not completely precise. Therefore, it can be concluded that the multiplex PCR established for dioxygenase detection is a reliable qualitative method to proof the presence of the

different plasmids, but a precise quantitative detection is not possible within this framework.

5.4 Mixed cultivation of *E. coli* BL21(DE3) 3Δ*sucA* [pET28a_AvLDO], 3Δ*sucA* [pET28a_BtIDO] and 3Δ*sucA* [pET28a_GriE] and selection assay setup

A mixed cultivation was performed to study dioxygenase enrichment in a mixed culture of the triple-knockout selection strain *E. coli* BL21(DE3) 3Δ*sucA* harbouring either BtIDO (L-isoleucine dioxygenase from *B. thuringiensis*), AvLDO (L-leucine dioxygenase from *A. variabilis*) or GriE (leucine hydroxylase from *Streptomyces sp*. DSM40835). Information obtained from this study should assist the implementation of a growth-based assay for the selection of improved enzyme variants. The established multiplex PCR enables a reliable qualitative detection of different dioxygenase harbouring strains. Whereas the enrichment of *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_GriE] in L-leucine and the enrichment of 3Δ*sucA*[pET28a_BtIDO] in L-isoleucine, already discussed in Chapter 5.2, is in accordance with previous studies, the early disappearance of 3Δ*sucA*[pET28a_AvLDO] in minimal medium supplemented with its preferred substrate L-leucine (Correia Cordeiro *et al*., 2018) was not expected in advance. As the multiplex PCR was established carefully, it could be excluded that the primer pairs for the detection of AvLDO work with a considerably lower efficiency compared to the other primers used. Therefore, a large excess of 3Δ*sucA*[pET28a_AvLDO] was applied in parallel to the standard experiment. However, it only led to a more intensive PCR band for the ONC of AvLDO compared to the approach using all three strains in equal concentrations, but the AvLDO harbouring strains disappeared nevertheless, after one cycle of cultivation. Further, the backbone of pET28a_AvLDO was sequenced and point mutations in the kanamycin resistance gene were detected. After repairing the construct, the experiment was repeated. Although a higher transformation efficiency of *E. coli* BL21(DE3) 3Δ*sucA* was observed with the new construct, it had no influence on the mixed growth experiment. Linking the results of the growth studies conducted by Hanreich (Hanreich, 2019) and the expression studies carried out in this work (Chapter 4.1.2), both performed under similar conditions applied for the mixed cultivation, revealed that the fast disappearance of 3Δ*sucA*[pET28a_AvLDO] could be explained by the different expression levels of AvLDO and GriE. Whereas

extremely high amounts of GriE are formed after 24 hours of cultivation in presence of L-leucine, expression of AvLDO is considerably lower. As the expression of Fe(II)/α-KG dependent dioxygenases is essential for the restoring of the impaired TCA cycle to reduce metabolic stress and enable biomass formation (Theodosiou *et al*., 2017), it is feasible that 3Δ*sucA*[pET28a_AvLDO] was overgrown by 3Δ*sucA*[pET28a_GriE] forming high amounts of enzyme at an earlier time. Of course, it could also be explained by different specific enzyme activities of the two enzymes. Since only the specific activity of AvLDO (0.22 ± 0.024 U/mg) (Correia Cordeiro *et al*., 2018) and BtIDO (0.31 ± 0.04 U/mg) (Hibi *et al*., 2011) towards L-leucine are known it is unclear whether the enrichment of GriE is caused by higher expression levels or higher specific enzyme activities. Comparing the presence of 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_BtIDO] in the mixed cultivation, the former strain got lost one cycle earlier than the latter cultivated in minimal medium supplemented with L-leucine (Figure 23 (a)). This observation could be linked to the slightly higher specific activity of BtIDO compared to AvLDO towards L-leucine. Regarding the expression after 24 hours cultivation in minimal medium with L-leucine, AvLDO seems to be expressed in slightly higher amounts than BtIDO. It indicates that in this case the difference in specific activity rather influences the enrichment in mixed cultivation than the different expression levels.

In contrast, the enrichment of BtIDO in minimal medium with L-isoleucine after one cycle of cultivation can be explained easily by previous studies. Studies of the substrate spectrum of AvLDO and GriE reported that both do not convert L-isoleucine (Correia Cordeiro *et al*., 2018; Zwick and Renata, 2018b) explaining why BtIDO with a specific activity of 0.71 ± 0.29 U/mg (Hibi *et al*., 2011) was enriched in the mixed culture.

Regarding the setup of the final assay, 72 hours adaption time for the first cultivation cycle was a trade-off between the long lag-phases ranging from 40 to 140 hours (Hanreich, 2019) and the fact that dioxygenase expression seemed to decline after 72 hours according to the expression studies under assay conditions conducted in this study (Chapter 4.1.2). Longer adaption times are not only disadvantageous due to a risk of contamination, but also risk of plasmid loss due to the ability of restoring cell growth via other side pathways, which were already discussed in detail in Chapter 5.1.

As Yun *et al.* (Yun *et al*., 2005) suggested an OD of 0.3 for the inoculation of fresh minimal medium of an enrichment culture for directed evolution an OD-based approach was carried out in this study initially. However, growth of mixed cultures in minimal medium varied from cycle to cycle and therefore, this approach appeared to be very impracticable. In contrast, Zhang *et al.* (Zhang *et al*., 2017) used a time-based approach with 24 hours cultivation for the enrichment of BtIDO in a microplate assay. As media and conditions for the enrichment culture used in this study were rather similar, following experiments focused on a time-based approach. The main issue is that cells should not have reached stationary phase when they are used for inoculation of fresh medium. Therefore, the OD was measured after each cycle and optical densities of 0.3 to 0.8 were reported for the mixed cultivation. Since final ODs measured under this setup were considerably higher, 24 hours of growth appeared to be suited intervals to avoid the stationary phase. Besides, it is known that IPTG used for induction in this study is stable at least 32 hours in minimal medium (Politi *et al*., 2014). Hence, it is reasonable to prefer 24 hours cultivation per cycle over longer periods. Regarding the stability of kanamycin, no decay was observed for incubation at 50 °C and pH 7.3 for 120 hours (Peteranderl *et al*., 1990). Hence, the medium used for the cultivation of thermostable *Clostridium thermohydrosulfuricum* was not the same than the minimal medium used in this study nor the conditions were exactly the same. However, it demonstrates that the stability of kanamycin is not the limiting factor of the determination of the cultivation times for the selection assay.

5.5 Separate cultivation of *E. coli* BL21(DE3) 3Δ*sucA* [pET28a_AvLDO], 3Δ*sucA* [pET28a_BtIDO] and 3Δ*sucA* [pET28a_GriE]

As the mixed cultivation (Chapter 4.5) raised several issues, for instance why 3Δ*sucA* cells harbouring pET28a_GriE enriched in minimal medium with L-leucine, although AvLDO is capable of converting L-leucine as well (Correia Cordeiro *et al*., 2018; Zwick and Renata, 2018b), separate cultivation was performed. More specifically, all three strains *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_AvLDO], 3Δ*sucA*[pET28a_BtIDO] and 3Δ*sucA*[pET28a_GriE] were grown in test tubes and ODs were determined after each cultivation cycle. As 3Δ*sucA*[pET28a_GriE] already showed 3-fold higher growth in minimal medium with L-leucine than

3Δ*sucA*[pET28a_AvLDO] in separate cultivation after 72 hours, it is obvious why the latter was overgrown by the former in mixed cultivation after switching to 24-hour cultivation intervals. As only 25 µL of the 72-hour culture (t_0) were used to inoculate t_1, the amount of AvLDO bearing cells used for inoculation must have been at minimum 3-fold lower in the mixed cultivation. Further, the OD of 3Δ*sucA*[pET28a_GriE] was only cut half switching from 72 to 24-hour cultivation, whereas the OD of 3Δ*sucA*[pET28a_AvLDO] almost approached zero in the separate cultivation. This observation demonstrated that the differences in growth are not only present in shake flasks, where 3Δ*sucA*[pET28a_AvLDO] had a 2-fold longer lag-phase in minimal medium than 3Δ*sucA*[pET28a_GriE] (Hanreich, 2019), but also in test tubes.

Although *E. coli* BL21(DE3) 3Δ*sucA*[pET28a_BtIDO] obtained the highest OD in minimal medium with L-leucine after 72 hours, the OD almost dropped to zero after the first 24-hour cultivation. These results coincide with previous results from Hanreich (Hanreich, 2019), where the lag-phase of 3Δ*sucA*[pET28a_BtIDO] was almost 2-fold higher than for 3Δ*sucA*[pET28a_GriE] and explain why 3Δ*sucA*[pET28a_GriE] cells overgrow the other L-leucine converting dioxygenase harbouring strains 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_BtIDO]. Further, it could also be linked to the strong expression of GriE after 24 hours in minimal medium with L-leucine, already discussed in Chapter 4.2.

While the results of the separate cultivation in minimal medium with L-leucine are in agreement with previous results, it is unexplained why 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_GriE] show considerable growth in the test tubes in minimal medium with L-isoleucine, even in the 24-hour cultivations, although they are not expected to convert L-isoleucine (Correia Cordeiro *et al*., 2018; Zwick and Renata, 2018b). Moreover, lag-phases of 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_GriE] are the same in minimal medium with L-isoleucine than lag-phases of 3Δ*sucA*[pET28a_AvLDO] and 3Δ*sucA*[pET28a_BtIDO] in minimal medium with L-leucine, therefore it is unclear why the former two strains show significantly better growth, especially in the 24-hour cultivation intervals. In particular for 3Δ*sucA*[pET28a_AvLDO] it is unclear why growth in minimal medium supplemented with L-isoleucine is 2.5-fold higher than in L-leucine, although the former should not be converted (Correia Cordeiro *et al*., 2018). Besides, in the

growth studies for AvLDO in shake flasks in this study, as well as in previous studies (Hanreich, 2019) growth started earlier in M9 medium with L-leucine. Moreover, in this study the specific growth rate of 3Δ*sucA*[pET28a_AvLDO] cultivated in shake flasks (Chapter 4.1.2) was 1.3-fold higher in minimal medium with L-leucine than in minimal medium with L-isoleucine. As the optical density of 3Δ*sucA*[pET28a_AvLDO] in the re-screening in minimal medium with L-isoleucine (Chapter 4.9) also was considerably lower, the experiment should be repeated in both media for 3Δ*sucA*[pET28a_AvLDO] before any firm conclusions can be drawn.

5.6 Generation of a mutant library of pET28a_AvLDO

As no published crystal structure of AvLDO was available, rational design could be excluded for enzyme engineering (Ibrahim and El-Dewany, 2018). However, some extent of mechanistic understanding was available due to crystal structures of other Fe(II)/α-KG dependent dioxygenases as GriE (Lukat *et al*., 2017) or SadA (Qin *et al*., 2013). Valuable information of the reaction mechanism and knowledge of the protein family obtained from the 3DM (3D molecular class specific information systems) (Kuipers *et al*., 2010) assisted in identifying hotspots being likely to influence the desired substrate promiscuity. The aim was to extend the substrate scope of AvLDO towards L-isoleucine and even towards α-substituted amino acids, like α-methyl-L-leucine for a broader industrial application. To increase the likelihood of generating variants with improved substrate promiscuity, two positions in close proximity to the substrate binding site (I70 and F78) were targeted for site-directed saturation mutagenesis (Figure 25), (Schweiger, unpubl.). According to information obtained from the 3DM, at position 70 the degenerate codon DYW coding for small or mostly hydrophobic amino acids was selected, whereas at positions 78 the codon NTT generating mostly hydrophobic amino acids as well as TAT (tyrosine) were selected (Schweiger, unpubl.). The library was generated with a conventional QuikChange protocol using a mix of 48 nucleotides long degenerate oligonucleotide primers, which is ideal for the creation of random mutagenesis libraries with only two mutations at positions close to each other. For instance, mutation efficiencies of greater than 80 % can be achieved with a simple, one-day protocol and the parental methylated and the parental/mutant hemi-methylated hybrid DNAs can be selectively eliminated by *Dpn*I digest. Applying this step, upon transformation the majority of clones contain the desired mutation and the enrichment of the wild type can be prevented (Hogrefe *et al*., 2002).

5.7 Selection of AvLDO variants and re-screening

As higher enzyme activity should lead to a greater supply of succinate and in turn to the formation of larger clones on agar plates of minimal medium supplemented with the desired substrate (Zhang *et al*., 2017), this strategy was applied to detect AvLDO variants with improved properties. Iterative growth cycles in liquid minimal medium were performed before plating to even enhance the selection pressure and improve the probability of enriching improved variants.

The large difference in optical densities obtained from the liquid culture of the library compared to the liquid culture of the wild type (Figure 26) already indicates the presence of variants with improved catalytic properties. The OD_{600} of the library is 2-fold higher after 72-hours of cultivation and differences even increase switching to the 24-hour cultivation cycles.

Whereas Zhang *et al.,* who generated a mutant library of BtIDO via error prone PCR, only obtained five clones with increased size on M9-αKG plates employing the knockout strain *E. coli* Δ*sucA*Δ*aceA* (Zhang *et al*., 2017)*,* the proportion of clones with larger size in this experiment was significantly higher as it can be seen on the example plate in Figure 27. The higher number of larger sized clones can be explained by the fact that in this study a semi-rational approach was applied already creating a smaller and smarter library . Further, it is assumed that some non-beneficial variants were already negatively selected in the iterative cultivation approach. This assumption is underlined by the sequencing results, where eight out of the 14 picked clones showed the same genotype.

14 clones were picked for re-screening to exclude the selection of false positive clones. Plasmids of the 14 clones were isolated and re-transformed in fresh 3Δ*sucA* cells to avoid falsified re-screening results caused by mutations in the knockout selection strain itself. As the differences in the optical densities between the library and the wild type were remarkable, it was assumed to detect improved clones using the same iterative cultivation approach for re-screening than it was used in the assay itself. Further, a growth curve of each mutant strain was created in parallel to increase the chance of detecting false positive clones.

After Zhang *et al.* (Zhang *et al*., 2017) only detected one false positive clone out of five clones by re-screening, a large proportion of actual positive clones was expected in this study as well. According to the re-screening results only three

mutants were detected to be false positive. However, sequencing revealed that the plasmids of two false positive strains (3Δ*sucA*_varE and 3Δ*sucA*_varF) could not be aligned to the wild type template and for one additional strain, which showed faster growth than the wild type (3Δ*sucA*_varG) no sequencing results were received at all. However, it is unclear if these strains were a result of contamination related to the re-transformation or contamination during the re-screening itself. By all means, they do not harbour any improved variant of AvLDO.

Therefore, only three different variants with improved growth were detected, whereby one variant was picked eight out of 14 times. The most prominent variant shows a L-threonine at position 70 and no exchange at position 78. Hence, the exchange of the hydrophobic L-isoleucine by the extremely hydrophilic L-threonine is assumed to have a positive effect on the substrate binding of L-isoleucine. According to information from the 3DM database, where a superfamily for AvLDO (Clavaminate synthase like superfamily), which is a group of structures with common structural fold (Kuipers *et al*., 2010), was created, L-isoleucine has only 1.07 % abundancy within the superfamily, whereas L-threonine shows a higher frequency of 9.19 % (Schweiger, unpubl.). The aromatic phenylalanine, which is pointing towards the active site according to structural information from the 3DM database (Figure 25), (Schweiger, unpubl.) seems to be quite important as it is contained in nine selected variants. Considering the abundancy of phenylalanine based on a multiple structural alignment of 76 members of the 3DM superfamily created for AvLDO, phenylalanine is conserved within the superfamily (73.68 %) (Schweiger, unpubl.). The conservation could indicate a fundamental function for substrate binding, which would explain the high abundancy of phenylalanine at position 78 within all the selected variants in this study. Interestingly, the double-leucine mutant 3Δ*sucA*_varC showed delayed growth compared to the wild type, whereas mutant 3Δ*sucA*_varD, which only showed an exchange to L-leucine at position 70, but conserved to L-phenylalanine at position 78 displayed earlier growth and a 3.5-fold higher specific growth rate than the wild type. According to the superfamily created by 3DM, the two aromatic amino acids phenylalanine and tyrosine are mainly found in this region of the structural fold (Schweiger, unpubl.). This observation indicates a negative effect of L-leucine at position 78 on substrate binding and conversion. While mutant 3Δ*sucA*_varD (I70L, F78F) showed exponential growth after 78 hours reaching a final OD of almost 1.5 (specific growth

rate: 0.087 h^{-1}), mutant 3Δ*sucA*_varA and B (I70T, F78F) only showed a specific growth rate of 0.076 h^{-1} and the final OD_{600} was not higher than the final OD_{600} of the wild type. However, the latter mutants started exponential growth already in the first 24 hours of cultivation. As 3Δ*sucA*_varD (I70L, F78F) would not have been detected in the iterative re-screening approach despite its improved growth behaviour, it is recommended to use the growth curves for re-screening rather than the iterative approach in prospective experiments. Mutant 3Δ*sucA*_varH (I70T, F78I) overall showed the best growth behaviour, as it showed an almost similar specific growth rate than 3Δ*sucA*_varD and started exponential growth in the first 24 hours obtaining a higher final OD_{600} (1.3) than the wild type (1.0). Further, it also exerted highest growth after 72 hours in the iterative re-screening approach. Consequently, the exchange of the hydrophobic aromatic phenylalanine by the hydrophobic aliphatic isoleucine at position 78 in combination with the threonine at position 70 is assumed to be particularly advantageous. Although, both re-screening systems (iterative approach and growth curve) revealed the same false negative variants, it is impossible to predict, which of the faster growing variants exerts higher activity towards the substrate. Therefore, the establishment of a proper quantification method as HPLC is indispensable for further characterisation of selected variants and verification of the results obtained from this study.

5.8 Expression of selected variants under assay conditions

Expression of faster growing variants detected in the re-screening (Chapter 4.9) was performed to investigate if the respective mutations have any influence on expression of the enzyme variants. As expression of the wild type pET28a_AvLDO is quite difficult and several rounds of expression optimisation performed by Hanreich (Hanreich, 2019) including variable inducer concentrations and temperatures, co-expression of chaperones or autoinduction did not lead to any improvement, it was decided to express the selected variants under assay conditions in minimal medium supplemented with L-isoleucine. According to the results of the expression studies under assay conditions (Chapter 4.2), highest amount of protein was formed after 24 and 48 hours of cultivation. Therefore, cells were harvested after 24 hours and both the whole cell protein profile and a separation of soluble and insoluble fraction were analysed via SDS-gel. Whereas clearly visible amounts of the wild type enzyme were detected in all fractions on the gel, it was not the case for the selected variants. Further, the protein profile of the

insoluble fraction of variant G, which could not be sequenced, looked completely different than the profiles of the other variants. This observation suggests that mutant G in the re-screening was no AvLDO bearing *E. coli* 3Δ*sucA* strain.

Although, the expression of the variants seems to be even worse than the expression of the wild type, it cannot be concluded that the amino acid exchanges have a negative effect on expression of the enzyme as the expression of the wild type enzyme is far from optimised and varied in the two expressions conducted in this study. At least, these findings do not indicate that faster growth of selected variants is caused by stronger expression of the respective variants. Therefore, it can be assumed that the earlier growth of selected variants is rather induced by an improved conversion of the substrate L-isoleucine.

Although, Cordeiro *et al.* (Correia Cordeiro *et al*., 2018) managed to express proper amounts for quantification of AvLDO using the pET28a vector system employing *E. coli* BL21(DE3) as expression host, in this study it was not possible using the same system and conditions. Therefore, the re-cloning of the gene in a different expression vector system, for instance a pASK vector, which uses the *tet* (tetracycline) promoter, is highly recommended. The promoter works independently from the genetic background of *E. coli* enabling expression control over an 5000-fold range (~2.8 times lower for *lac* with IPTG) with low concentrations of the highly affine anhydrotetracycline (Bertram and Hillen, 2008; Lutz and Bujard, 1997).

5.9 Verification of L-isoleucine conversion of selected variants

The conversion of L-isoleucine to L-hydroxyisoleucine by faster growing variants detected in the re-screening (Chapter 4.9) should be verified via TLC. Therefore, expression of selected variants as well as the wild type enzyme in *E. coli* BL21(DE3) was performed under optimised conditions obtained from previous studies on AvLDO (Correia Cordeiro *et al*., 2018; Hanreich, 2019). After cell disruption via sonication the CFE was applied to set up biocatalytic reactions. Unfortunately, conversion of the selected variants could not be verified. For all enzymes tested, large amount of substrate was detectable in all samples taken during the activity assay, but no product was formed. As unspecific compounds are mainly detected in the reaction samples, but were hardly visible in the samples taken from the blank reaction, it is assumed that it results from unspecific *E. coli* proteins rather than from nonenzymatic degradation products of reaction compounds. Since the t_0 sample

shows the same unspecific by-products than the t_{24} sample, it is assumed that the compounds result from *E. coli* and were already present in the CFE before it was used to set up biocatalytic reactions. However, these results do not imply the selected variants are unable to convert the substrate as the positive control using the wild type enzyme and L-leucine as a substrate did also not lead to any product formation. Previous results from Hanreich (Hanreich, 2019) showed product formation using purified protein to set up the biocatalytic reaction. However, small amounts of product were also formed in the reactions set up with the CFE (Hanreich, 2019). These findings indicate that AvLDO expression is rather weak under the conditions applied and a considerably high amounts of endogenous *E. coli* proteins is assumed to be present in the CFE leading to unspecific by-product formation. Of course, purified protein could be used to set up the reactions. However, the optimised protocol established by Hanreich (Hanreich, 2019) is time consuming and did not yield highly active and pure protein as it would be required for any quantitative detection method. The use of a different expression system with tighter control as already discussed above (Chapter 5.8) would enable a faster verification of conversion. Stronger expression and less basal activity than obtained from the pET28a expression system (Novy and Morris, 2001) would not require laborious purification steps. The strong expression would enable the dilution of the CFE despite having enough enzyme present to perform the reaction. Therefore, less background activity from *E. coli* enzymes would occur and product formation could be detected easier. Further, the use of a gentler cell disruption method as chemical or biological lysis (Shehadul Islam *et al*., 2017) should be considered to yield higher amounts of active enzyme in the CFE than obtained from cell disruption via sonication. Otherwise, at least the sonication protocol should be optimised before the experiment is repeated. It is assumed, that a large extent of the enzyme lost its activity during sonication, as samples of the CFE analysed via SDS-gel, which should contain the soluble and active enzyme, hardly showed any protein. Therefore, it is plausible why no conversion of L-isoleucine was detected via TLC-based screening as almost no soluble and active enzyme was present in the reaction.

6 Conclusion and outlook

In this study, a growth-based selection assay employing the triple-knockout strain *E. coli* BL21(DE3) 3Δ*sucA* was implemented for the selection of dioxygenase variants. As previous growth experiments in shake flaks led to long lag-times, inconvenient handling and in turn to a risk of contamination, test tubes were used to set up an iterative cultivation approach. A mixed cultivation of the selection strain 3Δ*sucA* harbouring either the Leucine dioxygenase from *A. variabilis* (AvLDO), the Isoleucine dioxygenase from *B. thuringiensis* (BtIDO) or the L-leucine dioxygenase from *Streptomyces* strain DSM40835 (GriE), followed by a qualitative detection via multiplex PCR, should assist in the setup of the assay finding useful cultivation periods. Mixed cultivation revealed a significantly faster growth of 3Δ*sucA*[pET28a_GriE] in minimal medium with L-leucine and of 3Δ*sucA*[pET28a_BtIDO] in minimal medium with L-isoleucine leading to enrichment of these strains in the respective media.

Cultivation for 72 hours in M9 minimal medium should enable adaption, whereas further inoculation of fresh minimal medium in 24 hour intervals was intended to increase the selection pressure and promote the enrichment of strains bearing improved dioxygenase variants. The liquid culture was plated on M9-agar and cultivated for three days. Higher enzyme activity should correspond to a greater supply of succinate, which is necessary to shunt the broken TCA and therefore larger colonies should be formed on the plate.

The selection assay itself was applied for the selection of a small library of AvLDO containing 35 different amino acid variants. The mutant library was created via site-directed saturation mutagenesis applying degenerated primers for the mutation of two amino acids in close proximity to the substrate binding site of the enzyme to enhance the narrow substrate scope of AvLDO towards L-isoleucine.

After the selection of larger clones, re-transformation, genotypic characterisation, and re-screening in test tubes under assay conditions was performed. In total, three different mutants displaying faster growth than the wild type enzyme were detected during the re-screening, whereby one variant showed a high abundancy and was present in eight out of 14 clones selected from the plate. The respective variant showed one exchange from L-isoleucine to L-threonine and exhibited a significantly

earlier start in exponential growth as well as a 3-fold higher specific growth rate in the re-screening.

Expression of the selected variants under assay conditions was also studied, but variants were expressed in even lower amounts than the wild type enzyme. Since expression studies of the three dioxygenases AvLDO, BtIDO and GriE under assay conditions showed higher expression of BtIDO and GriE compared to AvLDO, it is assumed that expression of AvLDO in adequate quantities using the pET vector system is quite difficult. Therefore, it is recommended to re-clone the gene into a different expression vector, for instance a vector with a tetracycline expression system enabling tighter expression control. It would also be beneficial for the TLC screening conducted in this study, as stronger expression would enable the dilution of the CFE reducing background activity of *E. coli*. In this study, conversion of L-isoleucine by improved variants from the re-screening could not be verified due to low amounts of active protein in the CFE only yielding by-product formation, but no conversion to L-hydroxyisoleucine.

Optimised expression would further be important to obtain larger amounts of purified and active enzyme for verification of the results of this study via quantification of the selected variants by HPLC. As soon as a proper quantification system is established and the improved conversion of selected variants can be verified, the assay could be further used to select variants converting α-methyl amino acids to even increase the attractiveness of the enzyme for industrial applications.

7 References

Adrio JL, Demain AL. 2014. Microbial Enzymes: Tools for Biotechnological Processes. *Biomolecules* **4**:117–139.

Aik WS, Chowdhury R, Clifton IJ, Hopkinson RJ, Leissing T, McDonough MA, Nowak R, Schofield CJ, Walport LJ. 2015. Introduction to structural studies on 2-oxoglutarate-dependent oxygenases and related enzymes. In: . *2-Oxoglutarate-Dependent Oxyg.* Royal Society of Chemistry, pp. 59–94.

Aik W, McDonough MA, Thalhammer A, Chowdhury R, Schofield CJ. 2012. Role of the Jelly-Roll Fold in Substrate Binding by 2-oxoglutarate Oxygenases. *Curr. Opin. Struct. Biol.* **22**:691–700.

Apte A, Daniel S. 2009. PCR primer design. *Cold Spring Harb. Protoc.* **2009**:pdb-ip65.

Arnold FH. 2015. The Nature of Chemical Innovation: New Enzymes by Evolution. *Q. Rev. Biophys.* **48**:404–410.

Aronson JN, Borris DP, Doerner JF, 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.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, Wiley CJ, Allison RD, Bittner M, Blackshaw S. 2003. Current Protocols in Molecular Biology. New York: John Wiley & Sons. Vol. 1 146–146 p.

Bajaj P, Singh NS, Virdi JS. 2016. *Escherichia coli* β-lactamases: What really matters. *Front. Microbiol.* **7**:1–14.

Bertram R, Hillen W. 2008. The application of Tet repressor in prokaryotic gene regulation and expression. *Microb. Biotechnol.* **1**:2–16.

Blank LM, Ebert BE, Buehler K, Bühler B. 2010. Redox Biocatalysis and Metabolism: Molecular Mechanisms and Metabolic Network Analysis. *Antioxid. Redox Signal.* **13**:349–394.

Blasiak LC, Vaillancourt FH, Walsh CT, Drennan CL. 2006. Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440**:368–371.

Blesl J, Trobe M, Anderl F, Breinbauer R, Strohmeier GA, Fesko K. 2018. Application of Threonine Aldolases for the Asymmetric Synthesis of α-Quaternary α-Amino Acids. *ChemCatChem* **10**:3453–3458.

Broca C, Gross R, Petit P, Sauvaire Y, Manteghetti M, Tournier M, Masiello P, Gomis R, Ribes G. 1999. 4-Hydroxyisoleucine: Experimental evidence of its insulinotropic and antidiabetic properties. *Am. J. Physiol. - Endocrinol. Metab.* **277**:617–623.

Brun KA, Linden A, Heimgartner H. 2008. Synthesis and Conformational Analysis of Pentapeptides Containing Enantiomerically Pure 2, 2-Disubstituted Glycines. *Helv. Chim. Acta* **91**:526–558.

Cerna JF, Nataro JP, Estrada-Garcia T. 2003. Multiplex PCR for detection of three plasmid-borne genes of enteroaggregative *Escherichia coli* strains. *J. Clin. Microbiol.* **41**:2138–2140.

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

Crisma M, Peggion C, Moretto A, Banerjee R, Supakar S, Formaggio F, Toniolo C. 2014. The 2.0_5-helix in hetero-oligopeptides entirely composed of Cα, α-disubstituted glycines with both side chains longer than methyls. *Pept. Sci.* **102**:145–158.

Davids T, Schmidt M, Böttcher D, Bornscheuer UT. 2013. Strategies for the discovery and engineering of enzymes for biocatalysis. *Curr. Opin. Chem. Biol.* **17**:215–220.

Dawson JH. 1988. Probing Structure-Function Relations in Heme-Containing Oxygenases and Peroxidases. *Science (80-.).* **240**:433–439.

Denard CA, Ren H, Zhao H. 2015. Improving and repurposing biocatalysts via directed evolution. *Curr. Opin. Chem. Biol.* **25**:55–64.

Dong JJ, Fernández-Fueyo E, Hollmann F, Paul CE, Pesic M, Schmidt S, Wang Y, Younes S, Zhang W. 2018. Biocatalytic Oxidation Reactions: A Chemist's Perspective. *Angew. Chemie - Int. Ed.* **57**:9238–9261.

Enoki J, Meisborn J, Müller AC, Kourist R. 2016. A Multi-Enzymatic Cascade Reaction for the Stereoselective Production of γ-Oxyfunctionalyzed Amino Acids. *Front. Microbiol.* **7**:1–8.

Faber K. 1992. Biotransformations in Organic Chemistry 6th ed. Berlin: Springer. Vol. 4.

Fowden L, Pratt HM, Smith A. 1973. 4-Hydroxyisoleucine from seed of *Trigonella foenum-graecum*. *Phytochemistry* **12**:1707–1711.

Gamenara D, Seoane G, Méndez PS, de María PD. 2012. Redox biocatalysis: fundamentals and applications. John Wiley & Sons.

Gao SS, Naowarojna N, Cheng R, Liu X, Liu P. 2018. Recent examples of α-ketoglutarate-dependent mononuclear non-haem iron enzymes in natural product biosyntheses. *Nat. Prod. Rep.* **35**:792–837.

Gunsalus IC, Pederson TC, Sligar SG. 1975. Oxygenase-Catalyzed Biological Hydroxylations. *Annu. Rev. Biochem.* **44**:377–407.

Hanreich S. 2019. Identification and selection of active α-ketoglutarate dependent dioxygenases in engineered *Escherichia coli*; Graz University of Technology.

Hara R, Kino K. 2009. Characterization of novel 2-oxoglutarate dependent dioxygenases converting L-proline to *cis*-4-hydroxy-L-proline. *Biochem. Biophys. Res. Commun.* **379**:882–886.

Hausinger RP. 2004. Fe(II)/α-Ketoglutarate-Dependent Hydroxylases and Related Enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39**:21–68.

Hausinger RP. 2015. Biochemical Diversity of 2-Oxoglutarate-Dependent Oxygenases. *R. Soc. Chem. London*:1–58.

Hegg EL, Que L. 1997. The 2-His-1-Carboxylate Facial Triad - An Emerging Structural Motif in Mononuclear Non-Heme Iron(II) Enzymes. *Eur. J. Biochem.* **250**:625–629.

Herr CQ, Hausinger RP. 2018. Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. *Trends Biochem. Sci.* **43**:517–532.

Hibi M, Kawashima T, Kodera T, Smirnov S V., Sokolov PM, Sugiyama M, Shimizu S, Yokozeki K, 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, Sokolov PM, Smirnov S V., Kodera T, Sugiyama M, Shimizu S, Yokozeki K, 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, Ogawa J. 2016. Novel Enzyme Family Found in Filamentous Fungi Catalyzing *trans*-4-Hydroxylation of L-Pipecolic Acid. *Appl. Environ. Microbiol.* **82**:2070–2077.

Hibi M, 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.

Hogrefe HH, Cline J, Youngblood GL, Allen RM. 2002. Creating Randomized Amino Acid Libraries with the Quikchange® Multi Site-Directed Mutagenesis Kit. *Biotechniques* **33**:1158–1165.

Hou CT. 1986. Recent Progress in Research on Methanotrophs and Methane Monooxygenases. *Biotechnol. Genet. Eng. Rev.* **4**:145–168.

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

Hutton JJ, Tappel AL, Udenfriend S. 1967. Cofactor and Substrate Requirements of Collagen Proline Hydroxylase. *Arch. Biochem. Biophys.* **118**:231–240.

Iacobellis NS, Lavermicocca P, Grgurina I, Simmaco M, Ballio A. 1992. Phytotoxic properties of *Pseudomonas syringae* pv. syringae toxins. *Physiol. Mol. Plant Pathol.* **40**:107–116.

Ibrahim A, El-Dewany AI. 2018. Recent Trends for Discovery and Enhancement of Enzyme Function: A Review. *Egypt. J. Microbiol.* **0**:0–0.

Islam S, Leissing TM, Chowdhury R, Hopkinson RJ, Schofield CJ. 2018. 2-Oxoglutarate-Dependent Oxygenases. *Annu. Rev. Biochem.*

Jemli S, Ayadi-Zouari D, Hlima H Ben, Bejar S. 2016. Biocatalysts: Application and Engineering for Industrial Purposes. *Crit. Rev. Biotechnol.* **36**:246–258.

Jette L, Harvey L, Eugeni K, Levens N. 2009. 4-Hydroxyisoleucine: a plant-derived treatment for metabolic syndrome. *Curr. Opin. Investig. Drugs (London, Engl. 2000)* **10**:353–358.

Kang SH, Kang SY, Lee HS, Buglass AJ. 2005. Total Synthesis of Natural *tert*-Alkylamino Hydroxy Carboxylic Acids. *Chem. Rev.* **105**:4537–4558.

Kling A, Lukat P, Almeida D V, Bauer A, Fontaine E, Sordello S, Zaburannyi N, Herrmann J, Wenzel SC, König C. 2015. Antibiotics. Targeting DnaN for Tuberculosis Therapy Using Novel Griselimycins. *Science (80-.).* **348**:1106–1112.

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

Koketsu K, Shomura Y, Moriwaki K, Hayashi M, Mitsuhashi S, Hara R, Kino K, Higuchi Y. 2015. Refined Regio- and Stereoselective Hydroxylation of L-pipecolic Acid by Protein Engineering of L-proline *cis*-4-Hydroxylase Based on the X-ray Crystal Structure. *ACS Synth. Biol.* **4**:383–392.

Kuipers RK, Joosten HJ, Van Berkel WJH, Leferink NGH, Rooijen E, Ittmann E, Van Zimmeren F, Jochens H, Bornscheuer U, Vriend G, Martins Dos Santos VAP, Schaap PJ. 2010. 3DM: Systematic analysis of heterogeneous superfamily data to discover protein functionalities. *Proteins Struct. Funct. Bioinforma.* **78**:2101–2113.

Leemhuis H, Kelly RM, Dijkhuizen L. 2009. Directed evolution of enzymes: Library screening strategies. *IUBMB Life* **61**:222–228.

Li F, Zhang X, Renata H. 2019. Enzymatic C–H functionalizations for natural product synthesis. *Curr. Opin. Chem. Biol.* **49**:25–32.

Li M, Ho PY, Yao S, 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 ^{13}C labeling experiments. *Biochem. Eng. J.* **30**:286–296.

Li W, Zhang T, Ding J. 2015. Molecular basis for the substrate specificity and catalytic mechanism of thymine-7-hydroxylase in fungi. *Nucleic Acids Res.* **43**:10026–10038.

Li Y, Cirino PC. 2014. Recent Advances in Engineering Proteins for Biocatalysis. *Biotechnol. Bioeng.* **111**:1273–1287.

Li Z, Nimtz M, Rinas U. 2014. The metabolic potential of *Escherichia coli* BL21 in defined and rich medium. *Microb. Cell Fact.* **13**:1–17.

Lindqvist R, Barmark G. 2014. Specific Growth Rate Determines the Sensitivity of *Escherichia coli* to Lactic Acid Stress: Implications for Predictive Microbiology. *Biomed Res. Int.* **2014**.

Lukat P, Katsuyama Y, Wenzel S, Binz T, König C, Blankenfeldt W, Brönstrup M, Müller R. 2017. Biosynthesis of methyl-proline containing griselimycins, natural products with anti-tuberculosis activity. *Chem. Sci.* **8**:7521–7527.

Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Res.* **25**:1203–1210.

Markoulatos P, Siafakas N, Moncany M. 2002. Multiplex Polymerase Chain Reaction: A Practical Approach. *J. Clin. Lab. Anal.* **16**:47–51.

Martinez S, Hausinger RP. 2015. Catalytic mechanisms of Fe(II)- and 2-Oxoglutarate-dependent oxygenases. *J. Biol. Chem.* **290**:20702–20711.

Mitchell AJ, Dunham NP, Martinie RJ, Bergman JA, Pollock CJ, Hu K, Allen BD, Chang W, Jr JMB, Krebs C, Boal AK. 2018. Visualizing the reaction cycle in an iron(II)- and 2-(oxo)-glutarate- dependent hydroxylase. *J. Am. Chem. Soc.* **139**:13830–13836.

Miyazaki K TM. 2002. Creating Random Mutagenesis Libraries Using Megaprimer PCR of Whole Plasmid. *Biotechniques* **1038**:33(5): 1033–4, 1036–8.

Morley KL, Kazlauskas RJ. 2005. Improving Enzyme Properties: When Are Closer Mutations Better? *Trends Biotechnol.* **23**:231–237.

Munro AW, Taylor P, Walkinshaw MD. 2000. Structures of redox enzymes. *Curr. Opin. Biotechnol.* **11**:369–376.

Neumann CS, Fujimori DG, Walsh CT. 2008. Halogenation Strategies In Natural Product Biosynthesis. *Chem. Biol.* **15**:99–109.

Novy R, Morris B. 2001. Use of glucose to control basal expression in the pET System. *Innovations* **13**:13–15.

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

Parmee ER, Tempkin O, Masamune S, Abiko A. 1991. New catalysts for the asymmetric aldol reaction: chiral boranes prepared from. alpha.,. alpha.-disubstituted glycine arenesulfonamides. *J. Am. Chem. Soc.* **113**:9365–9366.

Parra LP, Agudo R, Reetz MT. 2013. Directed evolution by using iterative saturation mutagenesis based on multiresidue sites. *ChemBioChem* **14**:2301–2309.

Peteranderl R, Shotts EB, Wiegel J. 1990. Stability of Antibiotics under Growth Conditions for Thermophilic Anaerobes. *Appl. Environ. Microbiol.* **56**:1981–1983.

Peters C, Buller RM. 2019. Industrial Application of 2-Oxoglutarate-Dependent Oxygenases. *Catalysts* **9**.

Politi N, Pasotti L, Zucca S, Casanova M, Micoli G, Cusella De Angelis MG, Magni P. 2014. Half-life measurements of chemical inducers for recombinant gene expression. *J. Biol. Eng.* **8**:1–10. Journal of Biological Engineering.

Porter JL, Rusli RA, Ollis DL. 2016. Directed Evolution of Enzymes for Industrial Biocatalysis. *ChemBioChem* **17**:197–203.

Pratter SM, Ivkovic J, Birner-Gruenberger R, Breinbauer R, Zangger K, Straganz GD. 2014. More than just a halogenase: Modification of fatty acyl moieties by a trifunctional metal enzyme. *ChemBioChem* **15**:567–574.

Qin HM, Miyakawa T, Nakamura A, Hibi M, Ogawa J, Tanokura M. 2014. Structural optimization of SadA, an Fe(II)- and a-ketoglutarate-dependent dioxygenase targeting biocatalytic synthesis of *N*-succinyl-L-*threo*-3, 4-dimethoxyphenylserine. *Biochem. Biophys. Res. Commun.* **450**:1458–1461.

Qin H-M, Miyakawa T, Jia MZ, Nakamura A, Ohtsuka J, Xue Y-L, Kawashima T, Kasahara T, Hibi M, Ogawa J. 2013. Crystal structure of a novel N-substituted L-amino acid dioxygenase from *Burkholderia ambifaria* AMMD. *PLoS One* **8**.

Quasdorf KW, Overman LE. 2014. Catalytic Enantioselective Synthesis of Quaternary Carbon Stereocenters. *Nature* **516**:181–191.

Reetz MT, Prasad S, Carballeira JD, Gumulya Y, Bocola M. 2010a. Iterative saturation mutagenesis accelerates laboratory evolution of enzyme

stereoselectivity: rigorous comparison with traditional methods. *J. Am. Chem. Soc.* **132**:9144–9152.

Reetz MT, Soni P, Fernández L, Gumulya Y, Carballeira JD. 2010b. Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem. Commun.* **46**:8657–8658.

Ruff AJ, Dennig A, Schwaneberg U. 2013. To get what we aim for–progress in diversity generation methods. *FEBS J.* **280**:2961–2978.

Sambrook J, Russel DW. 2001. Molecular Cloning: A Laboratory Manual, 3 edn. New York: Cold Spring Harbor Laboratory Press.

Samsonova NN, Smirnov S V., Novikova AE, Ptitsyn LR. 2005. Identification of *Escherichia coli* K12 YdcW protein as a γ-aminobutyraldehyde dehydrogenase. *FEBS Lett.* **579**:4107–4112.

Schließmann A. 2010. Protein engineering of a *Pseudomonas fluorescens* esterase. Alteration of substrate specificity and stereoselectivity; Ernst-Moritz-Arndt University Greifswald.

Shehadul Islam M, Aryasomayajula A, Selvaganapathy PR. 2017. A Review on Macroscale and Microscale Cell Lysis Methods. Ed. Aaron T Ohta, Wenqi Hu. *Micromachines* **8**:83.

Shen Q, Zhang Y, Yang R, Hua X, Zhang W, Zhao W. 2015. Thermostability enhancement of cellobiose 2-epimerase from *Caldicellulosiruptor saccharolyticus* by site-directed mutagenesis. *J. Mol. Catal. B Enzym.* **120**:158–164.

Singh V, Rakshit K, Rathee S, Angmo S, Kaushal S, Garg P, Chung JH, Sandhir R, Sangwan RS, Singhal N. 2016. Metallic/bimetallic magnetic nanoparticle functionalization for immobilization of α-amylase for enhanced reusability in bio-catalytic processes. *Bioresour. Technol.* **214**:528–533.

Sint D, Raso L, Traugott M. 2012. Advances in multiplex PCR: Balancing primer efficiencies and improving detection success. *Methods Ecol. Evol.* **3**:898–905.

Smirnov S V., Kodera T, Samsonova NN, Kotlyarova VA, Rushkevich NY, Kivero AD, Sokolov PM, Hibi M, Ogawa J, Shimizu S. 2010. Metabolic engineering of *Escherichia coli* to produce (2*S*, 3*R*, 4*S*)-4-hydroxyisoleucine. *Appl. Microbiol. Biotechnol.* **88**:719–726.

Thakker C, Martínez I, San KY, Bennett GN. 2012. Succinate production in *Escherichia coli*. *Biotechnol. J.* **7**:213–224.

Theodosiou E, Breisch M, Julsing MK, Falcioni F, Bühler B, Schmid A. 2017. An Artificial TCA Cycle Selects for Efficient a-Ketoglutarate Dependent Hydroxylase Catalysis in Engineered *Escherichia coli*. *Biotechnol. Bioeng.* **114**:1511–1520.

Vaillancourt FH, Yeh E, Vosburg DA, Garneau-Tsodikova S, Walsh CT. 2006. Nature's inventory of halogenation catalysts: Oxidative strategies predominate. *Chem. Rev.* **106**:3364–3378.

Vaillancourt FH, Yin J, Walsh CT. 2005. SyrB2 in syringomycin E biosynthesis is a nonheme Fe^{II} α-ketoglutarate- and O_2-dependent halogenase. *Proc. Natl. Acad. Sci. U. S. A.* **102**:10111–10116.

Walsh CT, O'Brien R V., Khosla C. 2013. Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angew. Chemie - Int. Ed.* **52**:7098–7124.

Wang X, Wang G, Li X, Fu J, Chen T, Wang Z, Zhao X. 2016. Directed evolution of adenylosuccinate synthetase from *Bacillus subtilis* and its application in metabolic engineering. *J. Biotechnol.* **231**:115–121.

Wu I, Arnold FH. 2013. Engineered thermostable fungal Cel6A and Cel7A cellobiohydrolases hydrolyze cellulose efficiently at elevated temperatures. *Biotechnol. Bioeng.* **110**:1874–1883.

Wu LF, Meng S, Tang GL. 2016. Ferrous iron and α-ketoglutarate-dependent dioxygenases in the biosynthesis of microbial natural products. *Biochim. Biophys. Acta - Proteins Proteomics* **1864**:453–470.

Xiao H, Bao Z, Zhao H. 2015. High Throughput Screening and Selection Methods for Directed Enzyme Evolution. *Ind. Eng. Chem. Res.* **54**:4011–4020.

Xu F. 2005. Applications of oxidoreductases: Recent progress . *Ind. Biotechnol.* **1**:38–50.

Yuan L, Kurek I, English J, Keenan R. 2005. Laboratory-directed protein evolution. *Microbiol. Mol. Biol. Rev.* **69**:373–392.

Yun H, Hwang BY, Lee JH, Kim BG. 2005. Use of enrichment culture for directed evolution of the *Vibrio fluvialis* JS17 ω-transaminase, which is resistant to product inhibition by aliphatic ketones. *Appl. Environ. Microbiol.* **71**:4220–4224.

Zhang C, Ma J, Li Z, Liang Y, Xu Q, Xie X, Chen N. 2017. A strategy for L-isoleucine dioxygenase screening and 4-hydroxyisoleucine production by resting cells. *Bioengineered* **9**:1–8.

Zhuang Y, Yang G-Y, Chen X, Liu Q, Zhang X, Deng Z, Feng Y. 2017. Biosynthesis of plant-derived ginsenoside Rh2 in yeast via repurposing a key promiscuous microbial enzyme. *Metab. Eng.* **42**:25–32.

De Zotti M, Bobone S, Bortolotti A, Longo E, Biondi B, Peggion C, Formaggio F, Toniolo C, Dalla Bona A, Kaptein B. 2015. 4-Cyano-α-methyl-L-phenylalanine as a Spectroscopic Marker for the Investigation of Peptaibiotic-Membrane Interactions. *Chem. Biodivers.* **12**:513–527.

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

Zwick CR, Renata H. 2018b. 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.

8 Appendix

8.1 Chemicals and laboratory equipment

Table 21: Chemicals used in this study.

Chemicals	Source	CAS number
Acetic acid	Roth, Germany	64-19-7
Agar-Agar	Roth, Germany	67-64-1
Ammonium chloride	AppliChem, Germany	12125-02-9
Ampicillin sodium salt	Roth, Germany	7783-20-2
Ascorbate	Roth, Germany	50-81-7
Biotin	Roth, Germany	58-85-5
BugBuster® Master Mix	Merck Millipore, USA	
Butanol	Roth, Germany	71-36-3
Calcium chloride dihydrate	Roth, Germany	10035-04-8
Coomassie Brilliant Blue	Sigma Aldrich, USA	6104-59-2
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
Glycerol	Roth, Germany	56-81-5
Iron(II) sulphate heptahydrate	Roth, Germany	7782-63-0
L-isoleucine	TCI, Japan	73-32-5
Isopropyl β-D-1-thiogalactopyranosid	Roth, Germany	357-93-1
Kanamycin sulfate	Roth, Germany	25389-94-0
α-ketoglutarate	Roth, Germany	328-50-7
L-leucine	TCI, Japan	61-90-5
Magnesium chloride tetrahydrate	Sigma-Aldrich, USA	13446-34-9
Magnesium sulfate heptahydrate	Roth, Germany	10034-99-8
Manganese(II) Chloride	Sigma-Aldrich, USA	60-24-2
β-Mercaptoethanol	Sigma-Aldrich, USA	60-24-2
Ninhydrin	Roth, Germany	485-47-2
Sodium acetate	Roth, Germany	127-09-3
Sodium molybdate dihydrate	Roth, Germany	10102-40-6
Potassium acetate	Roth, Germany	127-08-2
Potassium chloride	Roth, Germany	7447-40-7
Potassium dihydrogenphosphate	Roth, Germany	7778-77-0
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	67-03-8
Tryptone	Roth, Germany	91079-40-21
Yeast extract	Roth, Germany	8013-01-2
Zinc sulfate heptahydrate	Roth, Germany	7447-20-0

Table 22: Laboratory equipment used in this study.

Labware	Model	Source
Analytical balance	Practum®	Sartorius, Germany
Precision balances	AY-6000	Sartorius, Germany
Centrifuges	5415 R	Eppendorf, Germany
	5810 R	Eppendorf, Germany
	Avanti® J-20 XP	Beckman coulter, USA
Rotor	JA-10	Beckman coulter, USA
	JA-25.50	Beckman coulter, USA
Shaking incubator	Infors HT Orbitron	Infors HT, Switzerland
	Infors Multitron Standard	Infors HT, Switzerland
Sonifier	Branson Ultrasonics™ Sonifier S-250A	Thermo Scientific, USA
Spectrophotometer	BioPhotometer 6131	Eppendorf, Germany
	NanoDrop™ 2000	Thermo Scientific, USA
pH-electrode	pH 50+ DHS®	XS Instruments, Italy
Plate reader	FLUOstar Omega	BMG Labtech, Germany
ThermoMixers®	ThermoMixer® C	Eppendorf, Germany
	ThermoMixer® comfort	Eppendorf, Germany
	HTM-2	HTA-BioTec, Germany
Vortex	Vortex-Genie 2	Scientific industries, USA

8.2 Plasmid maps

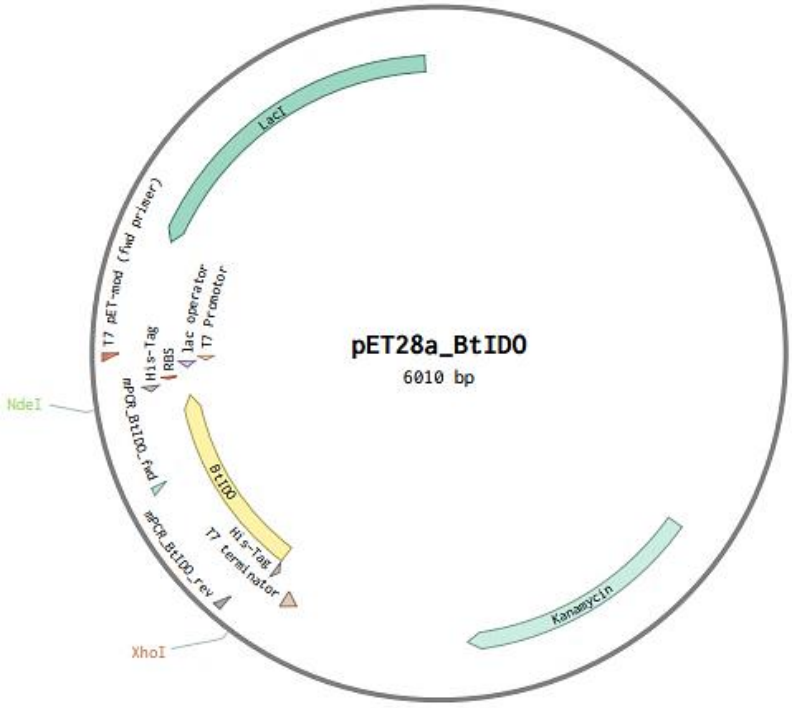

Figure 35: Ligation of the linearized vector backbone pET28(+) and the cut-out insert *BtIDO* resulted in the new vector construct pET28a_BtIDO. Restriction sites are shown in green (*Nde*I) and red (*Xho*I). Vector map of construct pET28a_BtIDO created with Benchling.

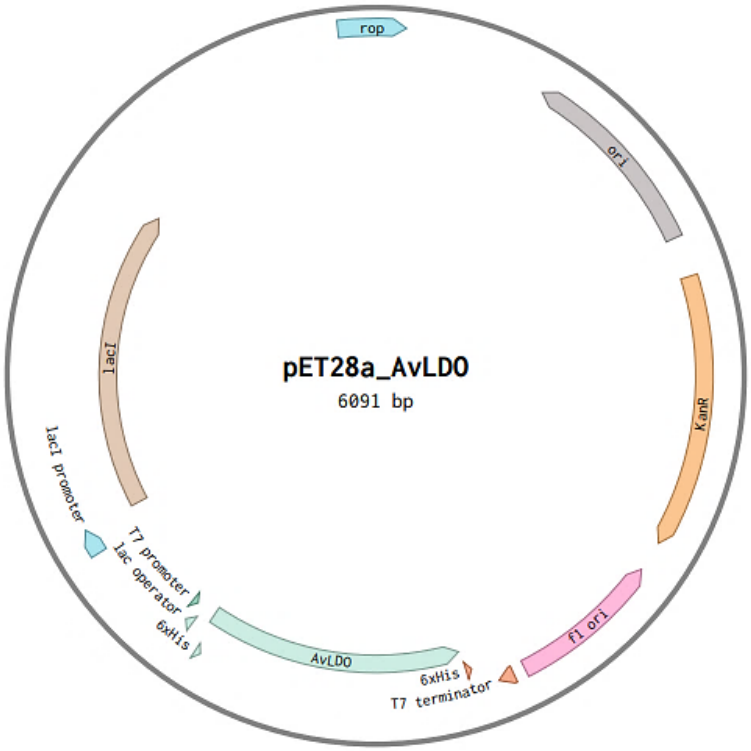

Figure 36: Vector map of construct pET28a_AvLDO created with Benchling.

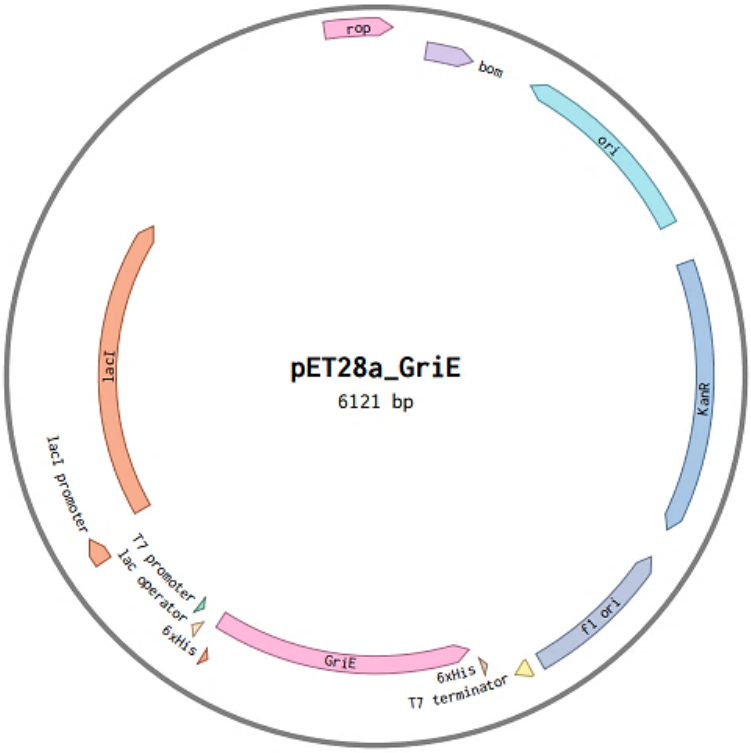

Figure 37: Vector map of pET28a_GriE created with Benchling.

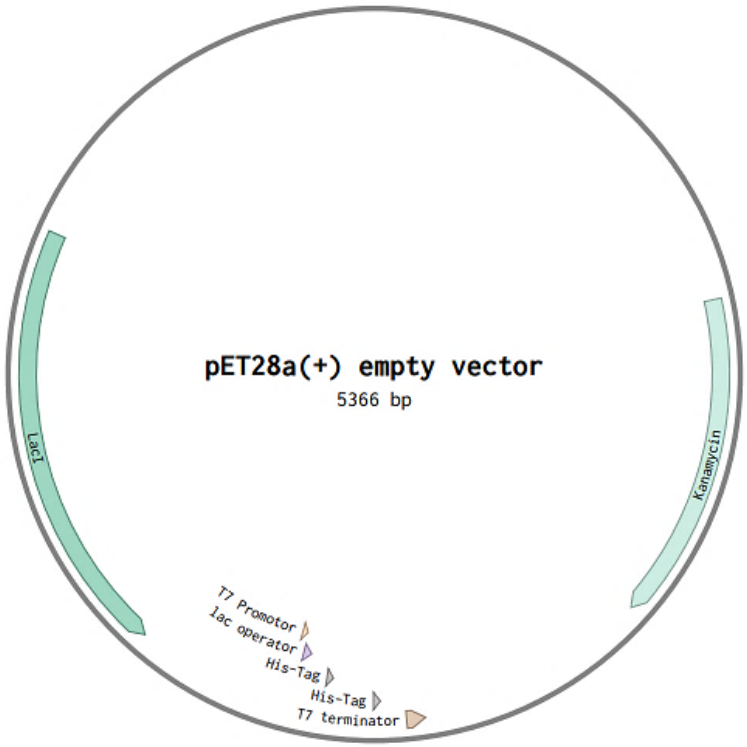

Figure 38: Vector map of the construct pET28a(+) created with Benchling.

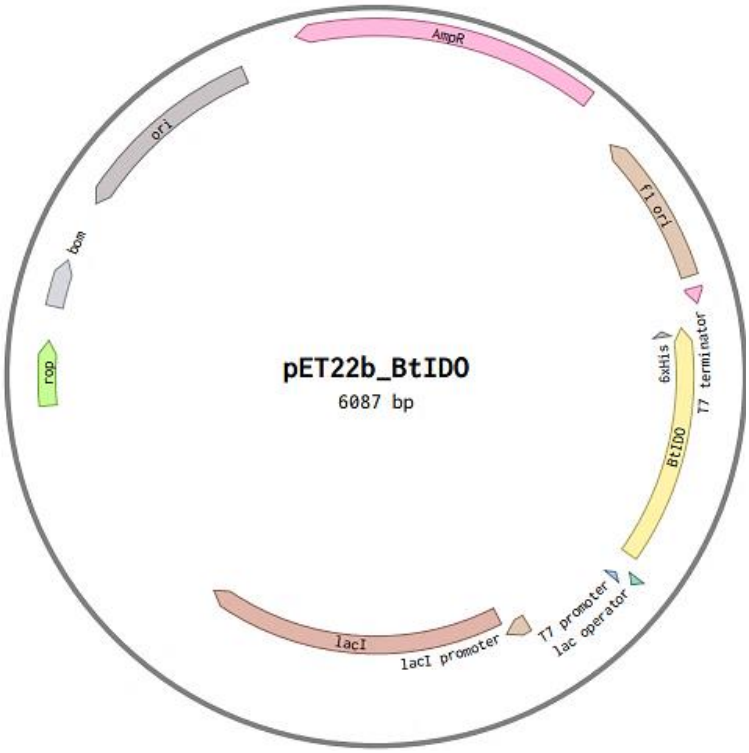

Figure 39: Vector map of construct pET22b_BtIDO created with Benchling.

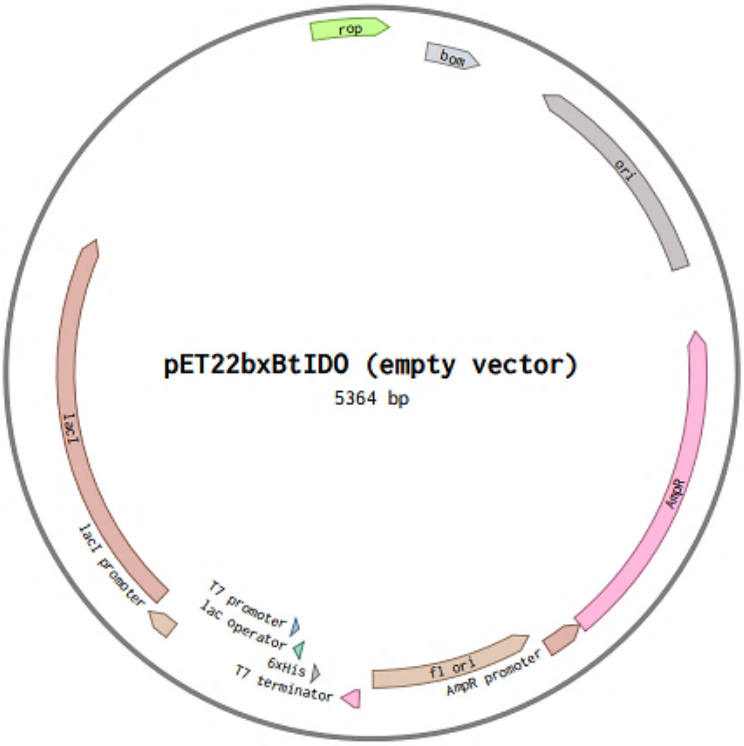

Figure 40: Vector map of pET22bxBtIDO (pET22b(+)) created with Benchling.

8.3 Sequence alignment of selected variants

Figure 41: Sequence alignment of selected variants created with Benchling. Sequences obtained from clones picked during the selection assay were aligned to the plasmid pET28a_AvLDO.

8.4 Verification of L-isoleucine conversion of selected variants

Figure 42: Identification of ascorbate as reaction component visible in all samples of the blank reaction.**[Leu]** Leucine standard 5 mM; **[B]** Blank reaction at t_1;**[Asc]** Ascorbate 10 mM.