Expression optimization of α-ketoglutarate dependent non-heme Fe(II) dioxygenases containing non-canonical amino acid in the facial triad

Smoljo, Josipa

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Food Technology and Biotechnology / Sveučilište u Zagrebu, Prehrambeno-biotehnološki fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:159:499801

UNIVERSITY OF ZAGREB
FACULTY OF FOOD TECHNOLOGY AND BIOTECHNOLOGY

GRADUATE THESIS

Zagreb, April 2021

Josipa Smoljo
1309/ MB

Expression optimization of α-ketoglutarate dependent non-heme Fe(II) dioxygenases containing non-canonical amino acid in the facial triad

Experimental work of this Graduate Thesis was done at the Institute of Molecular Biotechnology at Graz University of Technology. The Thesis was made under the guidance of prof. dr. sc. Robert Kourist and with the help of Dipl. ing. Kristin Bauer and Dr. Ivana Drienovska.

Acknowledgements

Firstly, I would like to express my deepest gratitude to prof. dr. sc. Robert Kourist for enabling this remarkable opportunity to work within his research group at the Institute of Molecular Biotechnology, Graz University of Technology and for providing everything needed for the research. Sharing valuable scientific knowledge and critical thinking about direction of the future research were one of the many opportunities in which I learned from him. Additionally, I am forever grateful to Dipl. ing. Kristin Bauer for selfless and precious help, for sharing her time and her abundant knowledge, valuable guidelines, and kindness whenever it was needed. She left a wonderful mark on my research experience. I would like to thank prof. dr. sc. Anita Slavica for being a teacher, a support and for many advice of hers. I want to thank Dr. Ivana Drienovska for everything I have learned from her. Moreover, I would like to say my truest thank you to everyone in the research group on TU Graz.

Na posljetku, neopisivo sam sam zahvalna svojim roditeljima i braći na svakoj riječi bezuvjetne podrške i ljubavi. Sretna sam i blagoslovljena što vas imam i što ste uz mene gdje god se nalazila i što god radila! Hvala mojoj široj obitelji Smoljića na svim trenucima neprocjenjive sreće i zajedništva. Hvala svim mojim prijateljima, na svakoj riječi utjehe i podrške i za svaki osmijeh koji mi obogaćuje život te ga čini vrijednim življenja.

Na kraju, hvala mome Mati što je moj mir. Hvala mu za svaki osmijeh, riječ podrške i za sve što je napravio za mene, a riječi je premalo da to opišu!

Bogu hvala!

BASIC DOCUMENTATION CARD

Graduate Thesis

University of Zagreb
Faculty of Food Technology and Biotechnology
Department of Biochemical Engineering
Laboratory of Biochemical Engineering, Industrial Microbiology and Malting and Brewing Technology

Scientific area: Biotechnical Sciences
Scientific field: Biotechnology

EXPRESSION OPTIMIZATION OF α-KETOGLUTARATE DEPENDENT NON-HEME FE(II) DIOXYGENASES CONTAINING NON-CANONICAL AMINO ACID IN THE FACIAL TRIAD

Josipa Smoljo, 1309/MB

Abstract: Expression of two L-leucine 5-hydroxylases, one isolated from *Anabaena variabilis* ATCC 29413 (AvLDO_wt) and second from *Streptomyces muensis* DSM 9461 (GriE_wt), were studied in two expression systems - pET28a(+) and pASKGATE. *In vitro* cloning methods FastCloning and Gateway® technology were performed in order to clone the genes in different vectors. Furthermore, site-directed mutagenesis QuikChange™ was done to clone the amber codon to the genetic code of the wild types at previously located sites, necessary for N-methylhistidine incorporation in the facial triad of the enzymes by the stop codon suppression methodology. The successful construction of these rationally designed enzymes was one of the rare examples of the incorporation of non-canonical amino acid in general, and N-methylhistidine specifically, in the active site of non-heme dioxygenases. Furthermore, the soluble expression of AvLDO and GriE (both wild types and mutants) was investigated using different concentrations of inducers and supplemented N-methylhistidine and successfully improved as observed by sodium dodecyl-sulfate polyacrylamide gel electrophoresis. Accordingly, the substrate scope of the GriE wild type and mutant was studied and assessed towards L-leucine and L-isoleucine by thin-layer chromatography.

Keywords: L-leucine 5-hydroxylases, AvLDO, GriE, stop codon suppression, N-methylhistidine
Thesis contains: 84 pages, 36 figures, 29 tables, 74 references, 10 supplements (appendices)
Original in: English
Graduate Thesis in printed and electronic (pdf format) version is deposited in: Library of the Faculty of Food Technology and Biotechnology, Kačićeva 23, Zagreb.

Mentor at Faculty of Food Technology and Biotechnology: P*rof. dr. sc. Anita Slavica*
Mentor at Graz University of Technology: *Prof. dr. sc. Robert Kourist*
Technical support and assistance: *Dipl. ing. Kristin Bauer*

Reviewers:
1. PhD *Renata Teparić*, Full professor, University of Zagreb
2. PhD *Anita Slavica*, Full professor, University of Zagreb
3. PhD *Robert Kourist*, Full professor, Graz University of Technology
4. PhD *Jadranka Frece*, Full professor (substitute), University of Zagreb

Thesis defended: 30th of April, 2021

TEMELJNA DOKUMENTACIJSKA KARTICA

Diplomski rad

Sveučilište u Zagrebu
Prehrambeno-biotehnološki fakultet
Zavod za biokemijsko inženjerstvo
Laboratorij za biokemijsko inženjerstvo, industrijsku mikrobiologiju i tehnologiju piva i slada

Znanstveno područje: Biotehničke znanosti
Znanstveno polje: Biotehnologija

OPTIMIZACIJA EKSPRESIJE *α*-KETOGLUTARAT-OVISNIH FE(II) HEM-NEOVISNIH DIOKSIGENAZA S NEPRIRODNOM AMINOKISELINOM U AKTIVNOM MJESTU

Josipa Smoljo, 1309/MB

Sažetak: Ekspresija dviju L-leucin 5-hidroksilaza, jedne izolirane iz *Anabaena variabilis* ATCC 29413 (naziva AvLDO_wt) te druge izolirane iz *Streptomyces muensis* DSM 9461 (GriE_wt), proučavana je u dva plazmida - pET28a (+) i pASKGATE. *In vitro* metode kloniranja, FastCloning i Gateway® tehnologija primijenjene su zbog kloniranja gena u različite vektore. Nadalje, mjesno specifična mutageneza QuikChange™ primijenjena je kako bi se klonirao stop kodon na prethodno odabrane pozicije u genima *AvLDO_wt* i *GriE_wt*, što je neophodno za ugradnju N-metilhistidina u aktivno mjesto ovih enzima pomoću metodologije stop kodon supresije. Uspješna konstrukcija ovih racionalno dizajniranih enzima jedan je od rijetkih primjera ugradnje neprirodne aminokiseline općenito, a posebno N-metilhistidina, u aktivno mjesto hem-neovisnih dioksigenaza. Nadalje, istražena je ekspresija AvLDO i GriE (divljih tipova i mutanata) u topljivoj frakciji korištenjem različitih koncentracija induktora i N-metilhistidina te je uspješno poboljšana, kako je i potvrđeno poliakrilamid gel elektroforezom u prisutnosti natrijeva dodecil-sulfata. Također, tankoslojnom kromatografijom je proučavana aktivnost GriE divljeg tipa i mutanta, prema supstratima L-leucinu i L-izoleucinu.

Ključne riječi: L-leucin 5-hidroksilaze, AvLDO, GriE, stop kodon supresija, N-methilhistidin
Rad sadrži: 84 stranica, 36 slika, 29 tablica, 74 literaturna navoda, 10 priloga
Jezik izvornika: engleski jezik
Rad je u tiskanom i elektroničkom (pdf format) obliku pohranjen u: Knjižnica Prehrambeno-biotehnološkog fakulteta, Kačićeva 23, Zagreb

Mentor pri Prehrambeno-biotehnološkom fakultetu: P*rof. dr. sc. Anita Slavica*
Mentor pri Graz University of Technology: *Prof. dr. sc. Robert Kourist*
Pomoć pri izradi: *Dipl. ing. Kristin Bauer*

Stručno povjerenstvo za ocjenu i obranu:
1. Prof. dr. sc. *Renata Teparić,* Sveučilište u Zagrebu
2. Prof. dr. sc. *Anita Slavica,* Sveučilište u Zagrebu
3. Prof. dr. sc. *Robert Kourist,* Graz University of Technology
4. Prof. dr. sc. *Jadranka Frece* (zamjena), Sveučilište u Zagrebu

Datum obrane:30. travnja 2020.

TABLE OF CONCENT

1. INTRODUCTION

α-Ketoglutarate dependent, non-heme Fe(II) dioxygenases are a large enzyme superfamily catalyzing a plethora of reactions, with hydroxylation being the most common one. These dioxygenases activate molecular oxygen for desirable oxidative transformations in which the substrate is usually hydroxylated, while a decarboxylation of α-ketoglutarate to succinate occurs. Even though diverse reactions are catalyzed by these enzymes, dioxygenases have a preserved 2-His-1-carboxylic motif that coordinates the Fe(II) atom in the active site. The real value of the reactions is the activation of the inert C-H bond, carried out by the non-heme Fe(II) center of the dioxygenases, which represents a powerful reaction in the organic synthesis chemistry (Wencel-Delord et al, 2011). Moreover, the C-H bond functionalization can provide valuable structural products, such as optically pure hydroxylated canonical and non-canonical amino acids, which are important for pharmaceutical purposes and complex molecule synthesis (Zwick and Renata, 2018; Zhu and Liu, 2017; Wu et al., 2016; Hutell, 2013; Wencel-Delord et al., 2011).

Non-canonical amino acids offer a great opportunity to enhance catalytic attributes of the enzyme, especially if incorporated in the active site of the enzyme. Changing the histidine in the facial triad of the dioxygenases possibly could widen catalytic features of the enzyme by expanding the limited repertoire of twenty naturally occurring amino acids (Drienovska et al., 2020). Having in mind that Fe(II) is crucial for the activation of dioxygen in the desired transformations carried by dioxygenases, the incorporation of N-methylhistidine (NMH) in the active triad may increase the activity of enzymes with a significant catalytic advantage (Green et al., 2016) by possibly achieving the stronger coordination of Fe(II) in the active site. In addition, the non-canonical amino acid NMH has improved electron-donative ability as a nucleophile and methylation prevents the formation of an unreactive acyl-enzyme.

The aim of this study was to incorporate non-canonical amino acid (NMH) in the facial triad of the two L-leucine 5-hydroxylases (AvLDO and GriE), belonging to the α-ketoglutarate dependent, non-heme Fe(II) dioxygenases superfamily - in order to catalyze abiological transformations, to potentially widen the substrate scope and overall investigate the potential of non-canonical amino acid in the active site of the α-ketoglutarate dependent, non-heme Fe(II) dioxygenases.

Accordingly, the aims of this Graduate Thesis were:

(a) to clone and express *AvLDO* wild type and the corresponding mutants (*AvLDO_1*, *AvLDO_2*; incorporated NMH at the locations 151 and 237, respectively) in the pASKGATE plasmid,

(b) to compare protein expression of the GriE wild type in the pET28a(+) and pASKGATE vectors in order to examine the difference, if there are any, between two different expression systems for GriE_wt expression of soluble,

(c) to optimize and improve the expression of soluble of the GriE mutant (GriE_1, H110NMH), utilizing NMH in the active triad with both pET28a(+) and pASKGATE plasmids,

(d) to determine what are the differences, if there are any, in the expressed protein concentrations, substrate scope and enzyme activity of the GriE_1 mutant in comparison to the wild type, expressed in both, pET28a(+) and pASKGATE.

To achieve the above-mentioned goals, different methods were used. The *in vivo* cloning techniques, FastCloning and Gateway®, were used to clone the two investigated hydroxylase genes into two plasmid systems. Furthermore, a site-specific mutagenesis QuikChange™ was performed to introduce the amber stop codon, which is necessary for stop codon suppression and incorporation of N-methylhistidine, in the active site of these enzymes. The expression of the wild type, as well as the mutants, was investigated and optimized using different expression protocols, different inducer and N-methylhistidine concentrations. The expression was verified and assessed using sodium dodecyl sulphate -polyacrylamide gel electrophoresis (SDS–PAGE). Purification of the protein was performed using Ni-affinity chromatography. The enzymatic activity was assessed by thin-layer chromatography.

2. LITERATURE REVIEW

2.1. ALPHA-KETOGLUTARATE DEPENDENT, NON-HEME FE(II) DIOXYGENASES

The α-ketoglutarate dependent, non-heme Fe (II) dioxygenases (further referred to as dioxygenases) are an enzyme superfamily, which are, by coupling the oxidative decarboxylation of an α-ketoglutarate, able to oxidatively transform a prime substrate. Hydroxylation reactions are primarily prominent, however, demethylation, ring formations, ring expansions, halogenations, and desaturations are also catalyzed by these dioxygenases (Herr and Hausinger, 2018; Islam et al., 2018; Wu et al, 2016; Hutell, 2013; Hoffart et al., 2006). The importance of this superfamily is affirmed by its role in nature. These enzymes are found to be involved in the repair or modification of nucleic acids and in the control of transcriptional and translational protein biosynthesis. Additionally, they have a role in oxygen sensing, fatty acid metabolism and secondary metabolite biosynthesis, as well as in modifications related to epigenetic regulation (Peters and Buller, 2019; Herr and Hausinger, 2018; Islam et al., 2018; Hausinger and Schofield, 2015; Martinez and Hausinger, 2015). Furthermore, the aforementioned rich oxidative potential makes dioxygenases interesting for industrial application (Peters and Buller, 2019).

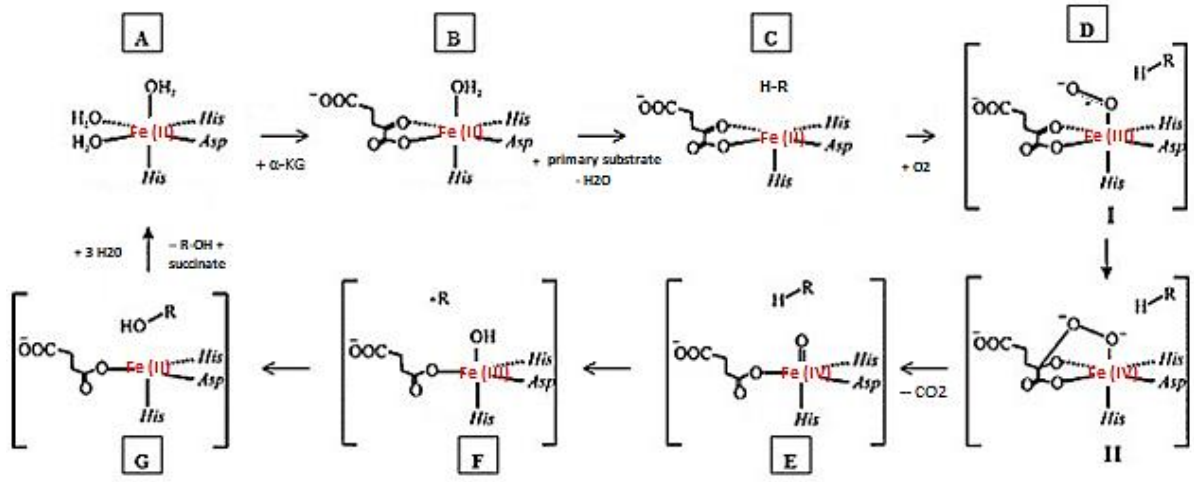


Figure 1. Consensus catalytic mechanism of oxygen activation by α-ketoglutarate dependent, non-heme Fe(II) dioxygenases. [A] Enzyme resting state, 2-His-1-carboxylic motif has three water molecules bound. [B] α-ketoglutarate binds and displaces two water molecules. [C] Substrate displaces last water molecule but does not bind to the iron. [D] Binding of the molecular oxygen. [E] Oxidative decarboxylation of an α-ketoglutarate yielding the Fe(IV)-oxo ferryl. [F] The ferryl species abstracts a hydrogen atom from the primary substrate producing the Fe(III)-OH and a substrate radical. [G] Rebinding of the substrate radical with the hydroxide, resulting in the hydroxylated substrate and formation of Fe(II). Hydroxylated substrate and succinate dissociate and [A] Resting state is regenerated (adapted from Krebs et al., 2007).

A structural characteristic of these dioxygenases is the conserved 2-His-1-carboxylic motif which enables the coordination of the Fe (II) atom in the active site (Figure 1 [A], Hegg and Que, 1997). With chelation of α-ketoglutarate, with its C1 carboxylate and C2 ketone in a bidentate configuration, at the metallocentre, two of three water molecules, are being replaced (Martinez and Hausinger, 2015; Rose et al., 2011; Hausinger, 2004, Figure 1 [B]). Introduction of the primer substrate to the active site, not binding to the metal ion directly, displaces the last molecule of water (Figure 1 [C]), triggering the binding of molecular oxygen (Figure 1 [D]). The binding of the oxygen consequently leads to the formation of the Fe (III) superoxo species (Martinez and Hausinger 2015). Oxidative decarboxylation of α-ketoglutarate is initiated by one of the Fe(III)-bound oxygen atom attack to the C2 of the co-substrate, and the reaction results in the release of CO_2 and the formation of succinate, while the iron oxidizes further to Fe(IV) (Figure 1 [E]). Additionally, it was highlighted in the literature that this ferryl species could be the intermediate for all the previously mentioned reactions which α-ketoglutarate dependent, non-heme Fe(II) dioxygenases catalyze (Herr and Hausinger 2018; Islam et al., 2018; Krebs et al., 2007). The Fe(IV)-oxo species abstracts a hydrogen atom from the substrate, generating Fe(III)-OH and a substrate radical (Figure 1 [F], Wu et al 2016, Martinez and Hausinger, 2015, Rose et al., 2011). Commonly, a rebound of the hydroxide from the Fe(III)-OH complex with the substrate radical, produces the hydroxylated product and formation of the Fe(II) alkoxo species. (Figure 1 [G]). Finally, it has been proved that this oxygen atom incorporated into the hydroxylated substrate is derived from the dioxygen (Bollinger et al., 2005). Furthermore, product dissociation, followed by succinate dissociation, and the rebinding of three water molecules, results in the regeneration of the enzyme resting state ((Figure 1 [A]).

Imperative in the reactions catalyzed by dioxygenases is the activation of the inert C-H bond. Functionalization of a C-H bond is an important challenge and powerful tool in synthetic organic chemistry because of its known low activity, low regio- and stereoselectivity (Frey et al., 2019; Zhu and Liu, 2017; Wu et al., 2016; Hutell, 2013; Wencel-Delord et al., 2011). Although, dioxygenases hold the desired specificity in regio- and stereoselectivity, which is highly important for industrial applications, industrial utilization of these enzymes is limited (Hutell, 2013). The limitation lies in a few reasons - firstly most of the industrially interesting dioxygenases have only recently been discovered and some dioxygenases are yet to be sequenced and characterized. For instance, the L-leucine 5-hydroxylase AvLDO was only characterized in 2018 (Correira Cordeiro et al., 2018). Secondly, reaction conditions are limited by the stability of the dioxygenases which can undergo oxidation or loss of the central iron and precipitation of the enzyme (Hutell, 2013). Until this date,

dioxygenases found their role in industry mainly on the hydroxylation of amino acids (proline hydroxylation and hydroxy isoleucine production) which are useful for chiral pharmaceuticals and chiral antibiotics (Chen et al., 2017; Hibi and Ogawa, 2014) and in the conversion of antibiotic penicillin G to a cephalosporin (Frey et al., 2019; Peters and Buller, 2019; Correira Cordeiro et al., 2018; Hutell, 2013). With gaining more knowledge about different dioxygenases, α-ketoglutarate dependent, non-heme Fe (II) dioxygenases have a promising future in the industry because of their diversity and beneficial property to activate the C-H bond. Implementation of asymmetric hydroxylation is especially needed in the chemical synthesis where a request for chiral compounds with defined stereochemical properties is highlighted. Certainly, dioxygenases can be used for up-scale catalysis of hydroxy aliphatic amino acids (Peters and Buller, 2019; Hibi and Ogawa, 2014; Hutell, 2013).

2.1.1. L-leucine 5-hydroxylases

As already mentioned, besides biocatalysis of secondary metabolites, amino acid hydroxylation is the most frequent and valuable reaction for industrial applications, catalyzed by α-ketoglutarate dependent, non-heme Fe(II) dioxygenases (Frey et al., 2019; Peters and Buller, 2019; Correira Cordeiro et al., 2018; Hutell, 2013). In this Graduate Thesis two highly stereoselective L- leucine hydroxylases were studied, namely AvLDO and GriE.

2.1.1.1. AvLDO

AvLDO is an L-leucine 5-hydroxylase isolated from *Anabaena variabilis* firstly characterized in 2018 (Correira Cordeiro et al., 2018). The gene is located on a gene cluster involved in the biosynthesis of (2S,4S)-4-methylproline, which is an important building block for bioactive peptides found in cyanobacteria (Correira Cordeiro et al., 2018). AvLDO has the characteristic 2-His-1-carboxylic motif of α-ketoglutarate dependent, non-heme Fe (II) dioxygenases and is formed by His151, Asp153 and His237.

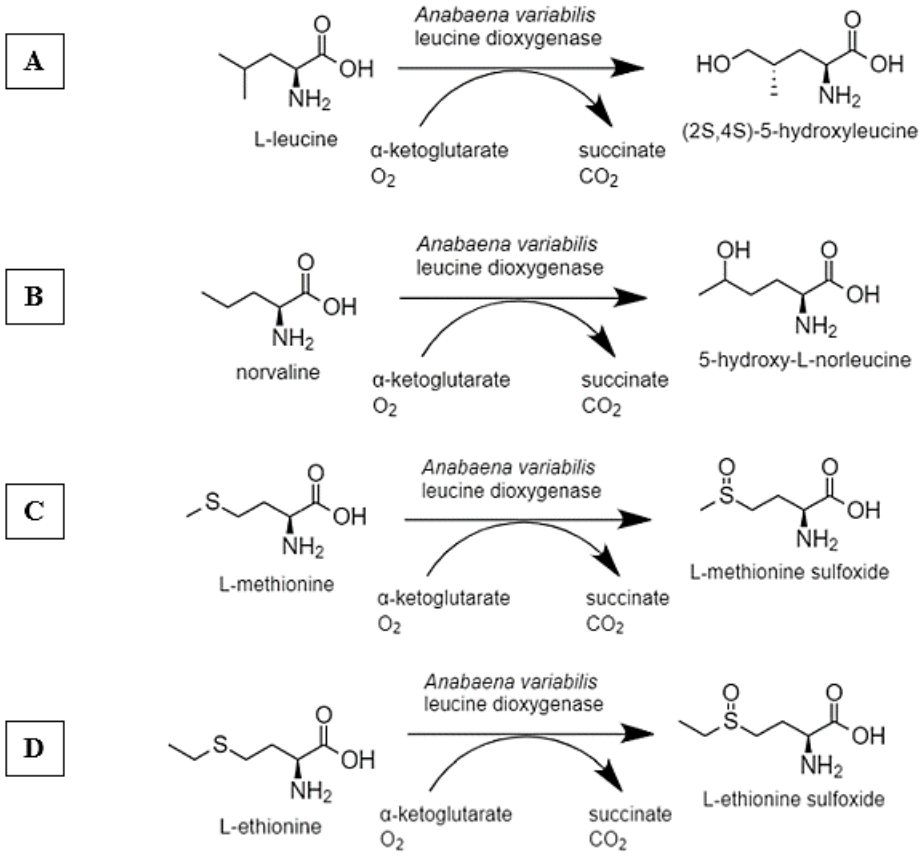


Figure 2. Examples of oxidative transformations catalyzed by AvLDO (adapted from Correira Corderio et al., 2018). [A] Conversion of L-leucine to (2S,4S)- 5-hydroxyleucine. [B] Conversion of L-norvaline to 5-hydroxy-L-norleucine. [C] Conversion of L-methionine to L-methionine sulfoxide. [D] Sulfoxidation of L-ethionine to L-ethionine sulfoxide.

Correira Corderio et al. (2018) investigated the substrate scope of AvLDO and have concluded that the enzyme strictly recognizes specific aliphatic side residues of L-amino acids in the active site. AvLDO catalyzes the hydroxylation of L-leucine, hydroxylation of L-norleucine, sulfoxidation of L-methionine and sulfoxidation of L-ethionine (Figure 2). L-isoleucine cannot be converted by AvLDO, which indicates that the regio- and stereoselective hydroxylation occurs in the C5-position (Correira Corderio et al., 2018).

2.1.1.2. GriE

Another L-leucine 5-hydroxylase investigated in this study is GriE. The *GriE* gene is a part of the griselimycin's biosynthetic gene cluster. The GriE enzyme is responsible for the hydroxylation of L-leucine into (2S,4R)-5-hydroxyleucine as the first step in the (2S,4R)-4-methyl-proline formation, which is a building block for the rare alkaloid manzacidin C and cavinafungin B in the griselimycin biosynthesis (Figure 3; Peters and Buller, 2019; Lukat et al., 2017). Cyclic peptide griselimycin was

isolated from *Streptomyces* sp. which showed antibacterial and antimicrobial activity. Additionally, Kling et al. (2015) as well as Holzgrabe in 2015 outlined that griselimicyn is highly active against *Mycobacterium tuberculosis*, both *in vitro* and *in vivo*.

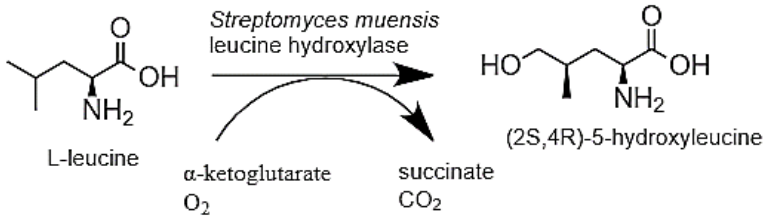


Figure 3. GriE catalyzes the hydroxylation of L-leucine to (2S,4R)-5-hydroxyleucine which further results in the production of (2S,4R)-4-methyl-proline, building block of the rare alkaloid manzacidin C and cavinafungin B synthesis (adapted from Lukat et al., 2017).

GriE belongs to the α-ketoglutarate-dependent, non-heme Fe (II) dioxygenase superfamily. Structure wise, it possesses the typical 2-His-1-carboxylic motif, and the canonical iron-coordinating fold is formed by His110, Asp112 and His210 (Lukat et al., 2017).

GriE hydroxylates the δ position of aliphatic amino acids and as investigated by Zwick and Renata in 2018, the enzyme shows remarkable catalytic efficiency and broader substrate scope relative to other dioxygenases. Among determined substrates is L-leucine, hydroxylated with complete diastereoselectivity and high (> 7,500) total turnover number. Furthermore, no type of conversion or activity was observed with D-leucine, L-valine, or L-isoleucine, while conversion of L-allo-isoleucine was catalyzed with a moderate total turnover number and 13 % yield, suggesting that the enzyme active site is strict towards differences at the β position of the substrate. Furthermore, activity towards γ-methyl-L-leucine and 4-hydroxy-L-leucine resulted in 54 % and 52 % yield, respectively, demonstrating enzyme flexibility regarding the γ position. In addition, γ-methyl-L-leucine was converted by a much higher total turnover number compared to the 4-hydroxy-L-leucine which hydroxyl group deactivates oxidation of the C-H bond. L-norleucine and 2-aminoheptanoic acid were hydroxylated to a yield of 80 % - 90 % and with moderate turnover number, again suggesting that the substitutions at distal positions are tolerated by the active site of GriE (Zwick and Renata, 2018).

2.2. NON-CANONICAL AMINO ACIDS

In order to potentially improve enzyme characteristics and widen the substrate scope, the non-canonical amino acid, N-methylhisitidine, was introduced at the active site of the aforementioned L- leucine 5-hydroxylases - GriE and AvLDO.

The use of non-canonical amino acid could overcome restrictions given by the limited number of natural amino acids, which could result in a greater novel site-specific manipulation and protein design versatility (Burke et al., 2019; Neumann-Staubitz and Neumann, 2016). Non-canonical amino acids are a powerful tool to potentially develop biorthogonal functions of the artificial enzyme in a fine-tuning manner, especially if incorporated in the active site (Drienovska and Roelfes, 2020; Neumann-Staubitz and Neumann, 2016; Wang et al., 2001). The exclusive embedding of these non-canonical amino acids, which are not genetically encoded by the organism itself, can be a useful tool for protein development and catalytic feature enhancement. For instance, improvement of thermostability (Deepankumar et al., 2014) and enhancement of structural stability compared to their natural counterparts (Ohtake et al., 2015; Ravikumar et al., 2015) were successfully achieved.

The ability to develop novel enzymes is highly beneficial for the generation of compounds needed for organic synthesis. Furthermore, non-canonical amino acids can allow immobilization of the enzymes with strictly specific and high efficiency, useful for bioprocess application. Representable case is incorporation of non-canonical tyrosine analogs such as 4-acetylphenylalanine, O-benzyl tyrosine, and 4-cyanophenylalanine into cytochrome P450, achieved by Ravikumar et al. in 2015. Additionally, non-canonical amino acids can be used, among others, to study enantioselective catalysis, protein behavior, dynamics, structure-function relationship, as well as serve for large-scale industry implementation including biopharmaceutical drug development. All the above is enabled through the agency of exquisite control over the protein structure (Drienovska and Roelfes, 2020; Saleh et al., 2019; Young and Schultz, 2018; Zou et al., 2018; Budisa et al., 2017; Neumann and Neumann, 2016; Drienovska et al., 2015; Chin et al., 2014; Wang et al., 2001). Technology and application of site-specific incorporated non-canonical amino acids grows rapidly, and the most appealing potential is the possibility of manipulating with protein structure and protein function, *in vivo* (Neumann and Neumann, 2016; Chatterjee et al., 2013).

2.2.1. Translation

Understanding the process of translation is the key to comprehend synthesis of the enzymes containing non-canonical amino acids. During translation, genetic information formed as messenger RNA (mRNA) dictates protein synthesis by codon - to - amino acid correspondence. Aminoacyl-tRNA synthetase is an enzyme that attaches a specific amino acid to the analogous transport RNA (tRNA), catalyzing the transesterification to the tRNA 3′-terminal adenosine via the amino acid carboxylic group. The mRNA codon interacts with the corresponding anticodon on the aminoacylated tRNA. Once the tRNA is charged, the ribosome can transfer the amino acid from the tRNA, which contains the anticodon complementary to the codon present on the mRNA, onto a growing peptide (Wang et al., 2012; Ambrogelly and Soll, 2007; Link and Tirrell, 2005; Ibba and Soll, 2000,).

The mRNA sequence carries the information for the amino acid sequence- three nucleotides can make a combination of sixty-four triplets. Three of these code for stop codons, which signal the termination of translation by the recruitment of one of two release factors. Sixty-one triplets, named "sense" codons encode for twenty natural amino acids. Twenty amino acids, alongside with rare but naturally occurring pyrrolysine and selenocysteine (shown in Figure 4), have a very limited range, especially if compared to the synthetic chemistry spectrum, as far as functional and structural diversity goes (Drienovska and Roelfes, 2020; Budisa et al., 2017; Fekner and Chan, 2011; Ibba and Soll, 2004).

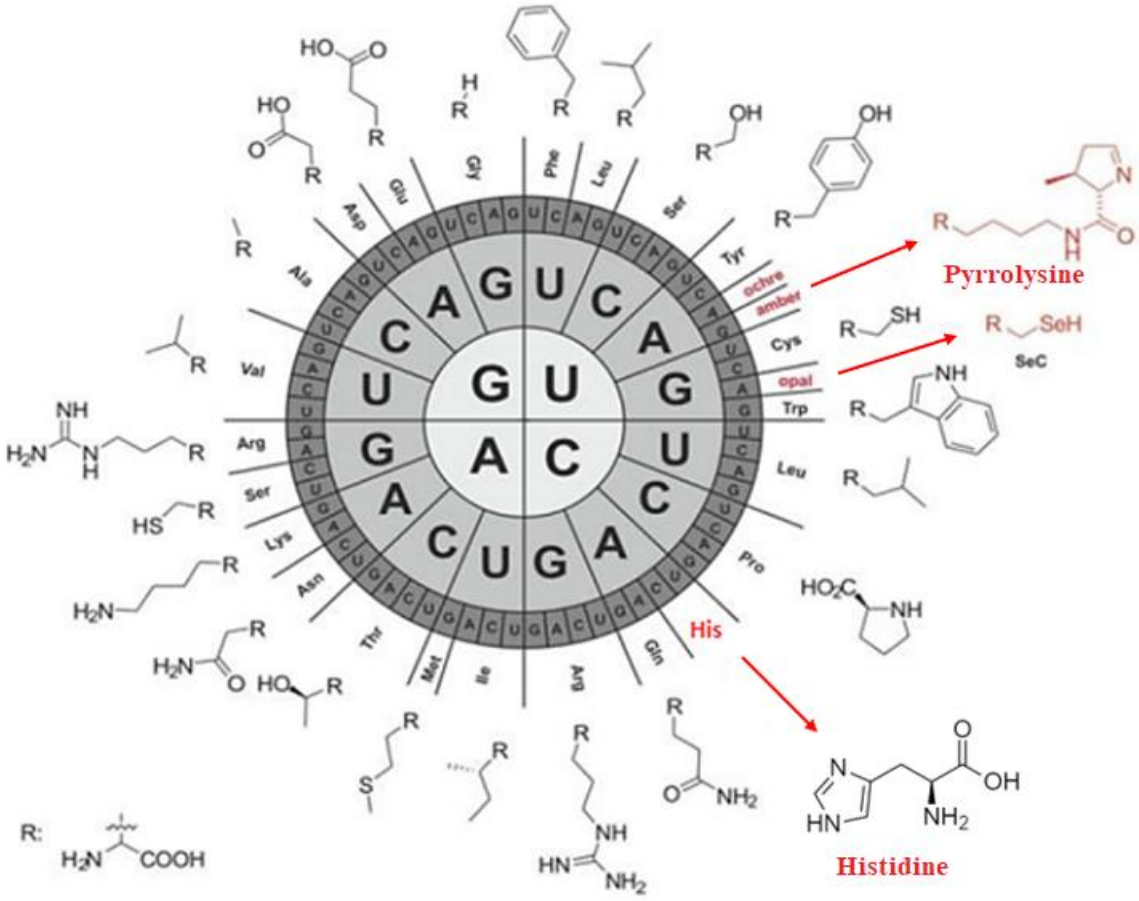


Figure 4. The genetic code chart (codons should be read from inside to outside). mRNA triplet (5'-3') corresponds to one of the 20 canonical amino acids or a stop codon. Rare but proteinogenic amino acids selenocystein (SeC) and pyrrolysine are also included in the chart. Incorporation of selenocystein is encoded by the opal stop codon (UGA) and pyrrolysine by the amber stop codon (UAG) (adapted from Agostini et al., 2017).

Most species use twenty aminoacyl-tRNA synthetases, one for each amino acid. Nevertheless, some species have developed additional aminoacyl tRNA synthetase for the incorporation of rare proteinogenic amino acids such as pyrrolysine. For example, in some *Methanosarcina* strains pyrrolysine is incorporated by the mRNA codon - pyrrolysl-tRNA anticodon correlation (Drienovska and Roelfes, 2020; Wals and Ovaa, 2014; Wang et al., 2012). These aminoacyl-tRNA synthetase / tRNA pairs naturally occur in these organisms and by transferring corresponding pairs into *E. coli*, successful introduction of non-canonical amino acids into the protein sequences can be achieved (Drienovska et al., 2018; Green et al., 2016; Wals and Ovaa, 2014).

2.2.1.1. Stop codon suppression

The methodology used for the specific incorporation of non-canonical amino acid in the protein sequences is the stop codon suppression. The method expands the genetic code by manipulating the natural translation process (Drienovska and Roelfes, 2018; Hofmann et al., 2018). With this approach, the stop codon can be reassigned to act as a new sense codon for site-specific incorporation of the non-canonical amino acid. The amber stop codon is used for the stop codon suppression as it shows the rarest use in *E. coli* when comparing it to the other stop codons (~ 7 %, Drienovska and Roelfes, 2020; Wals and Ovaa, 2014).

To incorporate the non-canonical amino acid at a pre-determined site by reading the stop codon, the aminoacyl-tRNA synthetase / tRNA pair evolved for translation of rare proteinogenic amino acids (such as pyrrolysyl-tRNA synthetase / pyrrolysyl-tRNA) and orthogonal to the host system needs to be introduced in the biosynthetic cell machinery (Figure 5).

Additionally, to successfully perform stop codon suppression (as shown in Figure 5), certain conditions must be met. The orthogonality of the aminoacyl-tRNA / tRNA pair ensures translation fidelity and incorporation of the correct non-canonical amino acid. It is necessary that the orthogonal aminoacyl-tRNA synthetase specifically achieves the exclusively linkage of the non-canonical amino acid, and not of the any canonical amino acids, to its cognate orthogonal tRNA, and not to any endogenous ones. In other words, the aminoacyl-tRNA synthetase must not aminoacylate any canonical amino acid to the orthogonal tRNA and it must not aminoacylate non-canonical amino acid to the endogenous tRNA. The orthogonal tRNA should not be recognized and aminoacylated by any endogenous synthetases but solely by the orthogonal synthetase. Furthermore, the amber stop codon needs to be recognized as an incorporation point for the non-canonical amino acid. The last factor which needs to be taken into account, is the trespassing of the cell membrane by the non-canonical amino acid. Only with an efficient transport, the non-canonical amino acid can be integrated (Drienovska and Roelfes, 2020; Hoffmann et al., 2018; Chin, 2017; Budisa et al., 2017; Neumann-Staubitz and Neumann, 2016; Chin, 2014; Wals and Ovaa, 2014; Wang et al., 2012; Link and Tirrell, 2005; Mehl et al., 2003; Chin et al., 2002; Wang et al., 2001).

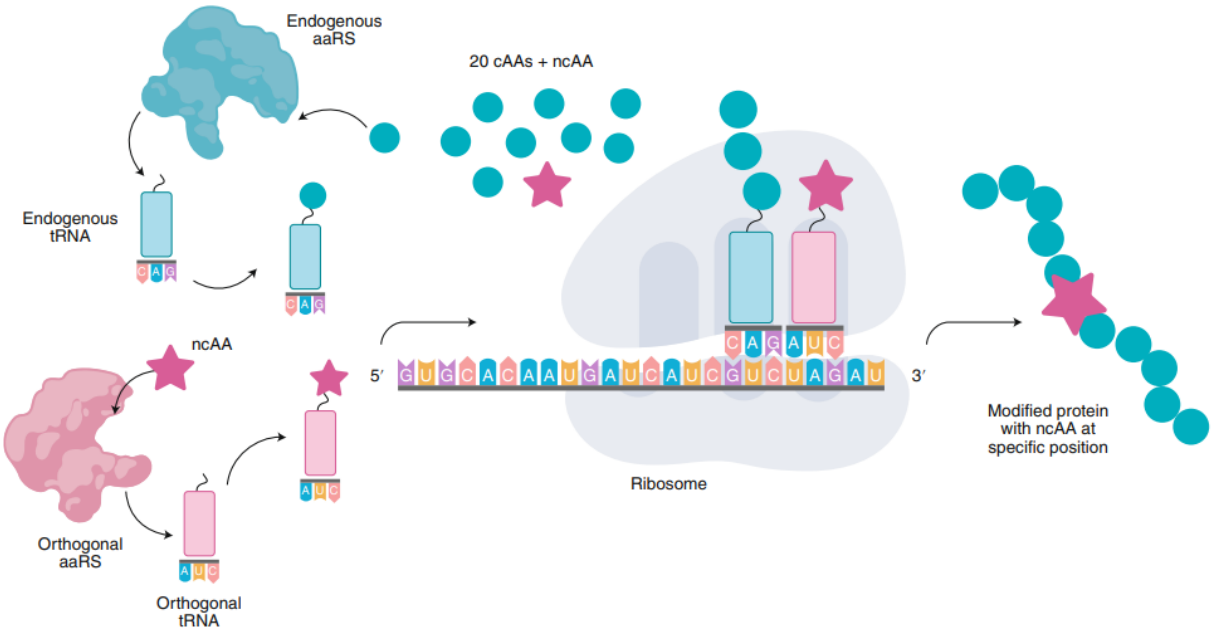


Figure 5. Use of stop codon suppression for the site-specific incorporation of non-canonical amino acids using an orthogonal aminoacyl-tRNA synthetase and transfer RNA (tRNA) pair (with permission, taken from Drienovska and Roelfes, 2020).

However, stop codon suppression has some disadvantages such as the competition between the release factor and the incorporation of the non-canonical amino. It bears the risk of termination of translation, resulting in inactive protein fragments, which further impact the expression yield of the target protein and the cell's viability (Drienovska et al., 2020; Budisa et al., 2017; Wang et al., 2012). Despite all of that, the method has already been applied for the incorporation of non-canonical amino acids in various enzymes (see Chapter 2.2. for examples).

2.2.2. N-methylhistidine

In this study, stop codon suppression was used to incorporate a non-canonical amino acid. The facial triad histidine, of the dioxygenases GriE and AvLDO, has been exchanged with the non-canonical analogue N-methylhistidine (NMH). The reason for investigating the NMH in place of the histidine lies in histidine's role in the active triad. Two histidines (Figure 6 [A]) are part of the characteristic 2-His-1-carboxylic motif that coordinates Fe (II) in the active side of dioxygenases (see Chapter 2.1.). Histidine is one of the most interesting and versatile catalytic residues due to its ability to serve both as an acid and a base. Additionally, as a result of containing two nitrogen atoms in the imidazole ring, it can participate in multiple interactions making it the most frequently encountered catalytic residue (Green et al., 2016).

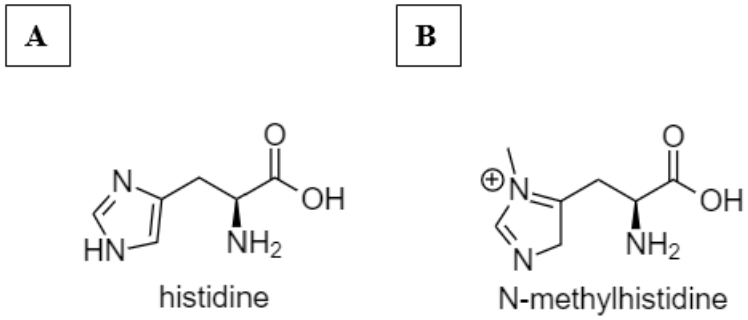


Figure 6. Comparison of [A] histidine and [B] N-methylhistidine.

Replacement of the N-δ hydrogen with an alkyl residue has several structural and functional consequences which directly affect histidine's relationships and bonding in the facial triad (Green et al., 2016). The replacement of histidine with NMH (Figure 6 [B]) may lead to a more stable enzyme and result in increased enzymatic activity. The methylation prevents the formation of the unreactive acyl-enzyme. Additionally, it has a catalytic advantage due to a better electron-donative ability as a non-canonical nucleophile. Overcoming the equivoque imidazole charge state, simplifies the construction of catalytic mechanisms, by NMH, enables a fixed single tautomeric state and secures the neutral charge of the imidazole ring (Burke et al., 2019; Budisa et al., 2017; Green et al., 2016).

Green et al. (2016) documented the incorporation of NMH in ascorbate peroxidase. The engineered enzyme showed a considerably higher turnover number (31,300 given by moles of product formed per mole of enzyme) compared to the wild type (6,200) with comparable catalytic efficiency. Additionally, the characterization of several variants containing NMH brought a further understanding of the active triad of the heme enzymes (Green et al., 2016). Pott et al. (2018) showed that containing NMH in the structure of myoglobin caused changes in the proximal pocket resulting in an increase of the heme redox potential and wider enzyme promiscuity. Furthermore, Burke et al. (2019) concluded that NMH has a great opportunity to impact the enzyme engineering being a valuable nucleophilic catalyst for chemical.

2.3. MOLECULAR CLONING METHOD

2.3.1. Gateway® technology

The Gateway® technology is an *in vitro* cloning method used in this study to clone L-leucine 5-hydroxylase gene *AvLDO* in pASKGATE vector. The Gateway® technology enables the fast transfer of genes between entry plasmid (silent carrier, pENTRE) and any needed destination vector for different research purposes and downstream applications (Katzen, 2007). Primarily, mentioned entry vector acts as an intermediate that enables integration of the insert gene into a variety of destinations vectors by using the Gateway® platform, meaning that the entry vector generated in this study (pENTRE_His_Leu_AvLDO_wt) has an added value as it is useful for the gene insertion and production of any destination vector and not solely designed for the pASKGATE plasmid.

The Gateway® reaction is carried out by the recombination of the attachment sites. In particular, the LR reaction applied from Gateway® technology took place between attL and attR sites and was catalyzed by the LR Clonase™ enzyme mix (Thermo Scientific, USA). As a result, expression vector was generated, with the gene of interest flanked by attB sites. Additionally, there was a double control system which enabled higher efficiency of destination vector generation - different antibiotic resistance between entry and expression vector and the toxin gene (*ccdB*) from destination vector which ended up incorporated in the toxin byproduct (Figure 7).

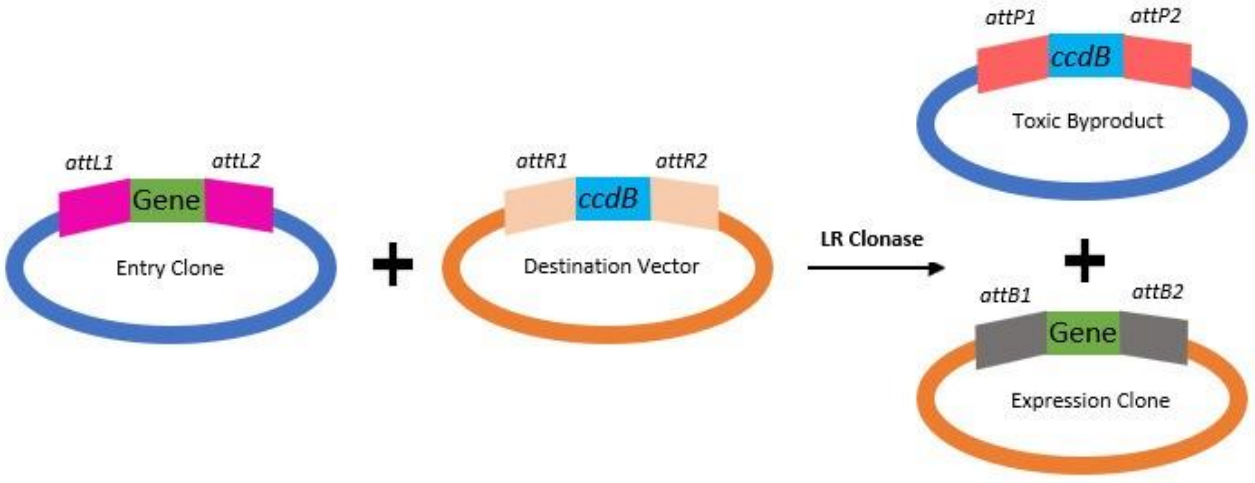


Figure 7. Scheme of LR reaction applied from Gateway® technology. Integration and excision LR recombinant reaction takes place in between attachment sites (*att; attL1, attL2 and attR1, attR2*) that are bordering the genes- in the entry clone that is the gene of the interest and in the destination vector, *ccdB* toxin gene. Recombination results in expression clone of interest and toxic byproduct (adapted from Anonymous 1, 2017).

3. MATERIALS AND METHODS

3.1. MATERIALS

3.1.1. L-leucine 5-hydroxylases

L-leucine 5-hydroxylases used in this study are listed in Table 1.
L-leucine 5-hydroxylase isolated from *Anabaena variabilis* AvLDO_wt and L-leucine 5-hydroxylase isolated from *Streptomyces muensis* GriE_wt, as well as corresponding mutants (AvLDO_1, AvLDO_2 and GriE_1, GriE_2, respectively) which contain N-methylhistidine (NMH), were studied. NMH was incorporated at the previously selected sites for both dioxygenases (locations H151NMH and H237NMH for AvLDO mutants, and H110NMH and H210NMH for GriE mutants).

In order to increase the expression of soluble L-leucine 5-dioxygenases (AvLDO and GriE wild type and AvLDO and GriE mutants which contain non-canonical amino acid NMH), were studied in two plasmids - pET28a(+) and pASKGATE. The former contains a T7 lac promoter, which is frequently used for high-level expression experiments (for plasmid maps see Appendices 2). The latter contains an anhydrotetracycline inducible promoter, which also has been shown to lead to high expression levels (Bertram and Hillenthe, 2008). Additionally, the pASKGATE vector is equipped with the Gateway® cloning system (see Chapter 2.3.1. and Appendices 3). All expression constructs used in this study are listed in Table 3.

Table 1. Overview of dioxygenases used in this study

Enzyme	**Abbreviation**	**Accession number**	**Kindly provided by**	**Name of the mutant**	**Position of N-methylhistidine incorporation**	**Literature**
L-leucine 5-hydroxylase from *Anabaena variabilis* ATCC 29413	AvLDO	GenBank Accession no.: ABA21237	Ruhr Universitat Bochum (Germany)	AvLDO_1	151	Correia Cordeiro et al., 2018
				AvLDO_2	237	
L-leucine 5-hydroxylase from *Streptomyces muensis* DS 9461	GriE	Protein Accession number: AKC91859.1	GenScript Biotech B.V. (Netherlands)	GriE_1	110	Lukat et al., 2017
				GriE_2	210	

3.1.2. Plasmids

In Table 2 an overview of all plasmids used in this study is listed.

Table 2. Overview of the plasmids used in this study

Plasmid	**Antibiotic resistance (abbreviation)**	**Source of the plasmid**	**Position of His-tag**	**Usage of the plasmid**
pEVOL_NMH	Chloramphenicol (Chl)	Kindly provided by Hilvert (Green et al., 2016)	N-terminus	Plasmid carries the gene encoding for tRNA synthetase / tRNA pair necessary for NMH incorporation
pET28a(+)	Kanamycin (Kan)	Institute for Molecular Biotechnology collection	N-terminus and C-terminus	Carries the gene encoding for L-leucine 5-hydroxylases: *AvLDO* or *GriE* (both wild type and mutants listed in Table 3; for plasmid map see Appendices 2)
pASKGATE	Spectinomycin (Sm)	Institute for Molecular Biotechnology collection	N-terminus	Carries the gene encoding for L-leucine 5-hydroxylases, *AvLDO* or *GriE* (both wild type and mutants listed in Table 3). Plasmid is used as the expression vector of Gateway® system (see Chapter 2.3.1. and for plasmid map see Appendices 3)
pENTRE	Kanamycin (Kan)	pENTREGFP2 was a gift from Nathan Lawson (Addgene plasmid #22450; RRID: Addgene_22450)	N-terminus	Entry vector for the Gateway® system – intermediate for pASKGATE construct

3.1.3. Expression constructs

Table 3 shows all expression constructs needed for this study. Genes encoding for L-leucine 5-hydroxylase (*AvLDO* and *GriE*) carried by different vectors are listed below.

Table 3. List of expression constructs used in this study

Lable of the constructs	**Gene**	**Wild type or mutant (location of amber stop codon necessary for NMH incorporation by stop codon suppression)**	**Plasmid backbone**	**Status of the plasmid (already available at the Institute / cloning experiments necessary to be done in this study)**
pENTRE_His_Leu_AvLDO_wt	*AvLDO_wt*	Wild type	pENTRE	This study
pENTRE_His_Leu_GriE_wt	*GriE_wt*	Wild type	pENTRE	Already available at the Institute
pASKGATE_His_Leu_AvLDO_wt	*AvLDO_wt*	Wild type	pASKGATE	This study
pASKGATE_His_Leu_AvLDO_1	*AvLDO_1*	Mutant 1- Amber stop codon at position 151	pASKGATE	This study
pASKGATE_His_Leu_AvLDO_2	*AvLDO_2*	Mutant 2-Amber stop codon at position 237	pASKGATE	This study
pET28a(+)_GriE_wt	*GriE_wt*	Wild type	pET28a(+)	Already available at the Institute
pET28a(+)_GriE_1	*GriE_1*	Mutant 1- Amber stop codon at position 110	pET28a(+)	Already available at the Institute
pET28a(+)_GriE_2	*GriE_2*	Mutant 2- Amber stop codon at position 210	pET28a(+)	Already available at the Institute
pASKGATE_His_Leu_GriE_wt	*GriE_wt*	Wild type	pASKGATE	Already available at the Institute
pASKGATE_His_Leu_GriE _1	*GriE_1*	Mutant 1- Amber stop codon at position 110	pASKGATE	Already available at the Institute
pASKGATE_His_Leu_GriE_2	*GriE_2*	Mutant 2- Amber stop codon at position 210	pASKGATE	This study

3.1.4. Bacterial strains

The bacterial strains used in this study are listed in Table 4.

Table 4. Overview of *E. coli* strains used in this study

Strain	**Genotype**	**Usage**
E. coli TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-	High-efficiency cloning
E. coli BL21 (DE3)	B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)	Transformation and high-level protein expression
E. coli BL21 (DE3) with pEVOL_NMH	B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS), additionally contains plasmid encoding for orthogonal tRNA synthetase	Strain used for expression experiments of the mutants (where NMH incorporation was required (Table 3))

3.1.5. Primers

Primers used for cloning experiments (see Chapters 3.2.1., 3.1.2. and 3.2.4.) or Sanger Sequencing by Microsynth AG (see Chapter 3.2.9.2.) are listed in the Table 5.

Table 5. List of primers used in this study

Name	**5'-3' Sequence**	**Purpose**
P30-pENTRE-AvlDO-fw	CTTTGTACAAAAAAGCAGGCTTGATGACTGCTAC TTCTCAAC	FastCloning (see Chapter 3.2.1.) and corresponding cPCR (see Chapter 3.2.1.1.)
P31-pENTRE-AvlDO-rv	CTTTGTACAAGAAAGCTGGGTTTAAGACCACAA AATAATTTC	FastCloning (see Chapter 3.2.1.) and corresponding cPCR (see Chapter 3.2.1.1.)
P5_pASK_Seq_rv	CGCAGTAGCGGTAAACG	cPCR of the Gateway® reaction (see Chapter 3.2.3.2.)
P20_pASKGATE_Seq_fw	TTTGCTCACATGACCCGAC	cPCR of the Gateway® reaction (see Chapter 3.2.3.2.)
AvLDO-His1-fw	GCAGGGTACTCAACTTTAGGTTGATTA	QuikChange™ (see Chapter 3.2.4.)
AvLDO-His1-rv	GGAGGATAATCAACCTAAAGTTGAGTA	QuikChange™ (see Chapter 3.2.4.)
AvLDO-His2-fw	ACGCGCAACTATTAGATTGTTGAGCCG	QuikChange™ (see Chapter 3.2.4.)
AvLDO-His2-rv	CGGCTCAACAATCTAATAGTTGCGCGT	QuikChange™ (see Chapter 3.2.4.)
GriE-His1-fw	GATTCCTGGGCCTGGTAGCAGGAC	QuikChange™ (see Chapter 3.2.4.)
GriE-His1-rev	GACGATGAAGTCCTGCTACCAGGCCCA	QuikChange™ (see Chapter 3.2.4.)
GriE-His2-fw	CATCCGGAGATCATCTAGGGATCGGCT	QuikChange™ (see Chapter 3.2.4.)
GriE-His2-rv	GTTGGGAGCCGATCCCTAGATGATCTC	QuikChange™ (see Chapter 3.2.4.)

Name	**5’-3’ Sequence**	**Purpose**
P12_pASKGATE_fw	CACCATCACATCGAAGGGCGCACAAGTTTGTACAAAAAAGCTG	Sanger Sequencing (see Chapter 3.2.9.2.)
P13_pASKGATE_rv	GAATTCGGGACCGCGGTCTCGACCACTTTGTACAAGAAAGCTG	Sanger Sequencing (see Chapter 3.2.9.2.)
P14_pENTRE_fw	ACCCAGCTTTCTTGTACAAAG	Sanger Sequencing (see Chapter 3.2.9.2.)
P15_pENTRE_rv	AGCCTGCTTTTTTGTACAAAG	Sanger Sequencing (see Chapter 3.2.9.2.)
P16_pENTRE_GriE_rv	GTACAAGAAAGCTGGGTTCATGCCAGCCTCGATTCGG	Sanger Sequencing (see Chapter 3.2.9.2.)
P19_pASKGATE_BbvCI_rv	GGACCGCGGTCTCGGCTGAGGACCACTTTGTACAAGAAAGCTG	Sanger Sequencing (see Chapter 3.2.9.2.)
P23_pENTRE_Leu_GriE_fw	CTTTGTACAAAAAAGCAGGCTTGATGCAGCTCACGGCCGATC	Sanger Sequencing (see Chapter 3.2.9.2.)
P24_GriE_Hismutant-1	GATTCCTGGGCCTGGTAGCAGGA	Sanger Sequencing (see Chapter 3.2.9.2.)
P25_GriE_Hismutant-2	CATCCGGAGATCATCTAGGGATCGGCT	Sanger Sequencing (see Chapter 3.2.9.2.)

3.1.6. Chemicals and labware

All chemicals and labware used in this study are listed in the Table 6 and Table 7, respectively.

Table 6. List of chemicals, suppliers and CAS numbers used in this study

Chemicals	Supplier	CAS number
10x Pfu Buffer	Promega, USA	/
1-Butanol	Roth, Germany	71-36-3
5x HF Phusion Buffer	Thermo Scientific, USA	/
Acetic Acid	Roth, Germany	64-19-7
Acetone	Roth, Germany	67-64-1
Acetonitrile	VWR International, USA	75-05-8
Agar-Agar	Roth, Germany	9002-18-0
Anti-foam SE-15	Sigma-Aldrich, USA	/
Ascorbate	Roth, Germany	50-81-7
BugBuster®	Merck Milllipore, USA	/
Chloramphenicol	Roth, Germany	56-75-7
Coomasie Brilliant Blue	Sigma-Aldrich, USA	6104-59-2
di-Kaliumhydrogenphosphat	Roth, Germany	7778-77-0
Dipotassiumhydrogen phosphate	Roth, Germany	7758-11-4.
Disodiumhydrogen phosphate	Roth, Germany	7558-79-4
DMSO	Sigma Aldrich, USA	67-68-5
dNTPs (100 mM each)	Thermo Scientific, USA	/
dpn1 (10 U/µL)	Thermo Scientific, USA	/
EDTA	Sigma Aldrich, USA	60-00-4
Ethanol	Chem-Lab, Belgium	64-17-5
Fe(III)-sulfat Heptahydrat	Roth, Germany	7782-63-0
Gateway® LR Clonase™ enzyme mix	Thermo Scientific, USA	/
Glycerol	Roth, Germany	56-81-5
H-His(3-Me)-OH	BACHEM, Switzerland	368-16-1
HS-Taq Mix	Biozym, Germany	CAT: 331126
Imidazole	Roth, Germany	288-32-4
Isopropyl β-D-1-thiogalactopyranosid	Roth, Germany	367-93-1
Kaliumdihydrogenphosphat	Roth, Germany	7778-77-0
Kanamycin sulfate	Roth, Germany	25389-94-0
Kanamycinsulfat	Roth, Germany	25389-94-0
L(+)-Ascorbin acid	Roth, Germany	134-03-2
LB-Medium (Lennox)	Roth, Germany	73049-73-7
L-Isoleucine	SERVA, Germany	73-32-5
L-Leucine	TCI, Japan	61-90-5
Natriumchlorid	Roth, Germany	7647-14-5
Ninhydrin	Roth, Germany	485-47-2

Chemicals	Supplier	CAS number
PfuPlus! Polymerase	EURX, Poland	/
Phusion Polymerase (2 U/µL)	New England Biolabs, USA	/
Potassium dihydrogenphosphate	Roth, Germany	7778-77-0
Sodium hydroxide	VWR International, USA	1310-73-2
Spectinomycin Dihydrochloride 5-hydrate	AppliChem, Germany	22189-32-8
TRIS	Roth, Germany	1185-53-1
Tryptone/Peptone ex casein	Roth, Germany	91079-40-2
Yeast extract	Roth, Germany	8013-01-2
α-Ketoglutaric acid	Roth, Germany	328-50-7
β-mercaptoenthanol	Sigma-Aldrich, USA	60-24-2

Table 7. List of labware used in this study

Labware	Model	Manufacturer
Shaking incubator		Infors HT, Switzerland
Centrifuges	5415 R	Eppendorf, Germany
	5810 R	Eppendorf, Germany
	Avanti® J-20 XP	Beckman coulter, USA
Rotors	JA-25.50	Beckman coulter, USA
	JA-10	Beckman coulter, USA
ThermoMixers®	ThermoMixer® C	Eppendorf, Germany
	ThermoMixer® comfort	Eppendorf, Germany
	HTM 2	HTA-BioTec, Germany
Analytical balance	Practum®	Sartorius, Germany
Precision balances	AY-6000	Sartorius, Germany
pH-electrode	inoLab® pH 720	VWR International, USA
Vortex	Vortex-Genie 2	Scientific industries, USA
Spectrophotometer	BioPhotometer 6131	Eppendorf, Germany
	NanoDrop™ 2000	Thermo Scientific, USA
	Plate reader	BioTek Instruments, USA
Sonifier	Branson Ultrasonics™ Sonifier S-250A	Thermo Scientific, USA
SDS page machine	PowerEase500 Invitrogen	Thermo Scientific, USA
Electroporator	Bio-Rad MicroPulser electroporator	Bio-Rad, USA

3.2. METHODS

In this Graduate Thesis used cloning technics were FastCloning, Gateway® technology and QuikChange™ site directed mutagenesis.

3.2.1. Construction of pENTRE_His_Leu_AvLDO_wt using FastCloning

Considering that *AvLDO_wt* was not available within the pASKGATE plasmid, the entry vector (pENTRE_His_Leu_AvLDO_wt which was necessary for Gateway® technology and cloning of pASKGATE_His_Leu_AvLDO_wt product) had to be cloned beforehand. The FastCloning technique was used for the generation of that construct (Li et al., 2011). FastCloning methodology implies PCR amplifications of the vector and insert - done separately, *dpn1* digestion of the amplified product pre-confirmed on agarose gel, and transformation of the amplified insert and linearized vector backbone directly into chemo competent *E. coli* Top 10. All these steps are described in the next subchapters.

3.2.1.1. PCR of insert amplification (AvLDO_wt) necessary for FastCloning technology

Components for the PCR were carefully mixed (Table 8) and the PCR program (
Table 9) was adjusted corresponding to the melting temperature of the used primers (Table 5); calculated with the online Tm Calculator provided by ThermoFisher Scientific (Anonymous 2, 2021). The amplification time was approximated to the size of the desired product using the manufacturer's recommendation with 1 kb / 30 sec for the Pfu polymerase (Anonymous 3, 2021). PCR reactions for the insert and backbone amplification, needed for entry vector (pENTRE_His_Leu_AvLDO_wt) generation, were done separately. Insert amplification was required and done in this study, while backbone pENTREGFP2 have already been linearized and amplified at the Institute (Table 2).

Table 8. PCR components of insert *AvLDO_wt* amplification (from pET28a(+)_AvLDO_wt) needed for generation of pENTRE_His_Leu_AvLDO_wt construct

Component	**Volume / mass**
5 x HF Phusion Buffer	20 µL
dNTPS (10 mM)	2 µL
P30-pENTRE-AvlDO-fw (10 µM)	5 µL
P31-pENTRE-AvlDO-rv (10 µM)	5 µL
Template DNA (pET28a(+)_AvLDO_wt)	20 ng
Phusion Polymerase (Pfu polymerase)	1 µL
ddH_2O	to 100 µL

Table 9. PCR program of insert *AvLDO_wt* amplification (from pET28a(+)_AvLDO_wt) needed for generation of pENTRE_His_Leu_AvLDO_wt construct

Temperature	**Time [min]**	**Cycles**
98 °C	00:30	1 x
98 °C	00:10	30 x
53.4 °C	00:15	
72 °C	00:30	
72 °C	5:00	1 x
4 °C	∞	1 x

3.2.1.2. Agarose gel electrophoresis

In order to confirm the PCR amplification (described above), a 1 % agarose gel [with 0.8 g agarose and 80 mL of 1x TAE Buffer (Tris-acetate-EDTA buffer; Table 10)] containing GelGreen (20,000 x) was prepared. Samples were prepared by mixing 5 µL of PCR sample with 1 µL of 6x Loading Dye (Biolabs New England, USA). As a standard, GeneRuler™ 1 kb DNA Ladder, was used (Thermo Fischer Scientific, USA). The gel was run for 40 min at 120 V and 400 mA.

Table 10. Components of TAE buffer

Component	**Concentration**
Tris base	2 M
Acetic acid	1 M
EDTA disodium salt hydrate	50 mM
ddH_2O	to 1 L

3.2.1.3. Dpn1 digestion

When the insert PCR amplification was confirmed as successful with an agarose gel electrophoresis, 1 µL of *Dpn1* methylation-sensitive restriction enzyme was added to the amplified insert and was incubated for 2.5 h at 37 °C, followed by heat inactivation of *Dpn1* at 80 °C for 20 min.

After *Dnp1* digestion, *E. coli* TOP10 cells were transformed with a mixture of 1 µL linearized vector backbone (already linearized at the Institute) and 4 µL of amplified insert (described above, see Chapter 3.2.1.) according to the general protocol for transformation of chemically competent *E. coli* TOP 10 strain described under Chapter 3.2.6. After transformation colony PCR for pENTRE_His_Leu_AvLDO_wt construct determination was done (see Chapter 3.2.3.1.).

3.2.2. Construction of pASKGATE_His_Leu_AvLDO_wt using the Gateway® cloning system

In order to perform the Gateway® reaction necessary for pASKGATE_His_Leu_AvLDO_wt construction, a mixture of 150 ng entry vector (pENTRE_His_Leu_AvLDO_wt created by FastCloning, see Chapter 3.2.1.) and 150 ng of the destination vector (pASKGATE; already available at the Institute) was prepared and if necessary, filled up to 8 µL with TE buffer pH 8 (Thermo Scientific, USA). 2 µL of LR Clonase™ enzyme mix was added to the mixture and incubated for 1 h at 25 °C. Afterwards 1 µL of Proteinase K was added to the mixture and incubated 10 min at 37 °C.

After the last incubation, 1 µL of the Gateway® reaction was used for the transformation of chemically competent *E. coli* TOP 10 cells (see Chapter 3.2.6.). After incubation of transformed *E. coli* TOP 10 cells (overnight on 37 °C), transformants were assessed with cPCR (see Chapter 3.2.3.2.).

3.2.3. Colony PCR

3.2.3.1. Colony PCR (cPCR) of pENTRE_His_Leu_AvLDO_wt construct

Colony PCR was performed to assess transformants on the correctly assembled pENTRE_His_Leu_AvLDO_wt plasmid cloned by FastCloning reaction (for plasmid map see Appendices 5). Colonies were picked from LB-agar plate with a sterile toothpick, suspended in 30 µL ddH_2O. The suspended cells were incubated for 10 min at 95 °C in ddH_2O and cell debris was obtained by spun down using a microcentrifuge (5415 R, Eppendorf; Germany), on maximal speed for 2 min. 5 µL of the supernatant was used as a template for the cPCR. Used PCR mixture is listed in Table 11 and PCR program is shown in Table 12. The amplification time for the cPCR was approximated to the size of the desired product using the manufacturer's recommendation of 1 kb/1 min for the New England Biolabs' Hot Start Taq 2X Master Mix (Anonymous 4, 2021). cPCR was confirmed by agarose gel electrophoresis as described in Chapter 3.2.1.2.

Table 11. PCR components for the cPCR of pENTRE_His_Leu_AvLDO_wt construct

Component	**Volume**
2X HS-Taq-PCR mix	7.5 µL
Template	5 µL
P30-pENTRE-AvlDO-fw (5 µM)	1.25 µL
P31-pENTRE-AvlDO-rv (5 µM)	1.25 µL
ddH_2O	to 15 µL

Table 12. PCR program for the cPCR of pENTRE_His_Leu_AvLDO_wt construct

Temperature	**Time [min]**	**Cycles**
95 °C	1:00	1 x
95 °C	0:15	30 x
64.2 °C	0:15	30 x
72 °C	1 min	
72 °C	5:00	1 x
4 °C	∞	1 x

3.2.3.2. cPCR of pASKGATE_His_Leu_AvLDO_wt construct

Colony PCR for pASKGATE_His_Leu_AvLDO_wt product confirmation cloned by Gateway® reaction (for plasmid map see Appendices 6) was performed as described in the previous subchapter 3.2.2. The PCR mixture was the same as the one listed in Table 11 except the used primers were P5_pASK_Seq_rv and P20_pASKGATE_Seq_fw (Table 5). The PCR program was adjusted and is shown in Table 13. Amplified PCR product was confirmed by agarose gel electrophoresis as described in Chapter 3.2.1.2.

Table 13. PCR program for the cPCR of pASKGATE_His_Leu_AvLDO_wt construct

Temperature	**Time [min]**	**Cycles**
94 °C	10:00	1x
94 °C	00:15	30x
62 °C	00:15	
68 °C	1:00	
68 °C	5:00	1x
4 °C	∞	1x

3.2.4. Site directed mutagenesis using QuikChange™

QuikChange™ is a site directed mutagenesis technique, which was used to change the codon for histidine in the facial triad of the dioxygenases to the amber stop codon. For *AvLDO* mutants these were at the positions 151 (H151NMH, *AvLDO_1*) and 237 (H237NMH, *AvLDO_2*); for *GriE* mutants at the positions 110 (H110NMH, *GriE_1*) and 210 (H210NMH, *GriE_2*). The introduction of the amber stop codon in the constructs transformed in the *E. coli* BL21 (DE3) pEVOL_NMH cells leads to a stop codon suppression and the subsequent incorporation of the non-canonical amino acid NMH.

The general reaction mixture for the QuikChange™ experiment can be seen in Table 14 and the corresponding PCR program in

Table 15. The amplification time was approximated to the size of the desired product using the manufacturer's recommendation with an amplification time of 1 kb/30 s for the PfuPlus! DNA polymerase (Anonymous 3, 2021). Substitution was created by incorporating desired nucleotide(s) in the center of the primers used in the reaction. Primers used for production of pASKGATE_His_Leu_AvLDO_1 were AvLDO-His1-fw and AvLDO-His1-rev, for production

of mutant pASKGATE_His_Leu_AvLDO_2 were used AvLDO-His2-fw and AvLDO-His2-rv primers and primers used in QuikChange™ reaction for pASKGATE_His_Leu_GriE_2 were GriE-His2-fw and GriE-His2-rv (Table 5).

Table 14. Components for QuikChange™ experiments to introduce the amber stop codon at the native facial triad histidine positions of *AvLDO_wt* and *GriE_wt* to produce corresponding mutants

Components	**Volume / concentration / mass (V_{total}=14.5 µL)**
PfuPlus! DNA Polymerase	1 µL
10 x Pfu reaction buffer	5 µL
Primer fw	10 µM
Primer rv	10 µM
2mM dNTPs	1 µL
DMSO	1 µL
Template (for *AvLDO* mutants: pASKGATE_His_Leu_AvLDO_wt, for *GriE* mutants: pASKGATE_His_Leu_GriE_wt)	10 ng

Table 15. Program for QuikChange™ experiments to introduce the amber stop codon at the native facial triad histidine positions of *AvLDO_wt* and *GriE_wt* to produce corresponding mutants

Temperature	**Time (min)**	**Cycles**
95 °C	0:30	1 x
95 °C	0:20	25 x
55 °C – 62 °C	0:45	
72 °C	8:30	
72 °C	10:00	1 x

Melting temperature of QuikChange™ reaction for production of mutant *AvLDO_1* was 55 °C, for *AvLDO_2* was 60 °C and for reaction for production of *GriE_2*, 60 °C.

3.2.5. Preparation of chemically competent *E. coli* BL21 (DE3) cells harboring pEVOL_NMH for overexpression of the mutants where NMH incorporation was required

An over-night culture of *E. coli* BL21 (DE3) harboring the pEVOL_NMH plasmid was prepared and incubated according to Chapter 3.2.8. 100 mL of LB medium was inoculated with 5 mL of the

ONC and incubated at 37 °C and 200 rpm. Once an optical density (OD_{600}) of culture reached 0.4 - 0.5 (measured on BioPhotometer 6131, Eppendorf, Germany), cells were harvested by centrifugation on 5810 R centrifuge (Eppendorf, Germany; 4,000 g, 4 °C, 15 min). All the following materials and solutions were pre-cooled. The pellet was suspended in 30 mL of sterile TfB1 buffer (Table 16) and in 3.2 mL of autoclaved 1 M $MgCl_2$, followed by an incubation of 15 min on ice. After centrifugation of pellet suspended in TfB1 buffer and $MgCl_2$ (4,000 g, 4 °C, 10 min) the supernatant was discarded. The pellet was than suspended in 4 mL of autoclaved TfB2 buffer (

Table 17) and incubated for 15 min on ice. 50 µL aliquots (of *E. coli* BL21 (DE3) harboring pEVOL_NMH pellet suspended inTfB2 buffer) were shock frozen in liquid nitrogen and stored at -80 °C.

Table 16. Components of TfB1 buffer used for preparation of chemically competent *E. coli* BL21 (DE3) cells

Component	**Mass / volume (V_{total}=500 mL)**
KCH_3COO	1.47225 g
KCl	3.7275 g
$CaCl_2$	0.55493 g
Glycerol	65 mL (13 %)

Table 17. Components of TfB2 buffer used for preparation of chemically competent *E. coli* BL21 (DE3) cells

Component	**Mass / volume (V_{total}=500 mL)**
KCl	0.37275 g
$CaCl_2$	5.5493 g
MOPS	1.04635 g
Glycerol	65 mL (13 %)

3.2.6. Transformation of (chemically and electro) competent *E. coli* TOP 10, *E. coli* BL21(DE3) and *E. coli* BL21(DE3) pEVOL_NMH strains

Transformation of *E. coli* TOP 10 was done for *in vivo* cloning experiments (with pENTRE and pASKGATE constructs), while transformation of *E. coli* BL21 (DE3) (or *E. coli* BL21 (DE3) pEVOL_NMH needed for mutants, prepared as described above) was necessary for expression purposes of the expression constructs used in this study: pASKAGATE_His_Leu_AvLDO_wt,

pASKAGATE_His_Leu_AvLDO_1, pASKAGATE_His_Leu_AvLDO_2, pET28a(+)_GriE_wt, pET28a(+)_GriE_1, pASKAGATE_His_Leu_GriE_wt and pASKAGATE_His_Leu_GriE_1. Transformation protocol for both chemically transformation and electroporation are the same for all three *E. coli* strains (protocol is shown for *E. coli* BL21(DE3) strain) and listed below.

3.2.6.1. Transformation of chemically competent E. coli strains

To the 50 µL of chemically competent *E. coli* BL21 (DE3), plasmid (50 ng) desired for overexpression was added (constructs listed in Table 3. List of expression constructs used in this study). The mixture was incubated for 30 min on ice. A heat shock was performed at 42 °C for 42 s followed by 2 min on ice. Afterwards 400 µL of SOC media was added for regeneration of cells (Table 18). The cells were regenerated for 60 - 90 min at 37 °C and 350 rpm. 50 µL of the transformation was plated on the agar plates (Table 19) containing corresponding antibiotics (40 µg mL^{-1} of Kan). The plates were incubated overnight at 37 °C.

Table 18. Components of SOC medium

Component	**Mass / volume**
SOC Broth (Lennox, Roth, Germany)	26.64 g
ddH_2O	up to 1 L

Table 19. Composition of LB – agar medium

Component	**Mass / volume**
LB powder (Lennox, Roth, Germany)	20 g
Agar-agar	15 g
ddH_2O	up to 1 L

3.2.6.2. Desalting of dpn1 digested PCR products following with electroporation of E. coli strains

Transformation of chemically competent *E. coli* stains was preferable compared to electroporation, because in case of transformation of chemically competent strains, desalting of *dpn1* digested PCR products (in this case - amplified insert and/or linearized backbone) was not necessary. If electroporation of the *E. coli* BL21 (DE3) strain was performed with the *dnp1* digested PCR products (amplified insert and/or linearized backbone), desalting of PCR products beforehand was mandatory. Additionally, transformation of electro- and chemically competent

E. coli strains did not show any difference in transformation efficiency, so both approaches could be used.

For desalting, a standard Petri dish was half-filled with deionized Milli-Q water. A nitrocellulose MF-Millipore Membrane filter (pore size 0.025 mm, diameter 13 mm, Bartelt, Austria) was placed on top of the water surface. An aliquot of the sample (above mentioned *dpn1* digested PCR products) was pipetted onto the membrane and left to dialyze for 1 h. After dialysis, the desalted sample was collected from the filter by pipetting.

Electrocompetent cells (50 µL aliquot) were thawed on ice and 100 ng of DNA of interest (in a volume of approximately 5 µL) was added to the *E. coli* TOP 10 cells. The mixture was carefully transferred to a pre-cooled electroporation cuvette without introducing bubbles and by covering the whole bottom of the cuvette. Electroporation was performed using a Bio-Rad MicroPulser electroporator (Bio-Rad, USA) with the following settings: 2.5 kV, 100 Ω, and 25 µF. After the electroporation 500 µl of SOC medium (Table 18) was immediately added to the cuvette. Electroporated cells were regenerated at 37 °C and 450 rpm for 1 h. After microcentrifugation (5415 R), on maximal speed for 1 min, part of the supernatant was removed and cells were re-suspended in the remaining supernatant volume and streaked out on LB-agar plates [Table 19, plates were pre-incubated on room temperature (RT)] containing the corresponding antibiotic for the selection of positive transformants (Kan 40 µg mL^{-1}). Incubation of the transformed cells plated on the LB-agar plates was done at 37 °C for 16-20 h.

3.2.7. Glycerol stocks

Glycerol cryopreservation stocks were prepared by mixing 1 mL of over-night culture of *E. coli* BL21 (DE3) (strains with/without pEVOL_NMH) transformed by plasmid desirable for overexpression in this study (Table 3) with 1 mL sterile glycerol (60%) in cryotubes. Afterwards, the cells were stored at -80 °C until further use.

3.2.8. Over-night cultures

Over-night cultures (ONCs) of *E. coli* BL21 (DE3) (strains with/without pEVOL_NMH) transformed by all plasmids desirable for overexpression in this study (Table 3) were prepared by supplementing approximately 15-20 mL of Lysogeny broth **(**LB medium, Table 20) with the corresponding antibiotics (Table 21). Inoculation of ONC was performed with a single colony or

25 µL of glycerol stock (see above) and cultures of transformed *E. coli* BL21 (DE3) were incubated overnight at 37 °C and 120 rpm.

Table 20. Composition of LB medium

Component	**Mass / volume**
LB powder (Lennox, Roth, Germany)	20 g
ddH_2O	to 1 L

Table 21. Plasmids, antibiotic resistances, and corresponding concentrations of used antibiotics

Plasmid	**Antibiotic**	**Concentration**
pET28a(+)	Kan	40 µg mL^{-1}
pASKGATE	Sm	50 µg mL^{-1}
pEVOL_NMH	Chl	50 µg mL^{-1}

3.2.9. Plasmid analysis

3.2.9.1. Plasmid isolation

Plasmid isolation (for all plasmids used in this study; Table 2 and Table 3) was performed with the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific, USA) according to the manufacturer's protocol (Anonymous 5, 2021).

3.2.9.2. DNA sequencing

For DNA sequence verification of the isolated plasmids (see above), Sanger Sequencing was performed by Microsynth AG (Balgach, Switzerland). Primers used for sequencing are listed in Table 5.

3.2.10. Cultivation of transformed *E. coli* BL21 (DE3) or *E. coli* BL21 (DE3) pEVOL_NMH cells

3.2.10.1. Inducers

Different concentrations of isopropyl *β*-D-1-thiogalactopyranoside (IPTG, Table 22) were used for the overexpression of the dioxygenases using the pET28a(+) plasmid. Anhydrotetracycline (ATET, Table 23) was used for the induction of the *tetA* promoter of the pASKGATE plasmids

and arabinose (ara; Table 22, Table 23) for the induction of pEVOL_NMH encoding for tRNA synthetase / tRNA pair for N-methylhistidine incorporation.

3.2.10.2. Non-canonical amino acid

N-methylhistidine (NMH, supplemented in concentrations as listed in Table 22, Table 23) was added at the same time as arabinose, before overexpression of the enzymes, in order to prepare environment for tRNA synthetase / tRNA for incorporation of N-methylhistidine.

3.2.10.3. Overexpression of AvLDO wild type and AvLDO mutants in pASKGATE

Heterologous expression was investigated in *E. coli* BL21 (DE3) for AvLDO_wt or in *E. coli* BL21 (DE3) pEVOL_NMH strain for AvLDO mutants (AvLDO_1, AvLDO_2), with the pASKGATE plasmid carrying the genes of the recombinant protein. Due to low expression of soluble levels of AvLDO protein in general, expression of soluble of AvLDO_wt and AvLDO mutants, induced by different inducer concentrations, was investigated. Expression was performed based on the conditions stated by Correia Cordeiro et al. (2018) and Hanreich (2019). Furthermore, Bertram and Hillenthe in 2008 declared that commonly applied concentration of ATET for inducement experiments is 200 ng mL^{-1}. For the study of heterologous overexpression of AvLDO_wt, AvLDO_1 and AvLDO_2 the following concentrations were investigated: 50 µg L^{-1}, 100 µg L^{-1}, 150 µg L^{-1} and 200 µg L^{-1} ATET.

Expression of AvLDO_wt and mutants in pASKGATE was done in a large-scale cultivation, for which 400 mL of LB-medium was supplemented with the corresponding antibiotic (50 µg mL^{-1} of Sm for wild type; 50 µg mL^{-1} of Sm and 50 µg mL^{-1} of Chl for mutants; Table 21). Furthermore 2.5 % v/v anti-foam SE-15 emulsion (Sigma-Aldrich, USA) was added to the medium and LB medium was inoculated with an ONC to an OD_{600} of 0.1 and incubated at 37 °C and 120 rpm until OD_{600} was not high enough for overexpression.

Overexpression was performed: (a) by induction with different concentrations of ATET; when an OD_{600} of 0.6-0.8 was reached, for AvLDO wild type, and (b) for the AvLDO mutants; firstly at OD_{600} of 0.4 – 0.6 the LB medium was supplemented by 1 mM of NMH and pEVOL_NMH was induced by 0.02 % arabinose. Afterwards, when the OD_{600} reached 0.6 – 0.8, AvLDO mutants were overexpressed by induction of pASKGATE with different concentrations of ATET.

After induction, for both cases (wild type and mutants), cultivation of *E. coli* BL21 (DE3) was continued overnight at 20 °C and 120 rpm. Samples of culture were taken immediately after the induction (t_0) and after 24 h (t_{24}) to analyze the expression of the recombinant proteins. The culture

was harvested by centrifugation (Avanti® J-20 XP, Beckman coulter, USA; 8,000 g, 15 min, 4 °C) and the cell pellet was suspended in 30 mL of 0.9 % NaCl, as an additional washing step. After centrifugation (5810 R, Eppendorf, Germany; 4,000 g, 15 min, 4 °C), the pellet was stored at -20 °C until further use.

3.2.10.4. Overexpression of GriE wild type in pET28a(+) and pASKGATE vectors

In order to test which plasmid system shows better expression of the L-leucine hydroxylase GriE_wt, the heterologous expression was investigated in *E. coli* BL21 (DE3) using pET28a(+) and pASKGATE. A large-scale cultivation using 400 mL of LB-medium was supplemented with the corresponding antibiotic [for pET28a(+) 40 µg mL^{-1} of Kan; for pASKGATE 50 µg mL^{-1} of Sm; as listed in Table 21] and 2.5 % v/v of anti-foam SE-15 emulsion (Sigma-Aldrich, USA) was added in LB medium. The LB medium was inoculated with an ONC of *E. coli* BL21 (DE3) to an OD_{600} of 0.1 and incubated at 37 °C and 120 rpm until OD_{600} reached value of 0.6-0.8. Overexpression was performed by induction of T7 promoter with 0.1 mM or 1.0 mM IPTG and by induction of *tetA* promoter with 50 µg L^{-1} or 100 µg L^{-1} ATET for pASKGATE, when an OD_{600} of 0.6-0.8 was reached (Correia Cordeiro et al., 2018; Bertram and Hillenthe, 2008). After induction, cultivation of *E. coli* BL21 (DE3) was performed overnight at 20 °C and 120 rpm. Samples of *E. coli* BL21 (DE3) cultivation were taken after the induction (t_0), after 4 h (t_4) and after 24 h (t_{24}) to analyze the expression of the recombinant protein [GriE_wt expressed from pET28a(+) and/or pASKGATE]. The culture was harvested by centrifugation (Avanti® J-20 XP, Beckman coulter, USA; 8,000 g, 15 min, 4 °C) and the cell pellet was suspended in 30 mL of 0.9 % NaCl, as an additional washing step. After centrifugation of suspension (5810 R, Eppendorf, Germany; 4,000 g, 15 min, 4 °C), the pellet was stored at -20 °C until further use.

3.2.10.5. Overexpression of GriE_1 mutant in pET28a(+) and pASKGATE vectors

In order to improve expression of soluble GriE_1, pET28a(+) and pASKGATE were investigated as expression systems. Different concentrations of arabinose, NMH and IPTG or ATET were tested (Table 22, Table 23).

Table 22. Six different conditions were tested to optimize the expression of soluble of pET28a(+)_GriE_1 / pEVOL_NMH

Label	**Arabinose**	**N-methylhistidine**	**IPTG**
1	0.2 %	1.0 mM	0.1 mM
2	0.2 %	1.0 mM	1.0 mM
3	0.02 %	6.0 mM	0.1 mM
4	0.2 %	6.0 mM	0.1 mM
5	0.02 %	6.0 mM	1.0 mM
6	0.2 %	6.0 mM	1.0 mM

Table 23 Six different conditions were tested to optimize the expression of soluble of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH

Label	**Arabinose**	**N-methylhistidine**	**ATET**
7	0.02 %	1.0 mM	50 µg L^{-1}
8	0.02 %	1.0 mM	25 µg L^{-1}
9	0.2 %	1.0 mM	50 µg L^{-1}
10	0.2 %	1.0 mM	25 µg L^{-1}
11	0.2 %	6.0 mM	50 µg L^{-1}
12	0.2 %	6.0 mM	25 µg L^{-1}

Heterologous expression was investigated in the *E. coli* BL21 (DE3) pEVOL_NMH strain, as this leads to the incorporation of NMH at the amber stop codon position. Expression of GriE mutant H110NMH (GriE_1) was done in a small-scale cultivation. 100 mL of LB-medium was supplemented with the corresponding antibiotic [Kan and Chl for pET28a(+)_GriE_1 and Sm and Chl for pASKGATE_His_Leu_GriE_1, as listed in Table 21] and 2.5 % v/v of anti-foam SE-15 emulsion (Sigma-Aldrich, USA) was added in LB medium. The medium was inoculated with an ONC to an OD_{600} of 0.1 and incubated until OD_{600} reached 0.4 - 0.6 [for pET28a(+)_GriE_1] and OD_{600} of 1.0 – 1.3 (for pASKGATE_His_Leu_GriE_1).

For (a) pET28a(+)_GriE_1, once an OD_{600} of 0.4 - 0.6 was reached, NMH was added and pEVOL_NMH plasmid was induced by arabinose. Afterwards (while still incubating on 37 °C and 120 rpm) when an OD_{600} of 0.6 - 0.8 was reached, overexpression of pET28a(+) plasmid was induced by IPTG (Table 22).

For (b) pASKGATE_His_Leu_GriE_1, high OD_{600} induction needed to be performed as an answer to a low cell growth. Once an OD_{600} of 1.0 – 1.3 was reached, NMH was added and pEVOL_NMH

plasmid was induced by arabinose. Afterwards (while still incubating on 37 °C and 120 rpm) when the OD_{600} reached 1.4 – 1.6, overexpression of pASKGATE plasmid was induced by ATET (Table 23).

After induction, for both cases, the cultivation was performed overnight at 20 °C and 120 rpm. Samples were taken after the induction (by IPTG or ATET) (t_0), after 4 h (t_4) and after 24 h (t_{24}) to analyze the expression of the recombinant protein.

3.2.11. Purification of expressed GriE (both wild type and mutant)

Expressed GriE_wt from pET28a(+) and pASKGATE was purified according to this protocol, while GriE_1 mutant expressed from either pET28a(+) or pASKGATE was used in further activity assays only as cell lysate (cell-free extract) achieved after sonication and centrifugation (according to the part of this protocol, described below).

The cell pellet obtained from the expression (see Chapters 3.2.10.3. and 3.2.10.4.) was thawn and suspended in lysis buffer (10 mL of lysis buffer per g of pellet; Table 24) to permeabilize the cells. Preparation of potassium phosphate buffer, pH 8, needed for preparation of lysis buffer, is described in Table 25.

Table 24. Components of the lysis buffer used for the purification of GriE protein

Component	**Concentration**
Potassium phosphate buffer (KPi), pH 8	50 mM
NaCl	500 mM
Glycerol	10 % (v/v)

Table 25. Components of Potassium phosphate buffer, pH 8

Component	**Volume**
K_2HPO_4 [1 M]	94.0 mL
KH_2PO_4 [1 M]	6.0 mL
ddH_2O	up to 1 L

Afterwards the cells were sonicated (Branson Ultrasonics™ Sonifier S-250A, Thermo Scientific, USA; output control 7-8, duty cycles 80 %, 5 min**,** two times). After removal of cell debris by centrifugation (Avanti® J-20 XP, Beckman coulter, USA; 15,000 g, 30 min, 4 °C), the target

enzyme was purified from the supernatant by Ni-affinity chromatography (GE Healthcare, USA). The Ni Sepharose™ 6 Flow affinity column (Anonymous 6, 2021; GE Healthcare, USA) was equilibrated with five column volumes (cv) of lysis buffer. The cell lysate (V=15 mL) was applied to the column and incubated on ice for 30 min on a rotary shaker at 200 rpm. After collecting the flow through, the column was washed with 10 cv of lysis buffer (Table 24) containing 10 mM imidazole. A second wash was implemented with 10 cv of lysis buffer containing 30 mM imidazole. Each flow-through was collected for SDS-PAGE analysis (see Chapter 3.2.13.). The protein was eluted with 5-7 cv of lysis buffer containing 300 mM imidazole and collected in 1.5 mL aliquots. In the end, the column was washed with 5 cv of lysis buffer containing 500 mM imidazole. This eluate fractions were pooled together and concentrated to a final volume of 2.5 mL (5810 R, Eppendorf, Germany; 4,000 rpm, 4 °C) using a Vivaspin® 20 centrifugal concentrator (10 kDa MWCO, Sartorius, Germany). For imidazole removal, the concentrated eluate was applied on a PD-10 desalting column (GE Healthcare, USA) according to the manufacturer's protocol. Each collected purification fraction was measured spectrophotometrically on NanoDrop™ 2000 for SDS-PAGE analysis, while the protein concentration of the purified and desalted eluate was measured using the bicinchoninic acid assay (BCA assay, see Chapter 3.2.12.). The purified protein was directly used for the activity assay (see Chapter 3.2.14.) or filtrated through 0.45 µm filter Minisart® Syringe Filter (Sartorius, Germany) and stored at 4 °C until further use.

3.2.12. Determination of protein concentration

The protein concentration of each collected purified elution fraction was determined spectrophotometrically with the Nanodrop™ 2000c spectrophotometer (Thermo Scientific, USA). For a more accurate determination of the protein concentration a bicinchoninic acid assay was performed (BCA assay, Thermo Fisher Scientific, USA) according to the manufacturer's protocol (Anonymous 7, 2021). The absorption at 562 nm was measured in triplicates with a plate reader (BioTek Instruments, USA).

3.2.13. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

In order to qualitatively determine, whether the expression of the desired gene was successful the samples were analyzed by SDS PAGE (described below).

3.2.13.1. SDS-PAGE of soluble and insoluble protein fractions from time samples

Cell pellets gathered during cultivations of (a) *E. coli* BL21 (DE3) transformed with pASKGATE_His_Leu_AvLDO_wt, pET28a(+)_GriE_wt or pASKGATE_His_Leu_GriE_wt and (b) *E. coli* BL21 (DE3) pEVOL_NMH transformed with pASKGATE_His_Leu_AvLDO_1, pASKGATE_His_Leu_AvLDO_2, pET28a(+)_GriE_1 or pASKGATE_His_Leu_GriE_1 (see Chapter 3.2.10.) were disrupted with BugBuster®. The following equation was used to calculate the necessary amount of BugBuster® for each sample taken during already mentioned cultivations of *E. coli* BL21 (DE3).

$$\text{V (BugBuster [µL])} = \frac{OD600 \times V(sample,\ \mu L)}{30}$$

The mixture of sample taken during cultivation with corresponding volume of BugBuster® was shaken at 300 rpm in a ThermoMixer® for 20 min at RT. Afterwards the sample was centrifuged (5415 R, Eppendorf, Germany) at 16,000 rpm and 4 °C for 20 min and the supernatant was transferred into a fresh Eppendorf tube (soluble fraction). After that centrifugation and supernatant transfer, remained cell pellet was suspended in lysis buffer (Table 24), in the same amount as BugBuster® was added in the first place (according to the above-mentioned equation). Cell pellet suspended in lysis buffer was marked as insoluble fraction. 5 µL of sample buffer (4x NuPAGE, Thermo Fisher Scientific, USA) and 2 µL of reducing agent (10 x, NuPAGE, Thermo Fisher Scientific, USA) were added to 13 µL of the sample, denaturated at 70 °C for 10 min and cooled down to the RT. As standard, 6 µL of PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA) was used. NuPAGE Bis-Tris Invitrogen gels (ThermoFisher Scientific, USA) were run in NuPAGE™ MES SDS Running Buffer (1 x, Thermo Fisher Scientific, USA) at 120 V and 250 mA for 1 h. They were stained with Coomassie Brillant Blue overnight (Table 26) and destained with glacial acetic acid/ ethanol (Table 27).

Table 26. Components of staining solution. Freshly prepared solution was filtrated and stored at room temperature.

Component	**Mass / volume (V_{total}=1 L)**
Coomassie Brilliant Blue R-250	1 g
Glacial acetic acid (conc.)	100 mL
Ethanol	300 mL
ddH2O	600 mL

Table 27. Components of destaining solution. Freshly prepared solution was filtrated and stored at room temperature

Component	**Volume (V_{total}=1 L)**
Glacial acetic acid (conc.)	100 mL
Ethanol	300 mL
ddH2O	600 mL

3.2.13.2. SDS-PAGE of purified protein fractions

After determination of the protein concentration of each purification fraction by BCA assay (see Chapter 3.2.12.), purified proteins were also analyzed by SDS-PAGE. For that, 5 µg total protein (max. 13 µL) was analyzed on the NuPAGE Bis-Tris Invitrogen gels (ThermoFisher Scientific, USA), following the same protocol as the one described under the subchapter 3.2.13.1.

3.2.14. Activity assay

In order to examine the substrate scope of the enzymes and to investigate if there are any differences between the activity of GriE wild type and GriE mutant (GriE_1) an activity assay was performed.

3.2.14.1. Enzymatic reaction

The setup for the enzymatic reaction can be found in Table 28.

Table 28. Components of activity assay

Component	**Concentration (V_{total}=1 L)**
Ascorbate	5 mM
α-Ketoglutarate	100 mM
L-leucine / L-isoleucine (substrate)	50 mM
$FeSO_4 \times 7H_2O$	1.7 mM
Enzyme	Purified: 2 mg mL^{-1}: cell lysate: 1.7 mg mL^{-1}
KPi buffer (see Table 25)	50 mM

α-ketoglutarate and ascorbate were prepared by dissolving the chemicals in 50 mM KPi buffer, while $FeSO_4 \times 7H_2O$ was prepared by dissolving the chemical in ddH_20 **(**Table 28). pH 8 of the enzymatic reaction was adjusted with HCl (37 % v/v) or NaOH (50 % v/v). The assay was

performed in 10 mL glass vials with a reaction volume of 1 mL at RT, at 120 rpm. Time samples were taken at the start of the enzymatic reaction (when the enzyme was added, t_0), after 1 h of the setup of the enzymatic reaction (t_1) and after 24 h (t_{24}). Samples were taken by quenching 50 µL of the enzymatic reaction with the 25 µL of 50 % v/v acetic acid. This mixture of the enzymatic reaction with acetic acid was stored at -21 °C, if not used directly for thin-layer chromatography. Time samples of activity assay was analyzed with thin-layer chromatography (TLC) and ninhydrin staining.

3.2.14.2. Thin-layer chromatography (TLC)

The hydroxylation of the substrates (L-leucine and L-isoleucine) to products (5-hydroxyleucine and 5-hydroxyisoleucine) can be confirmed by TLC and staining (of enzymatic reaction time samples as well as applied references on silica gel) with ninhydrin. 2 µL of the prepared sample (see Chapter 3.2.14.1.) was applied on a silica gel TLC plate (Classic silica plate without fluorescence indicators; Merck Millipore, USA) and run with 1- butanol : acetic acid : ddH_2O (22:3:5) as the mobile phase. 5 mM of L-leucine and 5 mM of L-isoleucine served as standards. Plates were dried and the TLC was stained with 0.2 % w/v ninhydrin. The retention factor (Rf) of developed samples was calculated for each with equation:

$$\text{Rf}= \frac{distance\ moved\ by\ solute\ (cm)}{distance\ moved\ by\ solvent\ (cm)}$$

4. RESULTS AND DISCUSSION

In order to achieve a more stable and stronger coordination of the Fe(II) atom in the 2-His 1-carboxylic motif of the α-ketoglutarate dependent dioxygenases (see Chapter 2.1.), two L-leucine 5-hydoxylases - AvLDO (isolated from *Anabaena variabilis*, AvLDO_wt) and GriE (isolated from *Streptomyces muensis*, GriE_wt), were studied. In the facial triad of these enzymes, the non-canonical amino acid N-methylhistidine (NMH), was introduced. The activity and the substrate scope were investigated for these wild types, as well as their mutants.

Accordingly, the aims of this study were:
(a) to clone and express L-leucine 5-hydroxylase *AvLDO* wild type and the corresponding mutants (*AvLDO_1* with incorporated NMH at the location 151, and *AvLDO_2* mutant with incorporated NMH at location 237; see Chapter 3.1.1.) in pASKGATE,
(b) to compare protein expression of GriE_wt in the pET28a(+) and pASKGATE vectors,
(c) to optimize and improve the expression of soluble of the GriE_1 mutant (*GriE_1* with H110NMH), utilizing NMH in the active triad of the enzymes, with both pET28a(+) and pASKGATE plasmids,
and (d) to determine what are the differences, if there are any, in the expressed protein concentrations, substrate scope and enzyme activity of the GriE_1 mutant in comparison to the wild type, expressed in both, pET28a(+) and pASKGATE.

To achieve these aims, (a) *AvLDO* was cloned in the pENTRE entry vector using FastCloning, and secondly pASKGATE_His_Leu_AvLDO was generated by the Gateway® cloning. Using QuikChange™, site-specific mutations were introduced, resulting in the construction of two mutants *AvLDO_1* (at the location H151) and *AvLDO_2* (H237), necessary for NMH incorporation by stop codon suppression (see Chapter 2.2.1.1.). Expression of AvLDO_wt, AvLDO_1 and AvLDO_2 was qualitatively compared using SDS-PAGE analysis.
(b) Expression of soluble GriE_wt, from pET28a(+) and pASKGATE, was investigated and quantitatively compared by determination of purified protein concentration. Additionally, (c) the expression of soluble of GriE_1 was optimized. Furthermore, (d) purified enzyme (GriE_wt) or cell lysate containing GriE_1 (see Chapter 3.2.11.) was used for biotransformations of L-leucine and L-isoleucine. The activity was assessed using thin-layer chromatography (TLC).

4.1. *IN VITRO* CLONING OF GENES CODING FOR L-LEUCINE 5-HYDROXLASES, WILD TYPES AND CORRESPONDING MUTANTS

The *GriE_wt* and *AvLDO_wt* and their corresponding mutants were already available in pET28a(+) (see Table 3). *GriE*_wt and the mutant H110NMH (*GriE_1*) were available within the pASKGATE plasmid at the Institute from previous research (Table 3). Accordingly, in this study it was necessary to firstly clone the *AvLDO_wt* in pASKGATE plasmid, as well as generate the mutants of *AvLDO_wt* (*AvLDO_1* and *AvLDO_2*) in the pASKGATE vector and secondly to produce the H210NMH mutant of *GriE_wt* (*GriE_2*) in pASKGATE vector.

4.1.1. Construction of pENTRE_His_Leu_AvLDO_wt using FastCloning

In order to achieve aim (a), cloning and expressing of *AvLDO*_wt and the corresponding mutants (*AvLDO_1*, *AvLDO_2*) in the pASKGATE plasmid, construction of the pENTRE_His_Leu_AvLDO_wt was performed. As explained in subchapter 2.3.1. it was important to clone the construct pENTRE_His_Leu_AvLDO_wt firstly, because it acts as an intermediate that enables further gene insertion in the pASKGATE destination vector by using the Gateway® platform and Gateway® methodology (see Chapter 2.3.1.).

The insert *AvLDO_wt* amplification, described in subchapter 3.2.1., was confirmed by agarose gel electrophoresis (Figure 8).

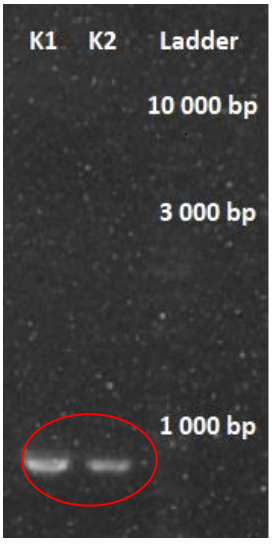


Figure 8. Agarose gel of PCR insert *AvLDO_wt* amplification from pET28a(+)_AvLDO_wt. Both reactions resulted in amplified *AvLDO_wt*, labeled as K1 and K2, correspond to the size of 800 bp (highlighted by red circle) which is the expected size of the PCR product. Ladder: GeneRuler 1kb (Thermo Fisher Scientific, USA).

The size of the amplified *AvLDO_wt* product, which can be seen in Figure 8, corresponded to 800 bp. That was the expected size of the *AvLDO_wt* gene with two additional base pairs, thymine and guanine added in 5' to 3' direction. The addition of these bases was necessary for maintaining the reading frame in pASKGATE plasmid, disturbed due to the 25 bp size of attB flanking sites necessary for Gateway® cloning system, cloned in the destination pASKGATE vector (see Chapter 2.3.1. and Figure 7). Addition of these two nucleotides which further translated for leucine, also had impact on the labeling of Gateway® constructs used in this Graduate Thesis. The name of plasmid backbone; His as 6xHis tag on N-terminus; Leu as additional amino acid for avoiding the frameshift; the name of the insert (*AvLDO* or *GriE*, wt or mutant) - resulted in the labeling such as pENTRE_His_Leu_AvLDO_wt or pASKGATE_His_Leu_GriE_wt.

Chemically competent *E. coli* TOP 10 cells were transformed with the *AvLDO_wt* amplified insert (above described) and the linearized vector backbone, according to the Chapter 3.2.6. and plated on LB agar plates containing the corresponding antibiotic (Table 19 and Table 21). After incubation (see Chapter 3.2.6.), colonies were picked, and the colony PCR was performed (see Chapter 3.2.3.1. and Figure 9).

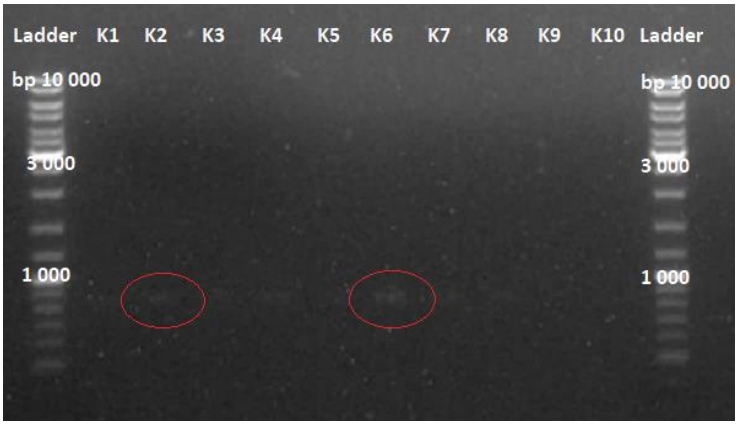


Figure 9. Agarose gel of *AvLDO_wt* gene incorporated in the pENTRE backbone by FastCloning. cPCR was done with ten colonies (K1-K10) picked after transformation of chemically competent *E. coli* TOP10 cells with the linear backbone (pENTRE) and amplified insert (*AvLDO_wt*) created by FastCloning, to achieve pENTRE_His_leu_AvLDO_wt construct. Expected size of the *AvLDO_wt* with the primers P30-pENTRE-AvlDO-fw and P31-pENTRE-AvlDO-rv (highlighted by red circles), indicating successful construction of pENTRE_His_leu_AvLDO_wt was 842 bp. Positive clones were confirmed for colonies labeled as K2 and K6. Ladder: GeneRuler 1kb (Thermo Fisher Scientific, USA).

AvLDO_wt with the primers P30-pENTRE-AvlDO-fw and P31-pENTRE-AvlDO-rv used for this cPCR (Table 5), had expected size of 842 pb (Figure 9). These primers were also overhangs of the flanking attL sites of the pENTRE backbone (for plasmid map see Appendices 5), necessary for

Gateway® cloning system (see Chapter 2.3.1. and Figure 7). To summarize, colony PCR (cPCR) resulted in the positive clones with an expected size of 842 bp, which indicated successful cloning of *AvLDO_wt* in pENTRE backbone and therefore resulting in the pENTRE_His_Leu_AvLDO_wt construct generation. Additionally, pENTRE_His_Leu_AvLDO_wt was confirmed with Sanger sequencing (see Chapter 3.2.9.2., data not shown).

4.1.2. Construction of pASKGATE_His_Leu_AvLDO_wt using Gateway®

The Gateway® reaction was performed with a confirmed positive clone pENTRE_His_Leu_AvLDO_wt labeled as K6 (Figure 9), constructed previously by FastCloning (see Chapter 4.1.1.). Colony PCR (cPCR) was performed and due to a double control of the Gateway® cloning system; the different antibiotic resistance between entry vector and expression vector (pASKGATE_His_Leu_AvLDO_wt) and because of the toxin gene (*ccdB*) which ends up incorporated in the toxin byproduct (Figure 7), higher cloning efficiency was enabled. Moreover, all colonies of transformed *E. coli* BL21 (DE3) where determined as positive (Figure 10).

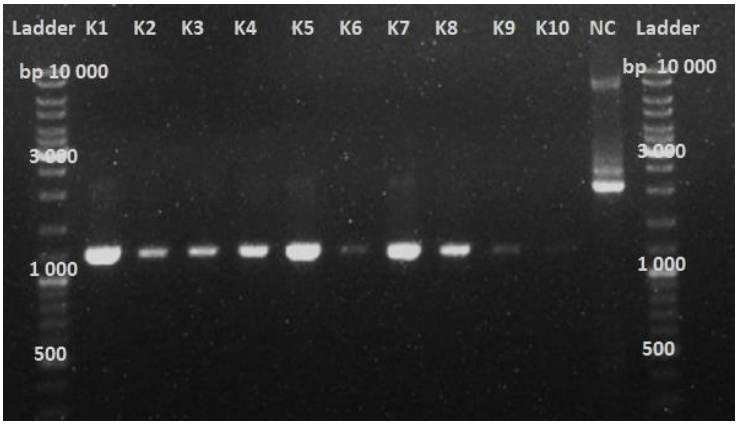


Figure 10. Agarose gel of *AvLDO_wt* incorporated in the pASKGATE backbone. cPCR was done with ten colonies (K1 – K10) picked after the transformation of *E. coli* BL21 (DE3) cells with the Gateway® reaction for production of pASKGATE_His_Leu_AvLDO_wt construct. Expected size of the *AvLDO_wt* with the P5_pASK_Seq_rv and P20_pASKGATE_Seq_fw, indicating successful construction of pASKGATE_His_Leu_AvLDO_wt was 1245 bp. Ladder: GeneRuler 1kb (Thermo Fisher Scientific, USA).

AvLDO_wt with the primers P5_pASK_Seq_rv and P20_pASKGATE_Seq_fw used for this cPCR (Table 5) had expected size of 1245 pb, as shown in Figure 10. PCR using these aforementioned primers enabled amplification of part of pASKGATE plasmid backbone, the 6xHis sequence, Xa cleavage site, *AvLDO_wt* gene and both attB flanking site of the pASKGATE backbone (Appendices 6) – in another words, 1245 bp size indicates successful incorporation of the gene

AvLDO_wt in the pASKGATE plasmid and therefore resulting in the pASKGATE_His_Leu_AvLDO_wt construct. The colony marked as K1 was selected for sequencing (see Chapter 3.2.9.2.), which confirmed the correct assembly of pASKGATE_His_Leu_AvLDO_wt (data not shown).

4.1.3. Site-directed mutagenesis using QuikChange™

Finally, in order to produce the *AvLDO* mutants, the site-directed mutagenesis QuikChange™ was performed on pASKGATE_His_Leu_AvLDO_wt. Mutagenesis resulted in substitution of histidine (CAC/CAT) with the amber stop codon (TAG) at the position H151 (labeled pASKGATE_His_Leu_AvLDO_1, Figure 11 [A]) and at the position H237 (pASKGATE_His_Leu_AvLDO_2, Figure 11 [B]).

Figure 11. Sequencing results of pASKGATE_His_Leu_AvLDO_1 and pASKGATE_His_Leu_AvLDO_2 cloning. [A] QuikChange™ mutagenesis resulted in an exchange of nucleotides coding for the amino acid histidine to the amber stop codon at the position of H151. [B] QuikChange™ mutation resulted in an exchange of nucleotides coding for the amino acid histidine to the amber stop codon at the position of H237. Sequences

pASKGATE_His_Leu_AvLDO_1 and pASKGATE_His_Leu_AvLDO_2 were aligned with pASKGATE_His_Leu_AvLDO_wt sequence. Primer: pASKGATE_Seq_fw (Microsynth AG, Switzerland).

The *GriE_2* mutant (H210NMH; pASKGATE_His_Leu_GriE_2) has not been successfully cloned using the QuikChange™ mutagenesis technique (pASKGATE_His_Leu_GriE_2). Accordingly, the initial aims of this Thesis were modified, leaving the expression of soluble GriE_2, and optimization of it, behind. The optimization and improvement of soluble expression, as well as investigating the substrate scope and enzymatic activity, were only investigated with the GriE_wt and the mutant GriE_1.

Attempts of the QuikChange™ mutagenesis (see Chapter 3.2.4.) to produce pASKGATE_His_Leu_GriE_2, included carrying out a gradient PCR under a temperature range around the calculated annealing temperature (55 °C-65 °C), increase in the *Dpn1* digestion time (digestion for 3 h at 37 °C) and increased amount of used *Dpn1* (double the amount). Redesigning of the QuickChange™ primers used for pASKGATE_His_Leu_GriE_2 should be considered as the next step. Redesigning the primers in a way that the non-overlapping sequence at the 3' end is larger than the overlapping sequences at 5' end, could potentially result in the successful site mutation - as this would lead to a complete suppression of the primer self-annealing (Xia and Xun, 2017; Liu and Naismith, 2008).

Alternative site-directed mutagenesis techniques could also be considered for the possible future cloning of *GriE_2*, such as the Q5® Site Directed Mutagenesis using the Q5™ High-fidelity DNA polymerase (known for its low error rate) and/or site-directed mutagenesis using two single-primer reactions in parallel (SPRINP). According to the NewEngland Biolabs, Q5® Site Directed Mutagenesis kit shows much higher efficiency compared to the QuikChange mutagenesis Agilent kit. Furthermore, in the case of substitution, 90% of the resultant colonies had the desired mutation incorporated by Q5® Site Directed Mutagenesis (Anonymous 8, 2021), while Liu et al. (2015) demonstrated that the accuracy and mutation rate of Q5® Site Directed Mutagenesis for the plasmid of typical size (6,4 kb) was 100 %. Another mentioned alternative could be Single-Primer Reactions in Parallel (SPRINP) mutagenesis which requires the usage of a single primer in one reaction (Edelheit et al., 2009).

4.2. EXPRESSION AND PURIFICATION OF RECOMBINANT L-LEUCINE 5-HYDROXYLASES

4.2.1. Protein expression of AvLDO_wt and its mutants AvLDO_1 and AvLDO_2 in pASKGATE vector

After successful *in vivo* cloning of *AvLDO_wt* in pASKGATE vector, as well as after successful site-directed mutagenesis resulting in the mutants pASKGATE_His_Leu_AvLDO_1 and pASKGATE_His_Leu_AvLDO_2 - the protein expression was studied.

The heterologous expression of the AvLDO_wt was investigated in *E. coli* BL21 (DE3), while heterologous expression of the mutants was investigated in the *E. coli* BL21 (DE3) with pEVOL_NMH (see Chapter 3.2.10.3.). pEVOL_NMH is plasmid carrying the gene which encodes for aminoacyl-tRNA synthetase / tRNA pair much needed for NMH incorporation in the protein sequence and therefore pEVOL_NMH was necessary for the overexpression of the AvLDO mutants (AvLDO_1 and AvLDO_2).

In order to optimize the expression of AvLDO wild type and mutants, in pASKGATE, different concentrations of the inducer anhydrotetracycline (ATET; 50, 100, 150 and 200 $\mu g\ L^{-1}$) were investigated. After cell disruption using BugBuster® (see Chapter 3.2.13.1.), the soluble fraction (supernatant given by centrifugation of disrupted cells) and insoluble fractions (cell pellet suspended in lysis buffer) of cell extract containing overexpressed proteins (AvLDO_wt, AvLDO_1 and AvLDO_2), taken during different time points (t_0 and t_{24}), were analyzed by SDS-PAGE (Figure 12 - 15). The t_0 samples served as controls since no protein was expected at the point of induction and only native *E. coli* BL21 (DE3) proteins were visible.

Even though expected molecular weight of overexpressed AvLDO, both wild type and AvLDO mutants was 30.5 kDa - based on the experience from previous research related to the overexpression of the same enzymes (Hanreich, 2019, is that our enzyme of interest will shown on SDS-PAGE gel in the level of between 35 and 40 kDa band compared to the used PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA). Also, it is important to note that the SDS-PAGE is a qualitative method. Furthermore, the aim of these experiments and electrophoresis was not the quantitative detection of the overexpressed proteins but the monitoring and comparison of expressions between wild type and mutants.

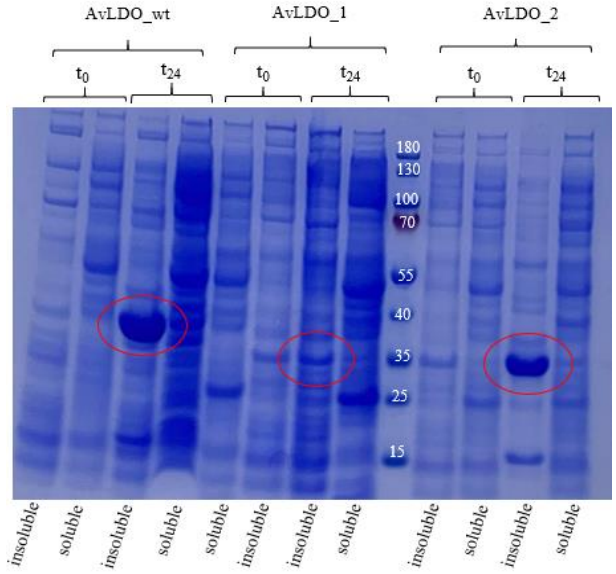


Figure 12. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) harboring pASKGATE_His_Leu_AvLDO_wt, and *E. coli* BL21 (DE3) with pEVOL_NMH harboring pASKGATE_His_Leu_AvLDO_1 and/or pASKGATE_His_Leu_AvLDO_2, taken at different time points (t0, t4, t24). Induction was performed with 50 µg L^{-1} ATET. Expected size of AvLDO_wt, AvLDO_1 and AvLDO_2 (highlighted by red circles) was approximately 30.5 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

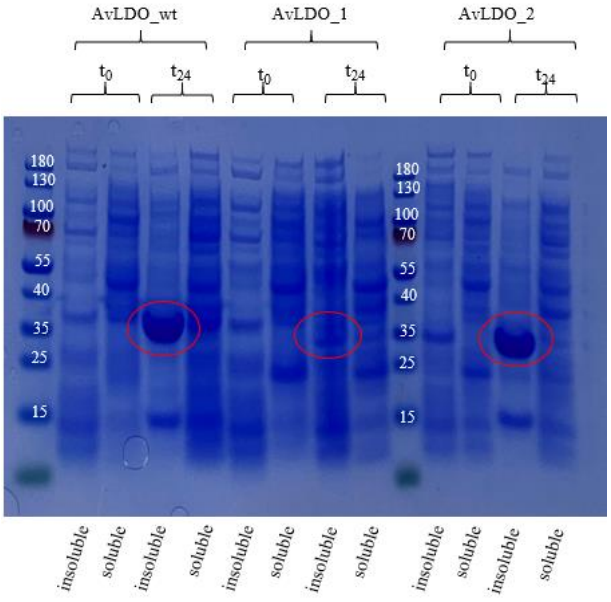


Figure 13. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) harboring pASKGATE_His_Leu_AvLDO_wt, and *E. coli* BL21 (DE3) with pEVOL_NMH harboring pASKGATE_His_Leu_AvLDO_1 and/or pASKGATE_His_Leu_AvLDO_2, taken at different time points (t0, t4, t24). Induction was performed with 100 µg L^{-1} ATET. Expected size of AvLDO_wt, AvLDO_1 and AvLDO_2 (highlighted by red circles) was approximately 30.5 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

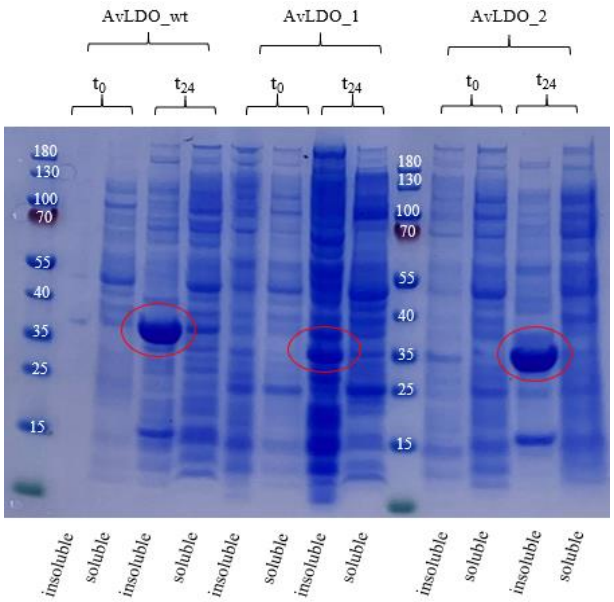


Figure 14. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) harboring pASKGATE_His_Leu_AvLDO_wt, and *E. coli* BL21 (DE3) with pEVOL_NMH harboring pASKGATE_His_Leu_AvLDO_1 and/or pASKGATE_His_Leu_AvLDO_2, taken at different time points (t0, t4, t24). Induction was performed with 150 µg L^{-1} ATET. Expected size of AvLDO_wt, AvLDO_1 and AvLDO_2 (highlighted by red circles) was approximately 30.5 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

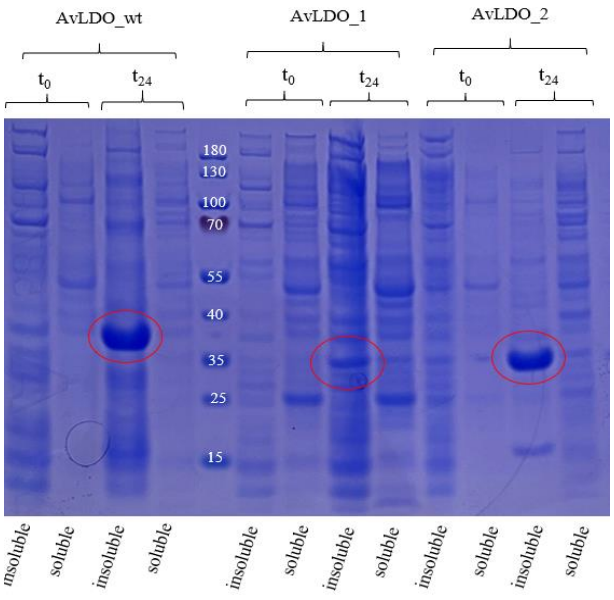


Figure 15. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) harboring pASKGATE_His_Leu_AvLDO_wt, and *E. coli* BL21 (DE3) with pEVOL_NMH harboring pASKGATE_His_Leu_AvLDO_1 and/or pASKGATE_His_Leu_AvLDO_2, taken at different time points (t0, t4, t24). Induction was performed with 200 µg L^{-1} ATET. Expected size of AvLDO_wt, AvLDO_1 and AvLDO_2 (highlighted by red circles) was approximately 30.5 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

As already said, expected molecular weight of overexpressed AvLDO, both wild type and mutants, was 30.5 kDa - based on the experience from previous research of the work group from the Institute of Molecular Biotechnology at TU Graz, it can be claimed that the circled bands are the protein of interest. In the insoluble fractions of t_{24} samples (Figure 12 -Figure 15), a high intensity full-length protein bands at the approximated size of 30.5 kDa (AvLDO_wt, AvLDO_1 and AvLDO_2) can be seen for all the tested ATET concentrations. This indicates that the incorporation of NMH, non-canonical amino acid of interest, was successfully incorporated at both positions (H151NMH and H237NMH).

In order to truly verify the correct incorporation of the NMH, control overexpression of the same constructs, pASKGATE_His_Leu_AvLDO_1 and pASKGATE_His_Leu_AvLDO_2 but without addition of NMH, should be done. If in that case, truncation of protein occurs, it can be claimed that no background incorporation of canonical amino acids took place and that NMH is the one incorporated at the locations of previously mutated amber stop codon (H151 and H237) of the mutants (AvLDO_1 and AvLDO_2, respectively).

However, the proteins AvLDO_wt, AvLDO_1 and AvLDO_2 were mostly detected in the insoluble fractions, indicating that the overexpressed proteins were mostly incorrectly folded. The heterologous high-level expression of AvLDO_wt, AvLDO_1 and AvLDO_2, may lead to inclusion body formation because of the disbalance between dynamic formation of soluble protein and aggregated protein (Singh et al., 2015; Villaverde and Carrio, 2003). It is one of the most common, but undesirable, outcomes of protein induction, while producing soluble recombinant protein (Garcia- Fruitos, 2010).

pASKGATE_His_Leu_AvLDO_1 certainly has shown a weaker intensity band, suggesting a lesser protein concentration compared to the wild type and AvLDO_2. For AvLDO_1, bands at 30.5 kDa for t_{24} insoluble fractions, were noticeably stronger when induced by 100 µg L^{-1} ATET (Figure 13) and 150 µg L^{-1} ATET (Figure 14), while AvLDO_wt and AvLDO_2 were expressed at a high level for all the inducer concentrations. AvLDO_wt and AvLDO_2 had better expression than AvLDO_1, the reason of which could be the positioning of the NMH incorporation. In similar case, Young et al. (2009) concluded that the protein context, meaning local protein structure and mRNA sequence, because of competition between release factor 1 (which recognizes the amber stop codon) and the aminoacyl-tRNA synthetase / tRNA pair for utilization of NMH, most likely affects the incorporation efficiency of non-canonical amino acid at specific location and it affects corresponding protein yields. To conclude, based on this qualitative expression comparison, it seems that the place of non-canonical amino acid incorporation at the H237NMH could be more favorable than the position H151NMH.

For the potential future studies regarding AvLDO mutants, soluble protein expression could be optimized by reducing the ATET concentration or by reducing cultivation temperature to 4 - 12 °C using *E. coli* Arctic Express as a host system, which would allow the growing polypeptide to fold correctly. Furthermore, the development of refolding strategy also could be an option to generate correctly folded protein.

4.2.2. Comparison of the GriE_wt expression in two vectors; pET28a(+) and pASKGATE

Taking into account the unsuccessful QuikChange™ reaction for acquiring *GriE_2* (see Chapter 4.1.3.), *GriE_wt* and only one *GriE_wt* mutant (*GriE_1*, H110NMH) were studied in two vectors [pET28a(+) and pASKGATE]. In order to determine which plasmid system is more suitable for the overexpression of GriE, the wild type was overexpressed in either pET28a(+) or pASKGATE, with different inducer concentrations. This knowledge should be further used to optimize the expression of the GriE_1 mutant.

4.2.2.1. Expression of pET28a(+) and pASK37(+)s as a negative control

Negative control of empty vector overexpression [pET28a(+) and pASK37(+)s (empty vector of pASKGATE)] in *E. coli* BL21 (DE3) was firstly performed. Empty vector means that the expression was performed only with the plasmid backbone, and no gene encoding for protein was carrying by that plasmid. Expression of pET28a(+) empty vector was induced by 0.1 mM and 1.0 mM IPTG, and expression of pASK37(+)s was induced by 50 µg L^{-1} and 100 µg L^{-1} ATET. Overexpression of empty vectors gave negative results (data not shown) which confirmed that the very sequence of the plasmid does not affect gene expression and protein culmination.

4.2.2.2. GriE_wt gene expression from pET28a(+)_GriE_wt

Heterologous expression of GriE_wt was investigated in *E. coli* BL21 (DE3) and the expression conditions were performed according to the Chapter 3.2.9.4. The time samples (t_0, t_4, t_{24}) from the pET28a(+)_GriE_wt overexpression, as well as the samples of the purified GriE_wt protein, were analyzed by SDS-PAGE and shown in Figure 16 - Figure 17. In addition, expected molecular weight of overexpressed GriE, wild type and mutant, expressed from pET28a(+) was 30.1 kDa. However, based on the experience from previous research related to the overexpression of the same enzyme (GriE; Hanreich, 2019) is that our enzyme of interest will show on SDS-PAGE gel between 35 and 40 kDa bands compared to the PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA). Again, it is important to note that the SDS-PAGE is a qualitative method for which the aim was not the quantitative detection of the overexpressed proteins, but the

monitoring of NMH incorporation in the active site of the mutant and comparison of expressions between wild type and mutant.

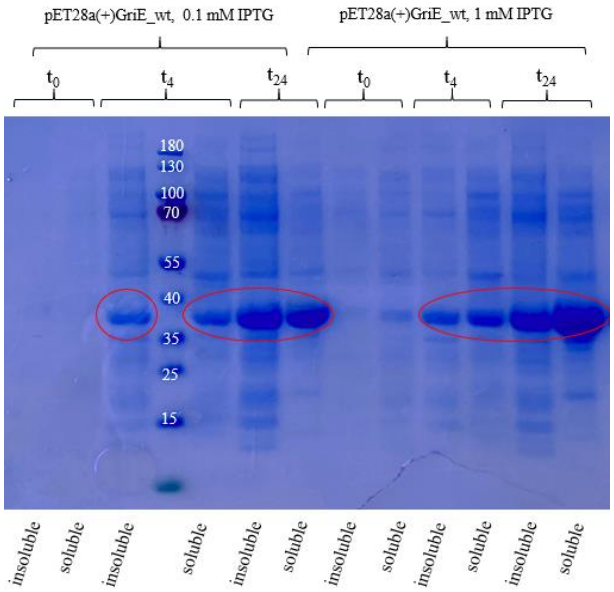


Figure 16. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pET28a(+)_GriE_wt taken at different time points (t0, t4, t24). Overexpression of pET28a(+)_GriE_wt was induced with 0.1 mM or 1.0 mM IPTG. Expected size of GriE_wt expressed from pET28a(+) (highlighted by red circles) was approximately 30.1 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

Regarding the pET28a(+)_GriE_wt expression (induced by either 0.1 mM or 1.0 mM IPTG), the highest amount of soluble protein was reached after 24 h (t_{24}), with no degradation of the protein observed after 4 h of expression - meaning that the concentration of the protein increased over time, as can qualitatively be seen comparing the thickness of the bands .of t_4 and t_{24} samples.

The same expression pattern can be seen for the overexpression induced by both IPTG concentrations (0.1 mM and 1.0 mM IPTG). Bands that approximately correspond to the size of 30.1 kDa and for which it can be claimed that it is our protein, based on previous research of Kourist's group on Institute for Molecular Biotechnology - were detected in the insoluble, as well as in the soluble fractions, indicating that the heterologous high-level expression of GriE_wt, may lead to partial inclusion body formation. By qualitative comparison, the concentration of the soluble protein induced by 1.0 mM IPTG is higher than the concentration of the soluble protein expressed by 0.1 mM IPTG (Figure 16).

Protein was purified according to the protocol described under the Chapter 3.2.11. Elution fractions (E_1 - E_8 and E_1 – E_9; Figure 17), obtained by purification of the GriE_wt protein expressed from pET28a(+), and induced by 0.1 mM and 1.0 mM IPTG, were also analyzed by SDS-PAGE (Figure 17).

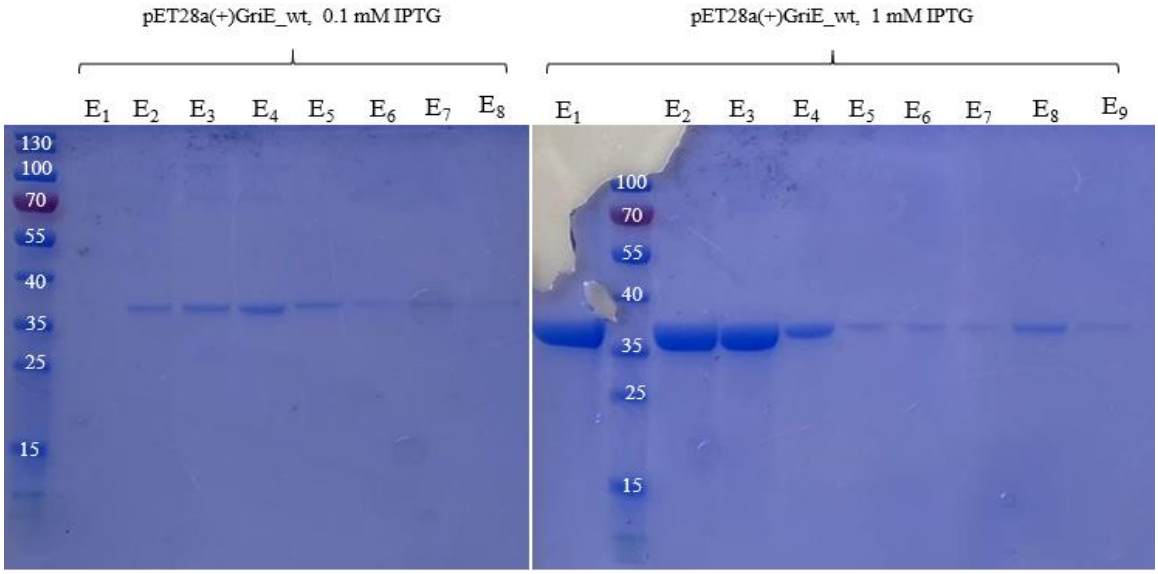


Figure 17. SDS-PAGE analysis of purified GriE_wt. Samples labeled as E_1 - E_8 and E_1 – E_9 are elution fractions obtained during protein purification - containing desired protein (pET28a(+)_GriE_wt) eluted with lysis buffer and 300 mM imidazole. Expression of GriE_wt from pET28a(+) was induced with 0.1 mM and 1.0 mM IPTG. Expected size of GriE_wt expressed from pET28a(+) was approximately 30.1 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

Generally, comparing all the fractions given by each step of protein purification; cell lysate, flow through, washing step, elution fractions and final wash (data not shown) - most of the desired protein was obtained in the elution fractions while using an imidazole concentration of 300 mM, as shown at Figure 17 and approximately corresponding to the size of 30.1 kDa, (see Chapter 3.2.11.). Additionally, as calculated using BCA assay (see Chapter 3.2.12.) - unspecific protein was mostly washed from the column by the wash one (flow through; lysis buffer containing 10 mM imidazole, see Chapter 3.2.11.) and almost no protein was washed by the wash two and by the final wash step (lysis buffer containing 30 mM imidazole / 500 mM imidazole, see Chapter 3.2.11.), indicating that the binding of the His-tagged dioxygenase works (data not shown).

4.2.2.3. GriE_wt gene expression from pASKGATE_His_Leu_GriE_wt

Heterologous expression of pASKGATE_His_Leu_GriE_wt was investigated in *E. coli* BL21 (DE3) by the induction of 50 µg L^{-1} and 100 µg L^{-1} ATET (Figure 18 - Figure 19) as described in Chapter 3.2.10.4.

GriE_wt expressed from pET28a(+) plasmid (30.1 kDa) was different in size comparing to the GriE_wt expressed from the pASKGATE (30.2 kDa), due to the addition of two base pairs, thymine and guanine. Addition was necessary for the maintaining the reading frame in pASKGATE, disrupted due to the Gateway system cloned in this vector (see Chapters 2.3.1. and 4.1.2.). As already mentioned, based on the experience from previous research related to the overexpression of the same enzyme (GriE; Hanreich, 2019) is that our enzyme of interest will show on SDS-PAGE gel between 35 and 40 kDa bands compared to the PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

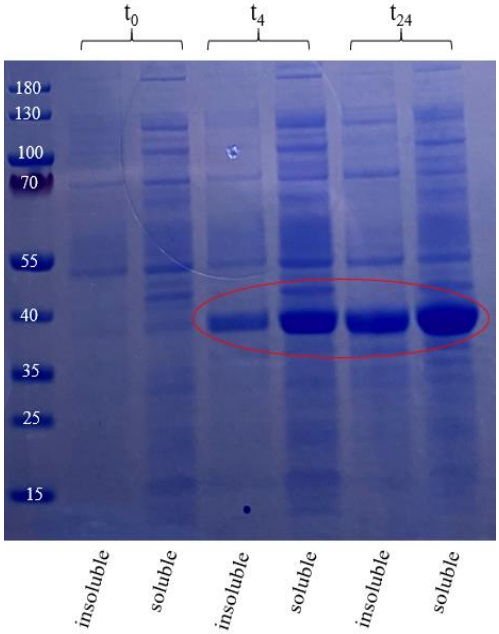


Figure 18. SDS-PAGE analysis of cell extract of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_wt taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_wt was induced with 50 µg L^{-1} ATET. Expected size of GriE_wt expressed from pASKGATE (highlighted by red circle) was approximately 30.2 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

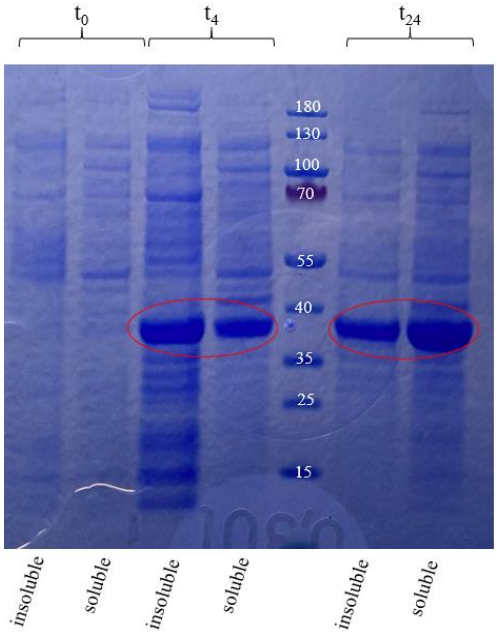


Figure 19. SDS-PAGE analysis of cell extract of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_wt taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_wt was induced with 100 µg L^{-1} ATET. Expected size of GriE_wt expressed from pASKGATE (highlighted by red circles) was approximately 30.2 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

Expression analysis of GriE_wt from pASKGATE_His_Leu_GriE_wt induced by 50 µg L^{-1} and 100 µg L^{-1} ATET resulted in a fully translated protein. GriE_wt protein expressed from pASKGATE, was approximately in the same amount in soluble and insoluble phase at the 4 h (t_4) and 24 h (t_{24}) samples, corresponding to the expected band, but only approximately corresponding to the size of 30.2 kDa (Figure 18 -Figure 19). Moreover, the increase of soluble protein can be observed at samples t_{24} in comparison to sample t_4, for both cases.

Protein (GriE_wt expressed from pASKGATE and induced by 50 and 100 µg L^{-1} ATET) was purified according to the protocol described under the Chapter 3.2.11. Elution fractions (E_1 – E_7), obtained by lysate buffer with 300 mM imidazole during protein purification, were also analyzed by SDS-PAGE (Figure 20).

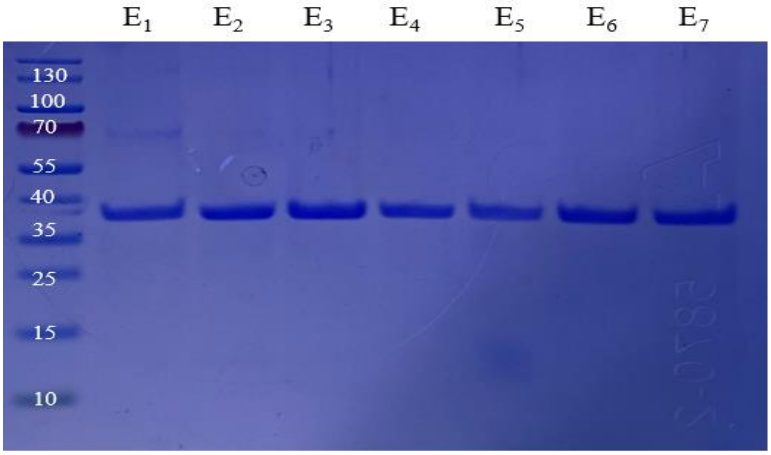


Figure 20. SDS-PAGE analysis of purified GriE_wt (E_1-E_7) induced by 100 µg L^{-1} ATET. Expected size of GriE_wt expressed from pASKGATE was approximately 30.2 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

Most of the desired GriE_wt protein expressed from pASKGATE (30.2 kDa) was obtained using an imidazole concentration of 300 mM (see Chapter 3.2.11., Figure 20). Most of the unspecific protein was washed by the wash one (lysis buffer containing 10 mM imidazole) and some was washed out by the last washing step (containing 500 mM imidazole), while no protein was washed by the second wash step (500 mM imidazole, see Chapter 3.2.11.).

Biological triplicates of *E. coli* BL21 (DE3) strain transformed with either pET28a(+)_GriE_wt or pASKGATE_His_Leu_GriE_wt were used for the overexpression comparison of GriE_wt. The BCA assay (see Chapter 3.2.12.) was used for the determination of the purified protein concentration (Figure 21).

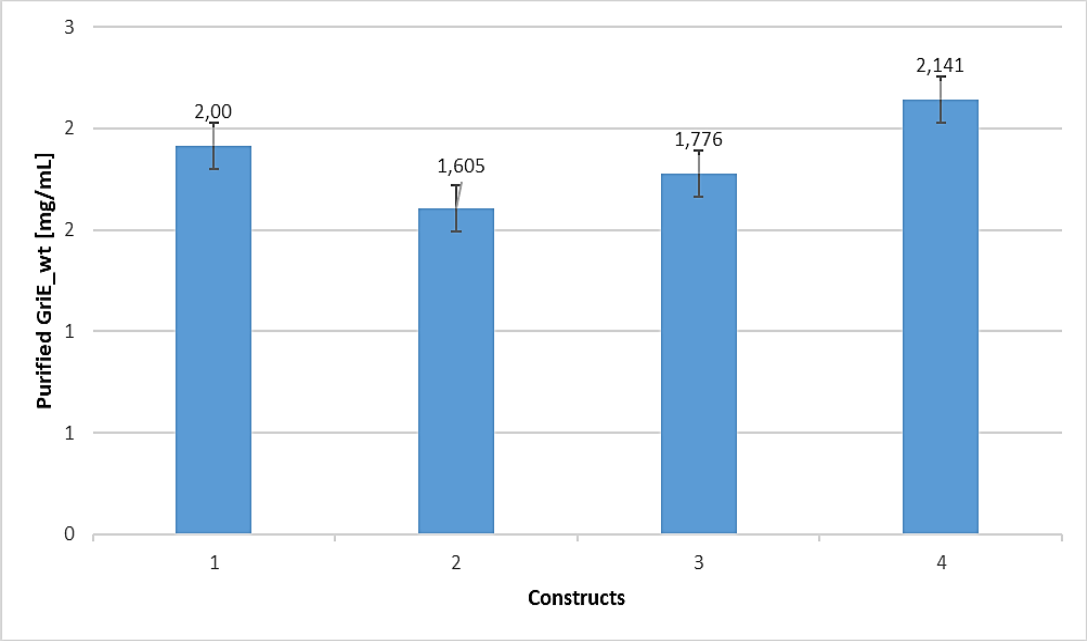


Figure 21. Comparison of protein concentrations [mg mL^{-1}] of GriE_wt overexpressed in different plasmid systems [pET28a(+) and pASKGATE], given by protein purification and determination of protein concentration by BCA assay. [1- pET28a(+)_GriE_wt induced by 1.0 mM IPTG; 2- pET28a(+)_GriE_wt induced by 0.1 mM IPTG; 3- pASKGATE_His_Leu_GriE_wt induced by 50 µg L^{-1} ATET; 4- pASKGATE_His_Leu_GriE_wt induced by 100 µg L^{-1} ATET.]

Comparison of the GriE_wt overexpression in two plasmids [pET28a(+) and pASKGATE] indicated that the expression induced by higher concentration of IPTG (1.0 mM) and ATET (100 µg L^{-1}) both resulted in a higher amount of expressed GriE_wt protein (2.0 mg mL^{-1} and 2.141 mg mL^{-1}, respectively), compared to the expression induced by 0.1 mM IPTG and 50 µg L^{-1} ATET (1.605 mg mL^{-1} and 1.776 mg mL^{-1}, respectively).

pET28a(+)_GriE_wt induced by 1.0 mM IPTG resulted in the concentration 2.0 ± 0.67 mg mL^{-1} of the protein, while the overexpression of pET28a(+)_GriE_wt induced by 0.1 mM IPTG resulted in a concentration of 1.605 ± 0.304 mg mL^{-1} of the protein. Overexpression of pASKGATE_His_Leu_GriE_wt induced by 50 µg L^{-1} ATET resulted in the production of 1.776 ± 1.080 mg mL^{-1} of the protein and overexpression of pASKGATE_His_Leu_GriE_wt induced by 100 µg L^{-1} ATET resulted in 2.141 ± 0.927 mg mL^{-1} of the protein.

pET28a(+) and pASKGATE plasmid systems showed a comparable amount of purified GriE_wt, because of tightly regulated high level protein expression allowed by both T7 and *tetA* promoters (Bertram and Hillenthe, 2008; Novagen, 2003; Skerra, 1994; Studier and Moffatt, 1986).
Having in mind above described results (similar concentrations of expressed protein with both plasmids) and that neither plasmid showed any kind of advantage during protein expression (such as higher concentration of expressed protein), distinctions that could be considered for potential future studies of GriE_wt expression between these two plasmids [pET28a(+) and pASKGATE], can be that for the expression with *tetA* promoter any strain of *E. coli* can be used, while for T7 promoter it is necessary to use the strain with λDE3 lysogen or strain supplemented with exogenous T7 RNA polymerase (Skerra, 1994).
Additionally, following optimization and improvement of expression of soluble GriE_1 (mutant with incorporated NMH at the location H110) was performed in both plasmids, pET28a(+) and pASKGATE.

4.2.3. Optimization of the expression of soluble GriE_1 in small scale

As already mentioned in the Literature review (see Chapter 2.2.1.1.) and in previous discussion regarding the expression results (see Chapter 4.2.1.), one of the bottlenecks of the stop codon suppression is the competition between the release factor and the successful incorporation of non-canonical amino acid of interest at the amber stop codon. Termination of translation results in inactive, truncated protein fragments which affect the overall concentration of expressed protein. The goal of the soluble expression optimization was to coordinate the translation rate of pEVOL (plasmid which carries the gene encoding for aminoacyl-tRNA synthetase / tRNA pair for NMH incorporation) and translational rate of GriE_1 in either pET28a(+) or pASKGATE. Different concentrations of arabinose; IPTG and ATET as well as higher concentration of NMH (6.0 mM) were examined.

4.2.3.1. GriE_1 gene expression from pET28a(+)_GriE_1

Combinations of different concentrations IPTG, arabinose (ara) and NMH tested for the improvement of expression of soluble pET28a(+)_GriE_1 are listed in Table 22. The expression was investigated in *E. coli* BL21 (DE3) with pEVOL_NMH (see Chapter 3.1.4.) and performed in a small scale, as described in subchapter 3.2.10.5. The SDS-PAGE analysis of the insoluble and soluble fractions of cell extracts, taken during different time-points during pET28a(+)_GriE_1 overexpression, can be seen in the Figure 22 -Figure 25.

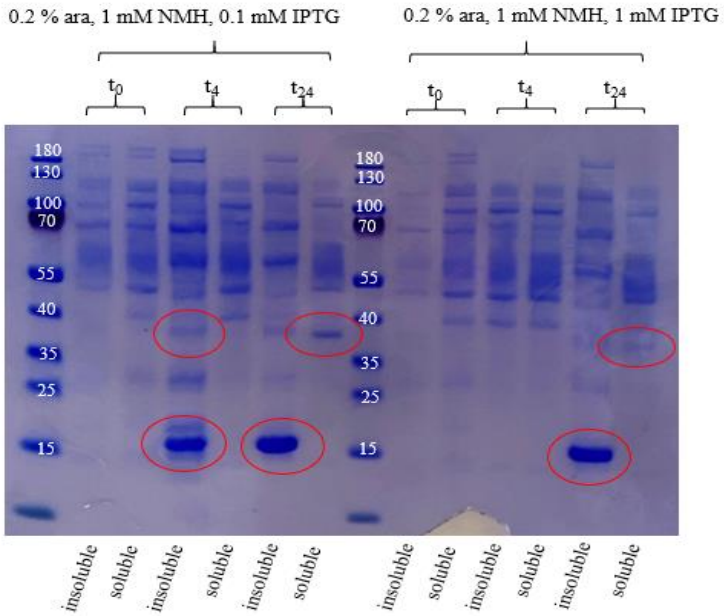


Figure 22. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pET28a(+)_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pET28a(+)_GriE_1 / pEVOL_NMH was induced with: (a) 0.1 mM IPTG and 0.2 % ara, respectively and supplemented by 1.0 mM NMH, and (b) 1.0 mM IPTG and 0.2 % ara, respectively and supplemented by 1.0 mM NMH. Expected size of GriE_1 (highlighted by red circles) was approximately 30.1 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.68 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

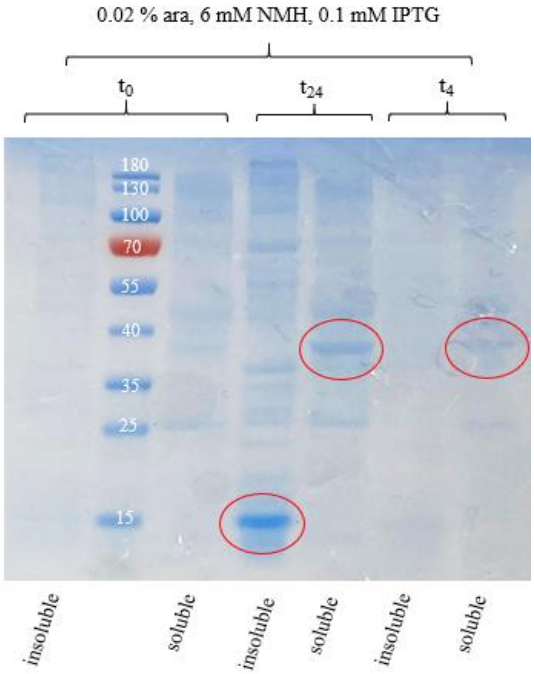


Figure 23. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pET28a(+)_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pET28a(+)_GriE_1 / pEVOL_NMH was induced with 0.1 mM IPTG and 0.02 % ara, respectively and supplemented by 6.0 mM NMH. Expected size of GriE_1 (highlighted

by red circle) was approximately 30.1 kDa, while size of truncated protein (highlighted by red circle) was approximately 12.68 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

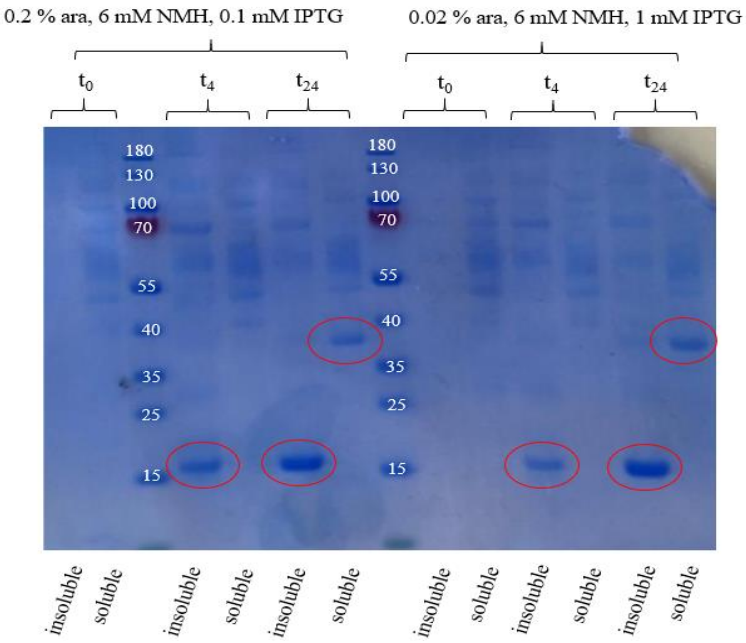


Figure 24. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pET28a(+)_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pET28a(+)_GriE_1 / pEVOL_NMH was induced with: (a) 0.1 mM IPTG and 0.2 % ara, respectively and supplemented by 6.0 mM NMH, and (b) 1.0 mM IPTG and 0.02 % ara, respectively and supplemented by 6.0 mM NMH. Expected size of GriE_1 (highlighted by red circles) was approximately 30.1 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.68 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

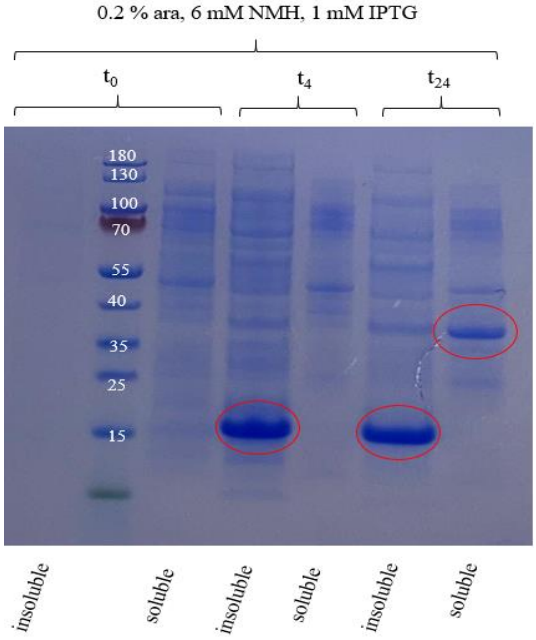


Figure 25. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pET28a(+)_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pET28a(+)_GriE_1 / pEVOL_NMH was induced with 1.0 mM IPTG and 0.2 % ara, respectively and supplemented by 6.0 mM NMH. Expected size of GriE_1 (highlighted

by red circle) was approximately 30.1 kDa, while size of truncated protein (highlighted by red circle) was approximately 12.68 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

For all six approaches listed in Table 22 and shown by Figure 22 -Figure 25, high concentrations of the truncated GriE_1 protein (expected molecular weight was 12.68 kDa, qualitatively shown at SDS-PAGE gels in approximately size of 15 kDa) were visible in the insoluble fractions. Especially for t_{24} samples, truncation is clearly visualized by large bands approximately in the size of 15 kDa. This indicates that the transcription of pEVOL_NMH is not as high as the transcription rate of pET28a(+)_GriE_1. Accordingly, NMH is not incorporated at the position of amber stop codon (H110NMH) in GriE_1, but termination of translation manifests (at location H110) and therefore the expression product is truncated, which approximately corresponds to the size of 12.68 kDa (starting from N-terminus to the stop codon). Furthermore, the insoluble fractions suggest inclusion body formation due to incorrect folding of polypeptides.

In the soluble fractions of all t_{24} sample, bands corresponding to the size of correctly folded protein (30.1 kDa) were obtained, indicating synthesis of some correctly folded protein with successfully integrated NMH. The highest concentration of full-length protein in the soluble phase was detected by inducing expression of pET28a(+)_GriE_1 with 1.0 mM IPTG, pEVOL_NMH by 0.2 % arabinose and supplemented by 6.0 mM of NMH (Figure 25).

From these experiments it can be concluded that a higher concentration of arabinose and added NMH, result in successful incorporation of NMH in the active triad of the protein. Even higher concentrations of soluble GriE_1, could be achieved by the addition of higher concentrations of NMH, *e.g.* 12 mM, as reported by Green et al. (2016). This was however not considered in this study, due to the price of NMH.

Furthermore, to further improve soluble protein expression, the cultivation temperature could be reduced to 4 – 12 °C using the *E. coli* Arctic Express. Lowering cultivation temperature would cause lower translational rate and therefore slower but more correct folding of the polypeptides.

4.2.3.2. GriE_1 gene expression from pASKGATE_His_Leu_GriE_1

One of the problems, which arose with pASKGATE_His_Leu_GriE_1, was low cell growth of *E. coli* BL21 (DE3) with pEVOL_NMH. This problem most likely arose due to the metabolic burden of having two plasmids in the host system: pASKGATE_His_Leu_GriE_1 and pEVOL_NMH. Interestingly, the same was not experienced while expressing pET28a(+)_GriE_1 in *E. coli* BL21 (DE3) with pEVOL_NMH.

Both plasmids [pET28a(+)_GriE_1 and pASKGATE_His_Leu_GriE_1] are considered small in size (4,240 bp and 6,121 bp, respectively). Additionally, there was no difference in *E. coli* BL21 (DE3) growth detected while expressing GriE_wt in each plasmid. Potential difference in cell growth between pASKGATE_His_Leu_GriE_1 and pET28a(+)_GriE_1 could be necessity for different antibiotic. pASKGATE carries the resistance for spectinomycin and combination of spectinomycin with chloramphenicol (which resistance carries pEVOL_NMH) could affect slow growth. Additionally, *tetA* promoter of pASKGATE plasmid, as experienced during studies, seemed to be so strong that almost no additional growth of *E. coli* BL21 (DE3) was observed after induction.

This problem was conquered by extending the incubation time of the over-night culture to approximately 25 h and using half of the antibiotic concentration (25 µg mL^{-1} of spectinomycin and 25 µg mL^{-1} chloramphenicol) in the LB medium. In addition, high OD_{600} induction was performed, which means that overexpression was induced at the end of the exponential phase or at the beginning of stationary *E. coli* BL21 (DE3) growth phase.

Combinations of different inducers (ATET and arabinose) and NMH concentrations are listed in Table 23. Expression was investigated in *E. coli* BL21 (DE3) with pEVOL_NMH (see Chapter 3.1.4.) and performed in small scale as described in subchapter 3.2.10.5. Expressions of GriE_1 from pASKGATE - induced by different concentrations of ATET, supplemented with different concentrations of NMH, and in addition of different concentrations of arabinose (which acts as inducer for pEVOL_NMH necessary for incorporation of NMH) - are shown by Figure 26 - Figure 31.

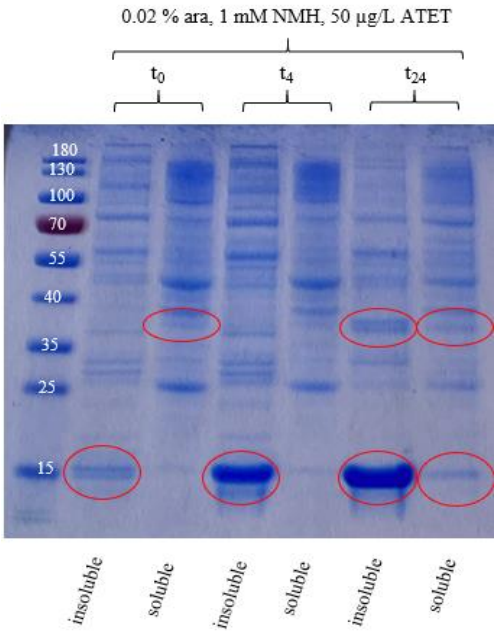


Figure 26. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH was induced with 50 µg L^{-1} ATET and 0.02 % arabinose, respectively, and supplemented by 1.0 mM NMH. Expected size of GriE_1 (highlighted by red circle) was approximately 30.2 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.8 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

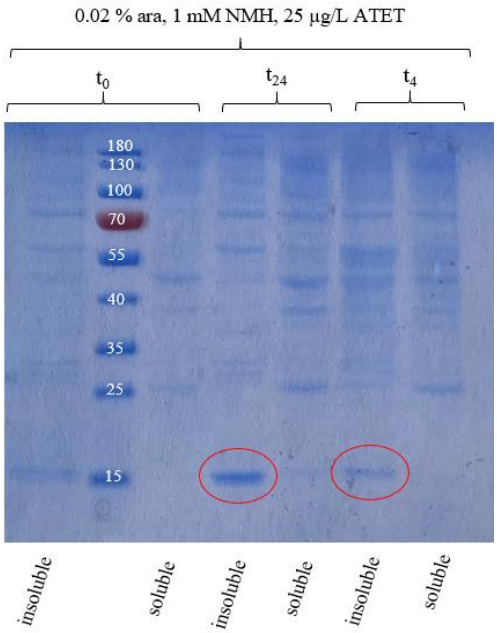


Figure 27. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH was induced with 25 µg L^{-1} ATET and 0.02 % arabinose, respectively, and supplemented by 1.0 mM NMH. Expected size of GriE_1 was approximately 30.2 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.8 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

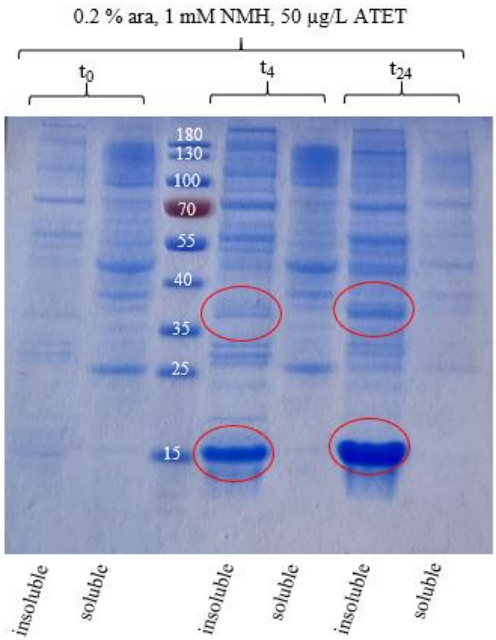


Figure 28. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH was induced with 50 µg L^{-1} ATET and 0.2 % arabinose, respectively, and supplemented by 1.0 mM NMH. Expected size of GriE_1 (highlighted by red circle) was approximately 30.2 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.8 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

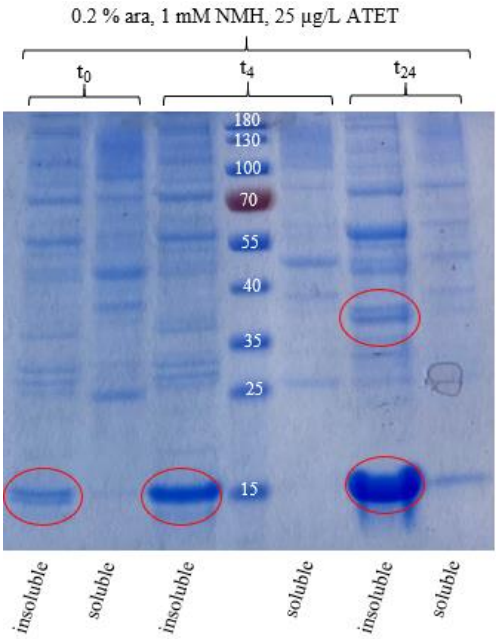


Figure 29. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH was induced with 25 µg L^{-1} ATET and 0.2 % arabinose, respectively, and supplemented by 1.0 mM NMH. Expected size of GriE_1 (highlighted by red circle) was approximately 30.2 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.8 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

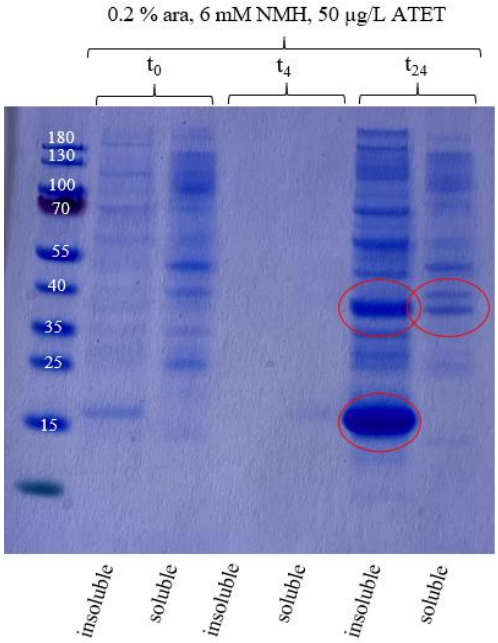


Figure 30. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH was induced with 50 µg L^{-1} ATET and 0.2 % arabinose, respectively, and supplemented by 6.0 mM NMH. Expected size of GriE_1 (highlighted by red circles) was approximately 30.2 kDa, while size of truncated protein (highlighted by red circle) was approximately 12.8 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

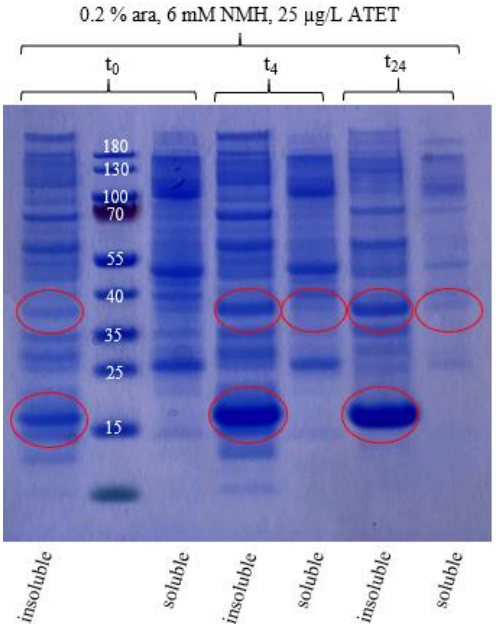


Figure 31. SDS-PAGE analysis of cell extracts of *E. coli* BL21 (DE3) transformed by pASKGATE_His_Leu_GriE_1 taken at different time points (t0, t4, t24). Overexpression of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH was induced with 25 µg L^{-1} ATET and 0.2 % arabinose, respectively, and supplemented by 6.0 mM NMH. Expected size of GriE_1 (highlighted by red circles) was approximately 30.2 kDa, while size of truncated protein (highlighted by red circles) was approximately 12.8 kDa. Marker: PageRuler™ Prestained Protein Ladder (Thermo Fisher Scientific, USA).

For all of the tested conditions for the expression of pASKGATE_His_Leu_GriE_1 (listed in Table 23 and shown by Figure 26 - Figure 31), results were similar as for pET28a(+)_GriE_1 (see Chapter 4.2.3.1.). Most of the GriE_1 protein from pASKGATE_His_Leu_GriE_1 at t_{24} samples is in the insoluble phase and in a truncated form (12.8 kDa, Figure 26 - Figure 31). Indicating incorrect protein folding and termination of translation due to failed NMH incorporation. Moreover, the highest concentrations of fully translated protein (30.2 kDa) were achieved by the induction of pEVOL_NMH by 0.2 % arabinose, induction of pET28a(+)_GriE_1 by 25 µg L^{-1} ATET (Figure 30) and 50 µg L^{-1} ATET (Figure 31) and in the addition of 6.0 mM of NMH. Likewise, the biggest improvement of expression of soluble full-length protein, corresponding to the size of 30.2 kDa, can be seen while inducing the overexpression by 0.2 % arabinose and 50 µg L^{-1} ATET and with addition of 6.0 mM NMH (Figure 30).

For two examples of overexpression in both of which pEVOL_NMH was induced by 0.2 % arabinose and pASKGATE_His_leu_GriE_1 with 50 µg L^{-1} ATET, but supplemented with different concentrations of NMH (first example; 1 mM of NMH, Figure 28 and secondly 6 mM of NMH, Figure 30), SDS-PAGE analysis showed wider and ticker band in insoluble fraction and in truncated form for the expression in which LB medium was supplemented with 6 mM NMH (Figure 30). Synthesis of the truncated GriE_1 protein might be seen as consequence of low NMH concentration, insufficient for the successful incorporation by stop codon suppression.

To summarize, besides higher aminoacyl-tRNA / tRNA pair expression important for the incorporation of NMH, concentration of available NMH also plays an important role (see Chapter 4.2.3.). Concentrations of NMH higher than 6.0 mM were not investigated due to the price of NMH.

There are few possibilities to overcome above mentioned problems. For potential future studies, cloning of the 6xHis tag on the C-terminus could help to improve the purification of the fully-translated protein, because in that case truncated protein should not be purified. Likewise, reducing the rate of protein synthesis can be achieved by lowering cultivation temperature to 4 - 12 °C by using *E. coli* Arctic Express as the host system. Moreover, the concentration of ATET could be further reduced (5 – 10 µg L^{-1} ATET) compared to the concentrations used in this study (25 – 50 µg L^{-1} ATET). Re-folding of the protein from inclusion bodies may also be considered, however, this strategy could have disadvantages such as poor yields and disturbance of the integrity of the refolded protein (Sorensen and Mortenses, 2005).

4.3. ACTIVITY ASSAY

4.3.1. Enzymatic reaction

In order to achieve the aim (d) of this Graduate Thesis - to determine what are the differences, if there are any, in the substrate scope and enzymatic activity of the GriE_1 mutant in comparison to the GriE_wt, [expressed usi ng either pET28a(+) or pASKGATE], activity assays were performed (see Chapter 3.2.14.). Hydroxylation product was detected by TLC and ninhydrin staining (see Chapter 3.2.14.2.). Reactions were set up with (a) purified enzyme with L-leucine and L-isoleucine as a substrate or (b) with the cell lysate containing GriE_1 enzyme with L-leucine and L-isoleucine as substrates.

4.3.2. Determination of activity by thin-layer chromatography

4.3.2.1. Activity assay with pET28a(+) and pASK37(+)s as negative control

Negative control of activity assays with empty vectors [pET28a(+) and pASK37(+)s] was performed in order to see if the potential conversion of substrates derives from enzyme of interest, GriE and to be able to claim that potential development of stain different from substrate or standard is derived from GriE and not caused by the any component of the enzymatic reaction setup. pET28a(+) empty vector was induced with 0.1 mM and 1.0 mM IPTG, and pASK37(+)S was induced with 50 µg L^{-1} and 100 µg L^{-1} ATET. Empty plasmids did not convert the substrate L-leucine (Appendices 9 and Appendices 10), that conversion derives from the enzyme of interest (GriE_wt or GriE_1).

4.3.2.2. Activity assays of GriE_wt expressed from pET28a(+) and pASKGATE

Enzymatic reactions of GriE_wt (see Chapter 4.2.1.) were performed with purified enzyme. The hydroxylation of the natural substrate (L-leucine) to product (5-hydroxyleucine) was confirmed by TLC and staining of enzymatic reaction applied on silica gel with ninhydrin (Figure 32 and 33). While as expected GriE_wt do not catalyze hydroxylation of the L-isoleucine to the product 5-hydroxyisoleucine (data not shown).

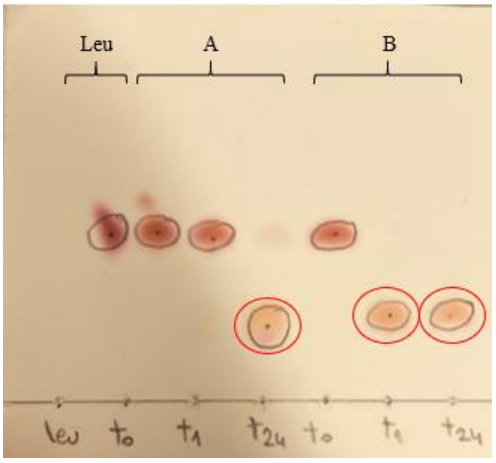


Figure 32. TLC of purified GriE_wt expressed from pET28a(+). Time sample analysis (t_0, t_1, t_{24}) of overexpression induced by [A] 0.1 mM and [B] 1.0 mM IPTG; L-leucine served as a reference and a substrate. Spots highlighted by red circles indicated formation of 5-hydroxyleucine. Spotting on TLC was detected with ninhydrin staining.

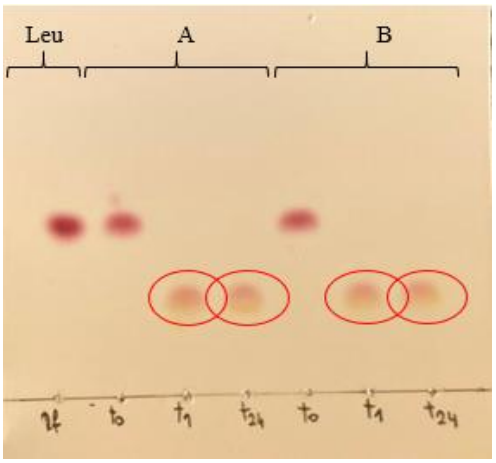


Figure 33. TLC of purified GriE_wt expressed from pASKGATE. Time sample analysis (t_0, t_1, t_{24}) of overexpression induced by [A] 50 µg L^{-1} and [B] 100 µg L^{-1} ATET; L-leucine served as a reference and a substrate. Spots highlighted by red circles indicated formation of 5-hydroxyleucine. Spotting on TLC was detected with ninhydrin staining.

Product formation catalyzed by GriE_wt expressed from pET28a(+) was detectable in t_{24} samples (Figure 32), while full conversion of the substrate with GriE_wt expressed from pASKGATE was

detected in one hour time range (Figure 33). This comparison indicated that the GriE_wt expressed in pASKGATE plasmid showed faster conversion of the substrate than the GriE_wt expressed in pET28a(+). Furthermore, these reactions were performed at the RT because this activity assay was also implemented in another study, while full conversion with GriE_wt, based on the literature data, can be achieved by incubation at 35 °C (Hanreich, 2019).

4.3.2.3. Activity assay of GriE_1 expressed from pET28a(+) and pASKGATE

Enzymatic reactions of cell lysate containing GriE_1 (expressed by the optimized soluble expression using different concentrations of IPTG/ATET, arabinose and NMH) were performed and corresponding activity was detected by TLC and ninhydrin staining (see Chapter 3.2.14.).

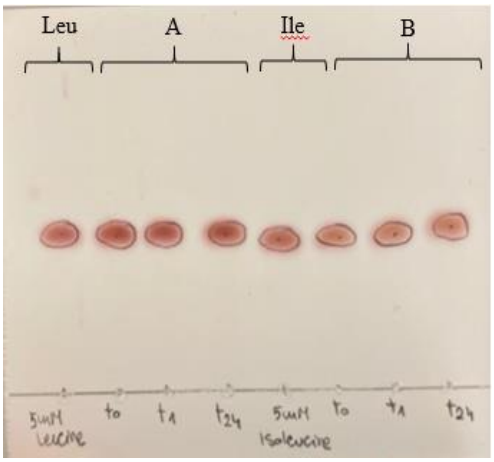


Figure 34. TLC of cell lysate containing GriE_1 expressed from pET28a(+). Time sample analysis (t_0, t_1, t_{24}) of activity assay was performed with the GriE_1 enzyme overexpressed by induction of pET28a(+)_GriE_1 / pEVOL_NMH by 1.0 mM IPTG and 0.2 % arabinose, respectively, and with addition of 6.0 mM NMH. Substrate of the assay was [A] L-leucine and [B] L-isoleucine. Both substrates served as a references. Spotting on TLC was detected with ninhydrin staining.

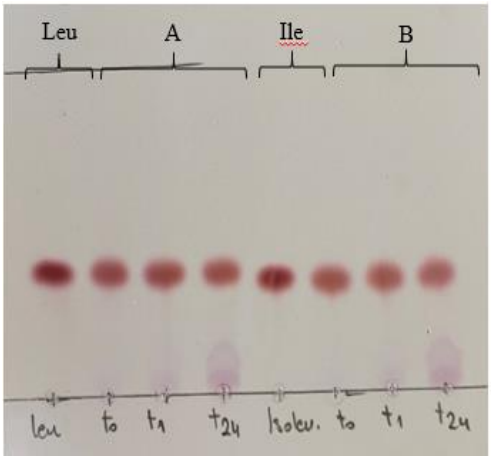


Figure 35. TLC of cell lysate containing GriE_1 expressed from pASKGATE. Time sample analysis (t_0, t_1, t_{24}) of activity assay was performed with the GriE_1 enzyme, overexpressed by induction of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH by 50 µg L^{-1} ATET and 0.2 % arabinose, respectively, and with addition of 6.0 mM NMH. Substrate of the assay was [A] L-leucine and [B] L-isoleucine. Both substrates served as a references. Spotting on TLC was detected with ninhydrin staining.

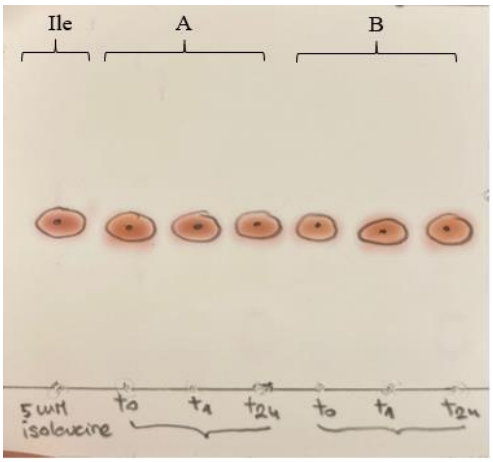


Figure 36. TLC of lysate containing GriE_1 expressed from pASKGATE. Time sample analysis (t_0, t_1, t_{24}) of activity assay was performed with the GriE_1 enzyme, overexpressed by induction of pASKGATE_His_Leu_GriE_1 / pEVOL_NMH by 25 µg L^{-1} ATET and 0.2 % arabinose, respectively, and with addition of 6.0 mM NMH. Substrate of the assay was [A] L-leucine and [B] L-isoleucine. Both substrates served as a references. Spotting on TLC was detected with ninhydrin staining.

Table 29. Overview of GriE expression systems, conditions of (over)expression, types of enzyme preparations, substrates, and quantitative results of TLC enzyme assay.

GriE expression system	Concentrations of inducers and/or NMH*	Enzyme preparation**	Substrate	TLC of enzyme assay ***
pET28a(+)_ GriE_wt	0.1 mM IPTG	Purified	L-leucine	+
	1.0 mM IPTG	Purified	L-leucine	+
pASKGATE_His_Leu_ GriE_wt	50 µg L^{-1} ATET	Purified	L-leucine	+
	100 µg L^{-1} ATET	Purified	L-leucine	+
pET28a(+)_GriE_1	1.0 mM IPTG, 0.2 % ara, 6.0 mM NMH	Cell lysate	L-leucine	−
		Cell lysate	L-isoleucine	−
pASKGATE_His_Leu_ GriE_1	50 µg L^{-1} ATET, 0.2 % ara, 6.0 mM NMH	Cell lysate	L-leucine	−
		Cell lysate	L-isoleucine	−
pASKGATE_His_Leu_ GriE_1	25 µg L^{-1} ATET, 0.2 % ara, 6.0 mM NMH	Cell lysate	L-leucine	−
		Cell lysate	L-isoleucine	−

*see Table 22 and Table 23,

**concentration of purified GriE_wt enzyme used in enzymatic reactions was 2.0 mg mL^{-1}, while concentration of GriE_1 enzyme in cell lysate was 1.7 mg mL^{-1},

*** active + ; inactive -.

Concentration of GriE_1 in cell lysate, expressed either from pET28a(+) or pASKGATE, was used for the reactions was 1.7 mg mL^{-1} (as measured using BCA assay). After negative results were detected on TLC, additional reactions were repeated with 8.5 mg $mL^{-1}$and 17 mg mL^{-1} of protein (data not shown) which as well did not result in the conversion of L-isoleucine. Concentrations of the GriE_1 were increased (8.5 and 17 mg mL^{-1}) in enzymatic reactions, because cell lysate, and not purified enzyme, was used for the enzymatic reaction and as described in subchapter 4.2.3., only small amount of correctly folded protein could be expected (Figure 25, Figure 30 - 31).

Even though activity of pET28a(+)_GriE_1 and/or pASKGATE_His_Leu_Grie_1 towards L-leucine and/or L-isoleucine, was not detected (Table 29), it cannot be claimed that the L-leucine and L-isoleucine are not substrates of the GriE_1 - considering the problem of soluble expression, the fact that cultivation was done in a small scale and that activity assays of optimized expressions were performed using cell lysate which possibly did not contain enough correctly folded protein that could convert the substrate. Additionally, having in mind that the mutant does not convert the

L-leucine which is substrate of the wild type - it means that the future studies must be performed. Moreover, future studies also should be performed using HPLC or LC-MS.

TLC is qualitative method in which detection is based on the visual spotting and with which common problems could be experienced, such as uneven advance of solvent front. However, HPLC and LC-MS are quantitative methods which can detect low concentration of converted substrate with high accuracy. These methods are also easily evaluated, while control parameters are easily and highly controlled.

In this study, NMH was successfully incorporated in the active site of AvLDO at both locations (H151NMH and H237NMH), while NMH was already successfully introduced in the active site of GriE (H110NMH, GriE_1 mutant). To the best of our knowledge, this is one of the first examples of the incorporation of NMH in the active site of non-heme dioxygenase. Accordingly, there are also some successful cases of incorporation NMH in heme enzymes (Pott et al., 2018; Green et al., 2016).

In order to investigate protein expression of GriE_1, firstly comparison of the expression of soluble GriE_wt within two plasmids [pET28a(+) and pASKGATE] was done. Additionally, the expression of soluble mutant (GriE_1) was optimized. Even though literature shows great examples of successful incorporation of different non-canonical amino acids in the protein sequence by stop codon suppression, this work showed, that under non-ideal conditions, it can also result in the truncation of the protein. In this study, undesired problems with truncation and expression of soluble enzymes needed to be solved. Improvement of the expression of soluble GriE_1 in two vectors, pET28a(+) and pASKGATE, was achieved by using 0.2 %arabinose (pEVOL_NMH), in addition of 6.0 mM NMH and while induced by 1.0 mM IPTG [pET28a(+)] or by 50 µg L^{-1} ATET [pASKGATE].

The performed activity assays using TLC, did not show any conversion of the substrates (L-leucine and L-isoleucine) by cell lysate containing GriE_1. The next step of the future studies would be to purify GriE_1 and repeat the experiments. Additionally, a more sensitive method should be established, in order to verify the activity. This should also be considered when determining any rate differences between wild type and mutant in the future.

Pott et al. (2018) achieved promiscuous peroxidase activity by significant changes in the proximal pocket by incorporation of N-methylhistidne in the heme myoglobin, and Green et al. (2016) successfully constructed ascorbate peroxidase with N-methylhistidine, which showed similar catalytic efficiency as wild type but with much higher turnover number. With the improvement of

soluble expression and more sensitive assay techniques, an improvement of catalytic properties might be notable for α-ketoglutarate, non-heme Fe(II) dioxygenases as well.

5. CONCLUSIONS

According to the results presented in this Graduate Thesis, several conclusions can be drawn, as follows.

(a) Cloning of wild type (wt) gene coding for L-leucine 5-hydroxylase from *Anabaena variabilis* (*AvLDO_wt*) in pASKGATE plasmid was successful. Additionally, site-directed mutagenesis of *AvLDO_wt* successfully resulted in the cloning of two *AvLDO* mutants - *AvLDO_1* (H151N-methylhistidine, NMH) and *AvLDO_2* (H237NMH) in pASKGATE.
AvLDO_wt was expressed in *E. coli* BL21 (DE3) strain, while mutants were expressed in double transformed *E. coli* BL21 (DE3) with pEVOL_NMH. Moreover, AvLDO_2 was expressed at the same level as the wild type, while overexpression of AvLDO_1 was detected at a substantially lower level. Qualitative analysis pointed out that the place of incorporation of the NMH at the location H237 (AvLDO_2) may be more favorable compared to the incorporation at the H151NMH (AvLDO_1). However, both wild type and the *AvLDO* mutants were expressed in the insoluble fraction, indicating inclusion body formation.

(b) The expression of soluble L-leucine 5-hydroxylase from *Streptomyces muensis* (GriE_wt) was investigated in two plasmid systems [pET28a(+) and pASKGATE]. Overexpression of GriE_wt induced by different concentrations IPTG [0.1 and 1.0 mM for pET28a(+)_GriE_wt] and ATET [50 and 100 $\mu g\ L^{-1}$ for pASKGATE_His_Leu_GriE_wt] resulted in approximately the same concentration of purified recombinant protein, 2.0 $mg\ mL^{-1}$.

(c) Guided by above-mentioned results, regarding the same translational rate of the different plasmids - optimization of expression of soluble mutant GriE_1 with incorporated non-canonical amino acid (H110NMH) was therefore performed in both vectors [pET28a(+) and pASKGATE]. Optimization and improvement of the soluble expression of GriE_1 was successfully achieved in both plasmids [pET28a(+) and pASKATE] by using different concentrations of inducers (IPTG or ATET and arabinose) and different concentrations of supplemented NMH. The addition of NMH (6.0 mM) and induction by arabinose (0.02 %) resulted in a stronger induction of pEVOL_NMH plasmid and consequently in a more efficient incorporation of NMH, which can be seen as a higher concentration of correctly folded (soluble) GriE_1 protein. Moreover, concentration IPTG and ATET of 1.0 mM and 25 $\mu g\ L^{-1}$, respectively, were necessary to induce the translation of the mutant GriE_1.

(d) GriE_wt from both plasmids [pET28a(+) and pASKGATE] showed conversion of L-leucine. While GriE_1 showed inactivity towards L-leucine and L-isoleucine (L-Ile). Nevertheless, it still cannot be claimed that the L-Ile is not the substrate of the GriE_1. The facts that cultivation of double transformed *E. coli* BL21 (DE3) with pEVOL was done in a small scale, the expression of soluble GriE_1 occurred, and that activity assays of optimized expressions were performed using cell lysate, which possibly did not contain enough correctly folded protein that could convert the substrate - affected the enzymatic efficiency and overall results of enzymatic activity .

6. REFERENCE

Anonymous 1 (2017) <https://blog.addgene.org/plasmids-101-gateway-cloning> . Accessed on the 26th of February 2021.

Anonymous 2 (2021) <https://www.thermofisher.com/hr/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html>. Accessed on the 18th of January 2021.

Anonymous 3 (2021) <https://international.neb.com/protocols/0001/01/01/pcr-protocol-m0530>. Accessed on the 18th of March, 2021.

Anonymous 4 (2021) <https://international.neb.com/protocols/2012/09/27/protocol-for-a-pcr-reaction-using-hot-start-taq-2x-master-mix-m0496>. Accessed on the 19th of March, 2021.

Anonymous 5 (2021) <https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0013117_GeneJET_Plasmid_Miniprep_UG.pdf&title=VXNlciBHdWlkZTogR2VuZUpFVCBQbGFzbWlkIE1pbmlwcmVwIEtpdA==>. Accessed on the 19th of March, 2021.

Anonymous 6 (2021) <http://www.prep-hplc.com/Uploads/ueditor/file/20190731/5d41050fbe33d.pdf>. Accessed on the 3rd of March, 2021.

Anonymous 7 (2021) <https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0011430_Pierce_BCA_Protein_Asy_UG.pdf>. Accessed on the 19th of March, 2021.

Anonymous 8 (2021) <https://www.neb-online.fr/en/neb-en/cloning-synthetic-biology/kits/q5-site-directed-mutagenesis-kit/>. Accessed on the 3rd of March, 2021.

Ambrogelly, A., Palioura, S., Söll, D. (2007) Natural expansion of the genetic code. *Nat. Chem. Biol.* **3**, 29–35.

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

Agostini, F., Völler, J., Koksch, B., Acevedo-Rocha, C., Kubyshkin, V., Budisa, N. (2017) Biocatalysis with Unnatural Amino Acids: Enzymology Meets Xenobiology. *Angew. Chem. Int.* **56**, 9680-9703.

Bollinger, J., Price, J., Hoffart, L., Barr, E., Krebs, C. (2005) Mechanism of Taurine: α-Ketoglutarate Dioxygenase (TauD) from *Escherichia coli*. *Eur. J. Inorg. Chem*. **2005**, 4245-4254.

Burke, A. J., Lovelock, S. L., Frese, A., Crawshaw, R., Ortmayer, M., Dunstan, M., Levy, C., Green, A. P. (2019) Design and evolution of an enzyme with a non-canonical organocatalytic mechanism. *Nature*. **570**, 219-223.

Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., Schultz, P. G. (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in *Escherichia coli*. *Biochemistry*. **52**, 1828-1837.

Chen, H., Bong, Y. K., Cabirol, F. L., Prafulchandra, A. G., Li, T., Moore, J. C., Quintanar-Audelo, M., Hong, Y., Collier, S. J., Smith, D. (2017) Biocatalysts and Methods for Hydroxylation of Chemical Compounds. U.S. Patent US2017/0121744 A1.

Chin, J. W. (2014) Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annu. Rev. Biochem*. **83**, 379-408.

Chin, J. W. (2017) Expanding and reprogramming the genetic code. *Nature*. **550**, 53-60.

Chin, J. W., Santoro, S.W., Martin, A. B., King, D. S., Wang, L., Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **124**, 9026-9027.

Correira Cordeiro, R. S., 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.

Deepankumar, K., Shon, M., Nadarajan, S. P., Shin, G., Mathew, S., Ayyadurai, N., Kim, B. G., Choi, S. H., Lee, S. H., Yun, H. (2014) Enhancing thermostability and organic solvent tolerance of ω-Transaminase through global incorporation of fluorotyro-sine. *Adv. Synth. Catal*. **356**, 993–998.

Drienovska, I., Rioz-Martínez, A., Draksharapua, A., Roelfes, G. (2015) Novel artificial metalloenzymes by in vivo incorporation of metal-binding unnatural amino acids. *Chem. Sci*. **6**, 770-776.

Drienovska, I., Gajdoš, M., Kindler, A., Takhtehchian, M., Darnhofer, B., Birner-Gruenberger, R., Dörr, M., Bornscheuer, U., Kourist, R. (2020) Folding Assessment of Incorporation of Noncanonical Amino Acids Facilitates Expansion of Functional-Group Diversity for Enzyme Engineering. *Chem. Eur. J.* **26**, 12338-12342.

Drienovska, I., Roelfes, G. (2020) Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nat. Catal*. **3**, 193-202.

Edelheit, O., Hanukoglu, A., Hanukoglu, I. (2009) Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol*. ***9***, 1-8.

Fekner, T., Chan, M. K. (2011) The pyrrolysine translational machinery as a genetic-code expansion tool. *Curr. Opin. Chem. Biol.* **15**, 387-391.

Frey, R., Hayashi, T., Buller, R. M. (2019) Directed evolution of carbon–hydrogen bond activating enzymes. *Curr Opin Biotechnol*. **60**, 29-38.

García-Fruitós, E. (2010) Inclusion bodies: a new concept. *Microb. Cell Fact.* ***9***, 1-3. doi: https://doi.org/10.1186/1475-2859-9-80

Green, A. P., Hayashi, T., Mittl, P. R. E., Hilvert, D. (2016) A Chemically Programmed Proximal Ligand Enhances the Catalytic Properties of a Heme Enzyme. *J. Am. Chem. Soc.* **138**, 11344–11352.

Hanreich, S. (2019) Identification and selection of active α-ketoglutarate dependent dioxygenases in engineered *Escherichia coli.* (Graduate Thesis) Graz University of Technology, Graz, Austria.

Hausinger, R., Schofield, C. (2015) 2-Oxoglutarate-Dependent Oxygenases, No. 3, The Royal Society of Chemistry Metallobiology, Cambridge.

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

Hegg, E. L., Que, L. Jr. (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, C. Q., Hausinger, R. P. (2018) Amazing Diversity in Biochemical Roles of Fe(II)/2 Oxoglutarate Oxygenases. *Trends Biochem. Sci.* **43**, 517-532.

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. Environ. Microbiol.* **98**, 3869-3876.

Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M., Krebs, C. (2006) Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc. Natl. Acad. Sci.* **103**, 14738-14743.

Hoffmann, F., Rinas, U. (2004). Stress Induced by Recombinant Protein Production in *Escherichia coli. Adv. Biochem. Eng. Biotechnol.* **89**, 73-92.

Hoffmann, J. E., Dziuba, D., Stein, F., Schultz, C. (2018) A Bifunctional Non-canonical Amino Acid: Synthesis, Expression, and Residue-Specific Proteome-wide Incorporation. *Biochemistry*. **57**, 4747-4752.

Holzgrabe, U. (2015) New Griselimycins for Treatment of Tuberculosis. *Cell Chem. Biol.*, **22**, 981-982.

Huttel, W. (2013) Biocatalytic production of chemical building blocks in technical scale with α-ketoglutarate-dependent dioxygenases. *Chem. Ing. Tech.* **85**, 809–817.

Ibba, M., Söll, D. (2000) Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem*. **69**, 617–650.

Ibba, M., Söll, D. (2004) Aminoacyl-tRNAs: setting the limits of the genetic code. *Genes Dev.* **18**, 731–738.

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

Katzen F. (2007) Gateway(®) recombinational cloning: a biological operating system. *Expert Opin. Drug. Discov*. **2**, 571-589.

Kling, A., Lukat, P., Almeida, D., Bauer, A., Fontaine, E., Sordello, S., Zaburannyi, N., Herrmann, J., Wenzel, S., Konig, C., Ammerman, N., Barrio, M., Borchers, K., Bordon-Pallier, F., Bronstrup, M., Courtemanche, G., Gerlitz, M., Geslin, M., Hammann, P., Heinz, D., Hoffmann, H., Klieber, S., Kohlmann, M., Kurz, M., Lair, C., Matter, H., Nuermberger, E., Tyagi, S., Fraisse, L., Grosset, J., Lagrange, S., Muller, R. (2015) Targeting DnaN for tuberculosis therapy using novel griselimycins. *Science*. **348**, 1106-1112.

Krebs, C., Galonić Fujimori, D., Walsh, C. T., Bollinger, J. M. Jr. (2007) Non-heme Fe(IV)-oxo intermediates. *Acc. Chem. Res*. **40**, 484-92.

Li, C., Wen, A., Shen, B., Lu, J., Huang, Y., Chang, Y. (2011) FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method. *BMC Biotechnol.* **11**, 1-10.

Link, A. J., Tirrell, D. A. (2005) Reassignment of sense codons in vivo. *Methods.* **36**, 291–298.

Liu, H., Naismith, J. H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol.* **8,** 1-10. doi: https://doi.org/10.1186/1472-6750-8-91

Liu, H., Ye, R., Wang, Y. (2015). Highly efficient one-step PCR-based mutagenesis technique for large plasmids using high-fidelity DNA polymerase. *Genet. Mol.* **14**, 3466-3473.

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.

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

Mehl, R. A., Anderson, J. A., Santoro, S. W., Wang, L., Martin, A. B., King, D. S., Horn, D. M., Schultz, P. G. (2003) Generation of a Bacterium with a 21 Amino Acid Genetic Code. *J. Am. Chem. Soc.* **125**, 935-939.

Neumann-Staubitz, P., Neumann, H. (2016) The use of unnatural amino acids to study and engineer protein function. *Curr. Opin. Struct. Bio.* **38**, 119-128.

Novagen (2003) pET system manual. <https://lifewp.bgu.ac.il/wp/zarivach/wp-content/uploads/2017/11/Novagen-pET-system-manual-1.pdf> Accessed on 22nd of February, 2021.

Ohtake, K., Yamaguchi, A., Mukai, T., Kashimura, H., Hirano, N., Haruki, M., Kohashi, S., Yamagishi, K., Murayama, K., Tomabechi, Y., Itagaki, T., Akasaka, R., Kawazoe, M., Takemoto, C., Shirouzu, M., Yokoyama, S., Sakamoto, K. (2015) Protein stabilization utilizing a redefined codon. *Sci. Rep*. **18**, 1-6. doi: https://doi.org/10.1038/srep09762

Ravikumar, Y., Nadarajan, S., Hyeon Yoo, T., Lee, C., Yun, H. (2015) Incorporating unnatural amino acids to engineer biocatalysts for industrial bioprocess applications. *Biotechnol. J.* **10**, 1862-1876.

Peters, C., Buller, R. (2019) Industrial Application of 2-Oxoglutarate-Dependent Oxygenases. *Catalysts*. **9**, 1-20.doi: https://doi.org/10.3390/catal9030221

Pott, M., Hayashi, T., Mori, T., Mittl, P. R. E., Green, A. P., Hilvert, D. (2018) A Noncanonical Proximal Heme Ligand Affords an Efficient Peroxidase in a Globin Fold. *J. Am. Chem. Soc.* **140**, 1535-1543.

Rose, N. R., McDonough, M. A., King, O. N. F., Kawamura, A., Schofield, C. J. (2011) Inhibition of 2-oxoglutarate dependent oxygenases. *Chem. Soc. Rev.* **40**, 4364-4397.

Saleh, A. M., Wilding, K. M., Calve, S., Bundy, B. C., Kinzer-Ursem, T. L. (2019) Non-canonical amino acid labeling in proteomics and biotechnology. *J. Biol. Eng.* **13**, 1-14. doi: https://doi.org/10.1186/s13036-019-0166-3

Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., Panda, A. K. (2015) Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microb. Cell Fact.* **14**, 1-10. doi: https://doi.org/10.1186/s12934-015-0222-8

Skerra, A. (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene.* ***151***, 131–135.

Studier, F. W., Moffatt, B. A (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–130.

Villaverde, A., Mar Carrió, M. (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol. Lett*. **25**, 1385–1395.

Wals, K., Ovaa, H. (2014) Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Front. Chem.* **1**, 2-15.

Wang, A., Nairn, N., Marelli, M., Grabstein, K. (2012) Protein Engineering with Non-Natural Amino Acids. In: Protein Engineering (Kaumaya, P., ed.), InTech, London, 254-290.

Wang, L., Brock, A., Herberich, B., Schultz, P.G. (2001) Expanding the Genetic Code of *Escherichia coli*. *Science*. **292**, 498-500.

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

Xia, Y., Xun, L. (2017) Revised Mechanism and Improved Efficiency of the QuikChange Site-Directed Mutagenesis Method. In: In Vitro Mutagenesis (Reeves, A., ed.), Humana press, Hatfield, 367-375.

Young, D. D., Schultz, P. G. (2018) Playing with the Molecules of Life. *ACS Chem. Biol.* **13**, 854−870.

Young, T. S., Ahmad, I., Yin, J. A., Schultz, P. G. (2010) An Enhanced System for Unnatural Amino Acid Mutagenesis in *E. coli*. *J. Mol. Biol*. **395**, 361-374.

Zhu, Q., Liu, X. (2017) Characterization of non-heme iron aliphatic halogenase Wel05* from Hapalosiphon welwitschii IC52-3: identification of a minimal protein sequence motif that confers enzymatic chlorination specificity in the biosynthesis of welwitindolinones. *Beilstein J. Org. Chem.* **13**, 1168–1173.

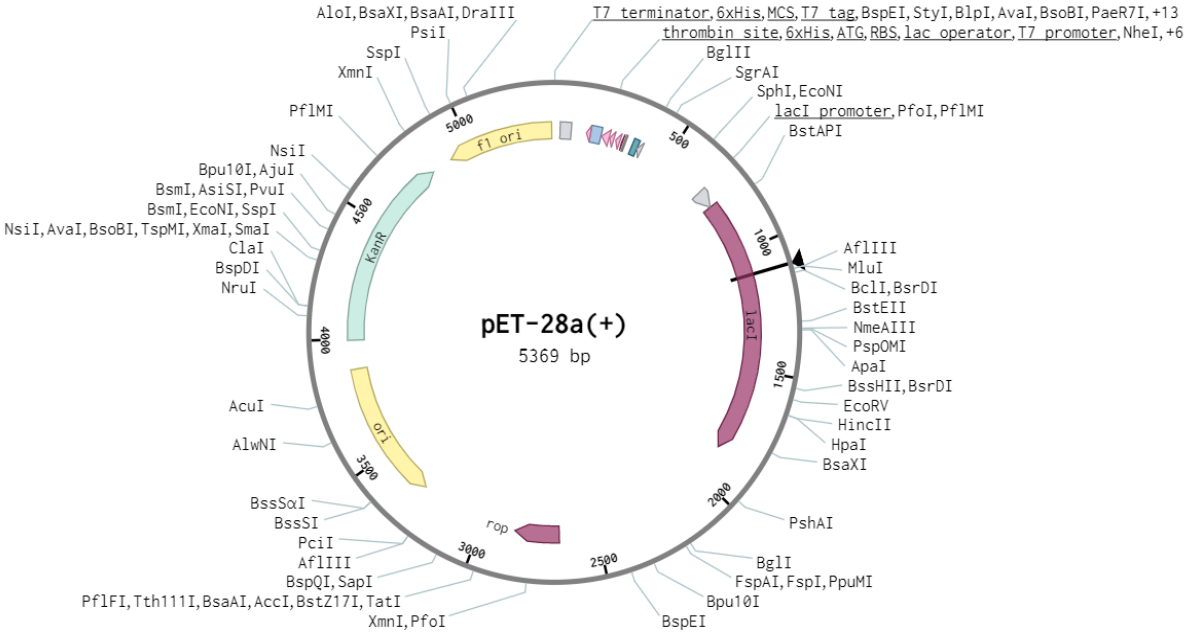
Zou, H., Li, L., Zhang, T., Shi, M., Zhang, N., Huang, J., Xian, M. (2018) Biosynthesis and biotechnological application of non-canonical amino acids: Complex and unclear. *Biotechnol. Adv*. **36**, 1917-1927.

Zwick, C. R., Renata H. (2018) 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.

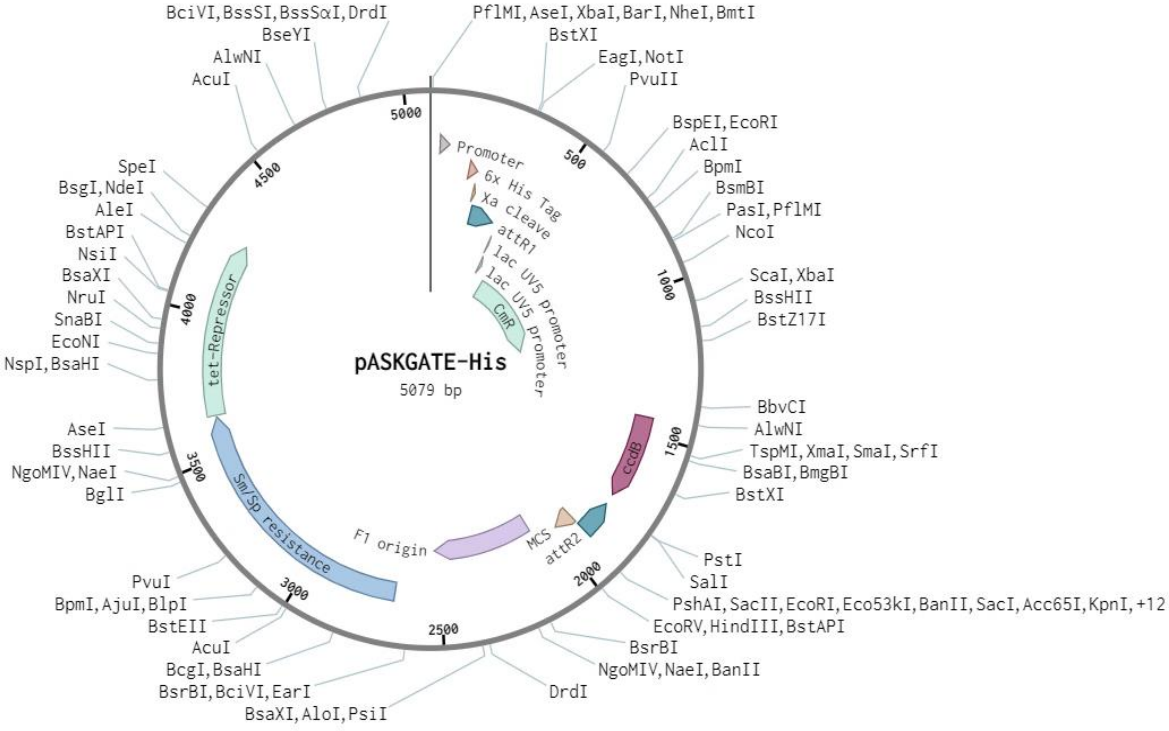
7. APPENDIX

Appendices 1. Abbreviation list

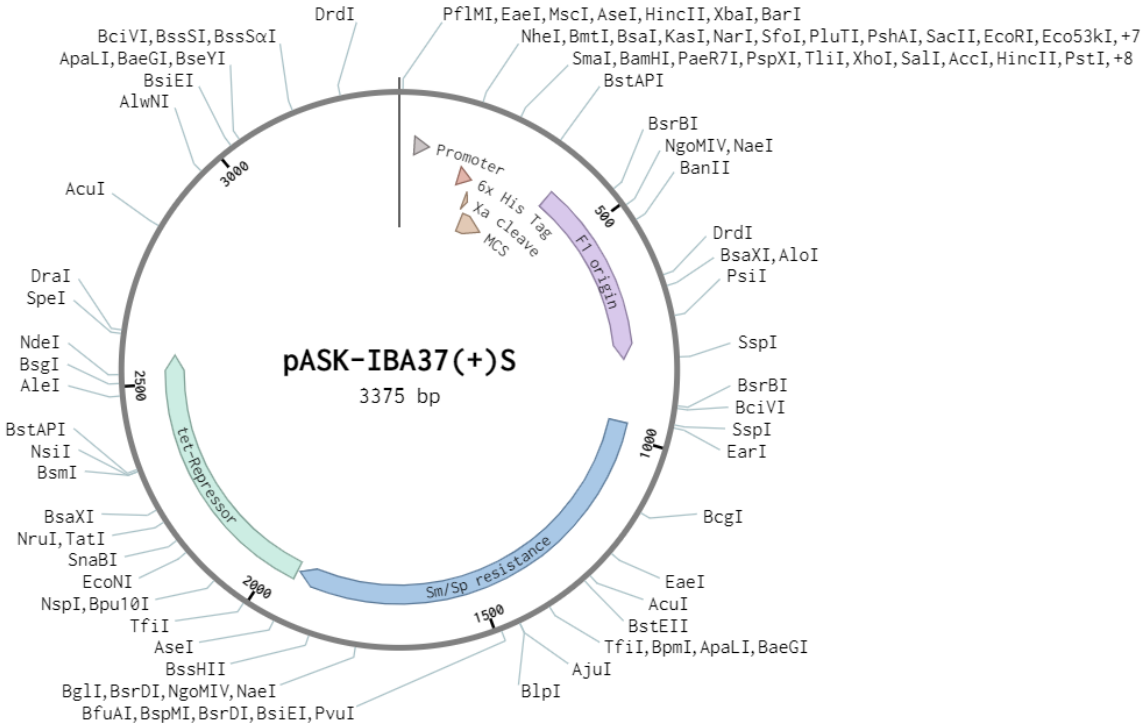
α-ketoglutarate dependent, non-heme Fe(II) dioxygenases	dioxygenases
2-His-1-carboxylic motif	2-histidine-1-carboxylic motif
Asp	Aspartic acid
AvLDO	L-leucine 5-hydroxylase isolated from *Anabaena variabilis*
ATET	anhydrotetracycline
Chl	Chloramphenicol
cPCR	Colony polymerase chain reaction
cv	column volumes
DMSO	Dimethyl sulfoxide
E. coli	*Escherichia coli*
GriE	L-leucine 5-hydroxylase isolated from *Streptomyces muensis*
H	Histidine
His	Histidine
h	hour
Institute	Institute for Molecular Biotechnology, University of Technology Graz
L-Ile	L-isoleucine
IPTG	Isopropyl β- D-1-thiogalactopyranoside
Kan	kanamycin
M	molar
NMH	N-methylhistidine
OD	Optical density
ONC	over-night culture
PCR	Polymerase chain reaction
RT	Room temperature
SDS-PAGE	sodium dodecyl sulphate -polyacrylamide gel electrophoresis
Sm	Spectinomycin
v/v	Volume percentage



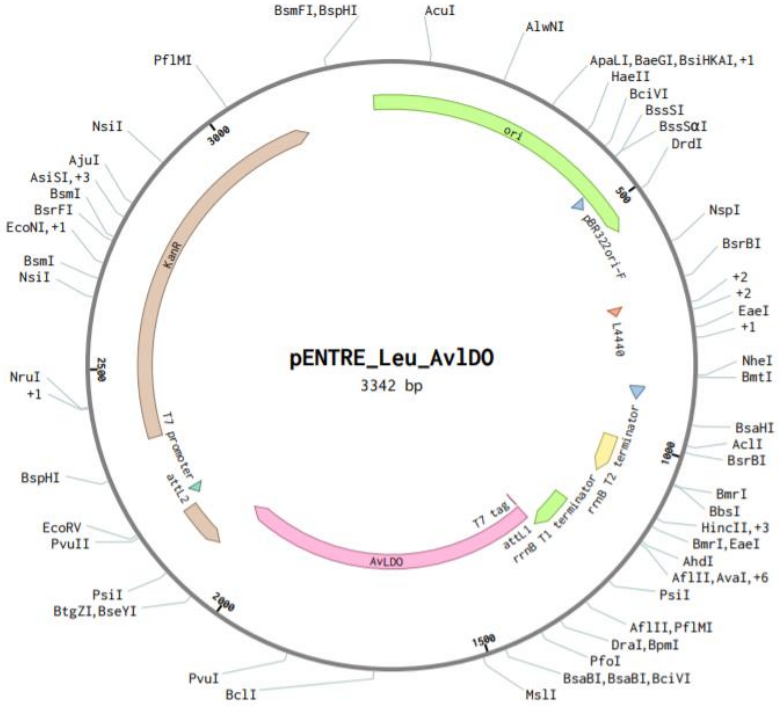
Appendices 2. pET28a(+) plasmid map



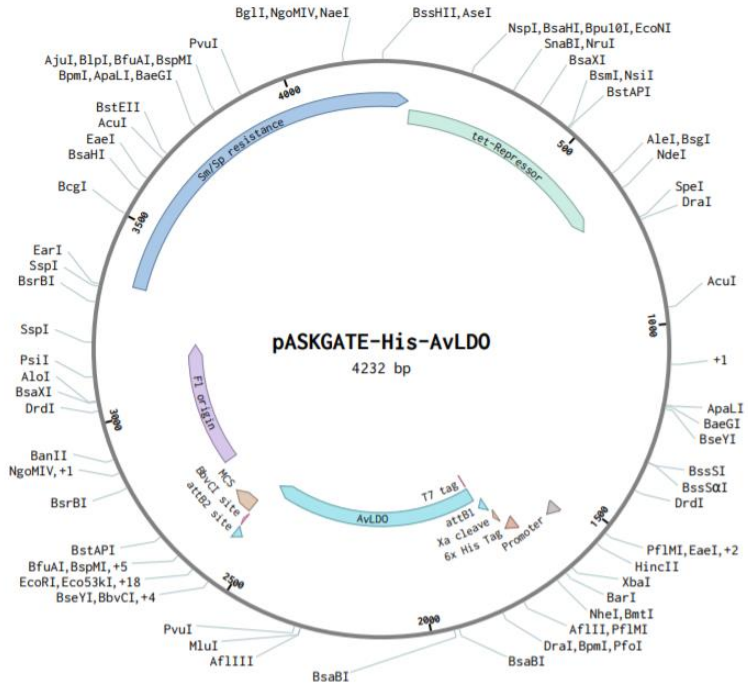
Appendices 3. pASKGATE plasmid map



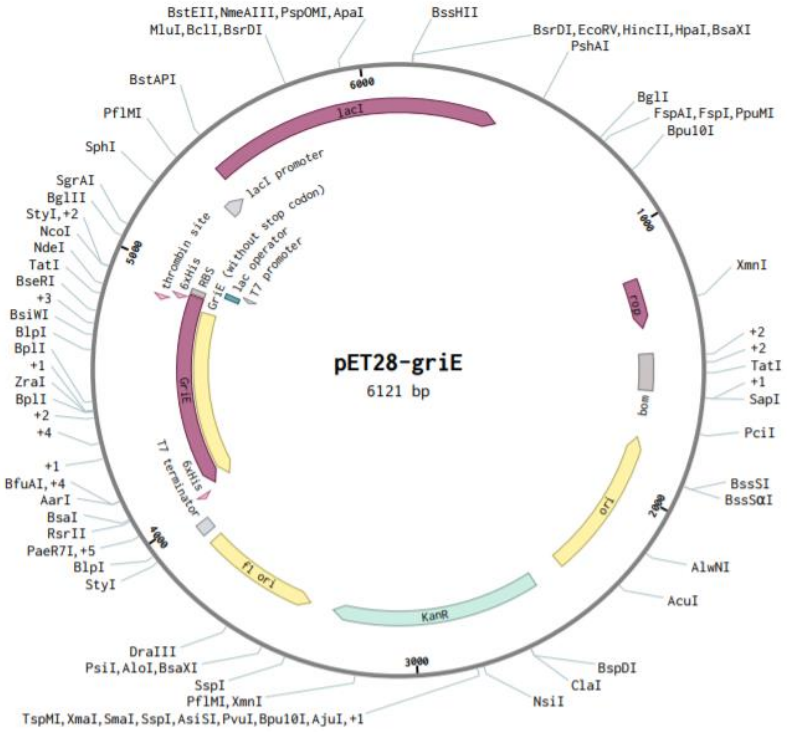
Appendices 4. pASK-IBA37(+)S plasmid map



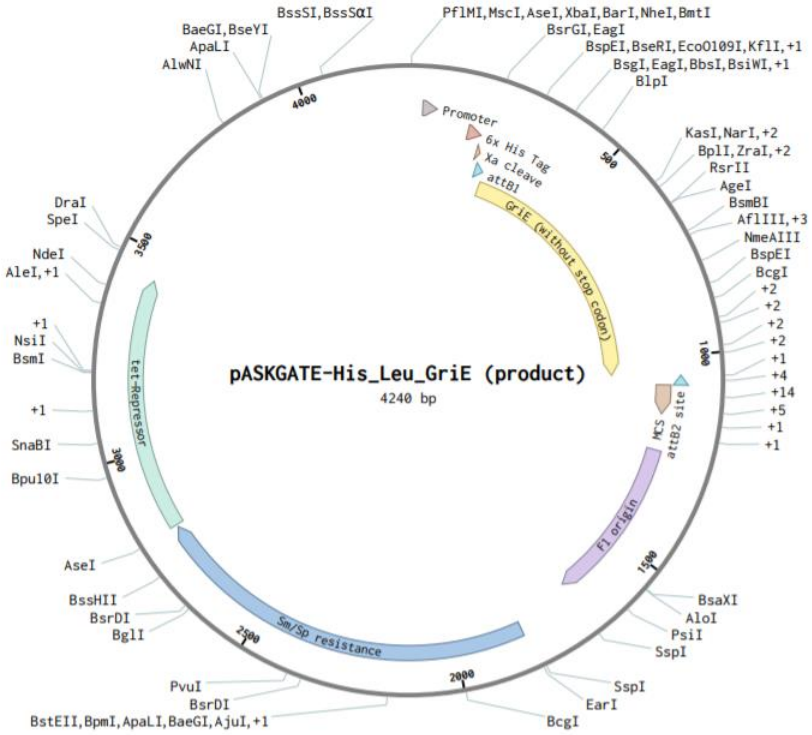
Appendices 5. pENTRE_His_Leu_AvLDO_wt _wt plasmid map



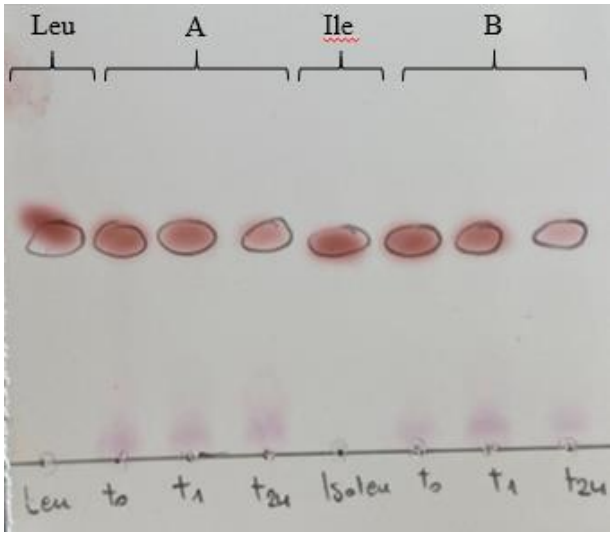
Appendices 6. pASKGATE_His_Leu_AvLDO_wt plasmid map



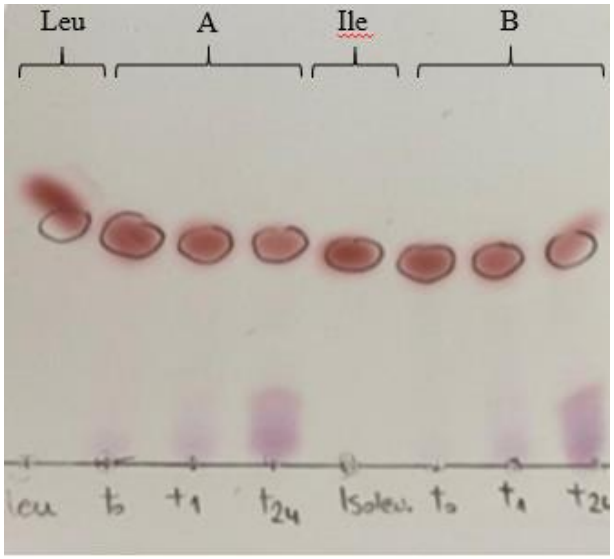
Appendices 7. pET28a(+)_GriE_wt plasmid map



Appendices 8. pASKGATE_His_GriE_wt plasmid map



Appendices 9. TLC of empty plasmid pET28a (+). Time sample analysis (t_0, t_1, t_{24}) of overexpression induced by [A] 0.1 mM and [B] 1 mM IPTG; L-leucine served as a reference and a substrate. Spotting on TLC was detected with ninhydrin staining.



Appendices 10. TLC of empty plasmid pASK37(+). Time sample analysis (t_0, t_1, t_{24}) of overexpression induced by [A] 50 µg L^{-1} and [B] 50 µg L^{-1} ATET; L-leucine served as a reference and a substrate. Spotting on TLC was detected with ninhydrin staining.

STATEMENT OF ORIGINALITY

This is to certify, that the intellectual content of this thesis is the product of my own independent and original work and that all the sources used in preparing this thesis have been duly acknowledged.

_____

Josipa Smoljo