THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Microbial Synthetic Biology, Systems Metabolic Engineering and Enzyme Engineering for Advanced Microbial Biodiesel Production with *Saccharomyces cerevisiae*

JUAN VALLE-RODRIGUEZ

Systems and Synthetic Biology

Department of Life Sciences

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2023

Microbial Synthetic Biology, Systems Metabolic Engineering and Enzyme Engineering for Advanced Microbial Biodiesel Production with *Saccharomyces cerevisiae*
JUAN VALLE-RODRIGUEZ
ISBN 978-91-7905-885-2

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5351
ISSN 0346-718X

Division of Systems and Synthetic Biology
Department of Life Sciences (Biology and Biological Engineering) Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:

Synthesis of Advanced Microbial Biodiesel (FAEEs) utilizing Mutant Yeast by screening Several Organisms, applying Disciplines and Engineering Areas

From left down and clockwise:
Atomic force microscopy image, strain ADP1; from Reinchert et al. 2013.
Microscopy image, strain PD630, grown on gluconate; from Reichelt at Westfälische Wilhelms University (WWU) of Münster, Germany.
Scanning electron microscopy image, grown on Halomonas complex; from the imagebase of the Joint Genome Institute, US Department of Energy.
Transmission electron microscopy image, strain 273-4 grown in NaCl 5%; from Casillo et al. 2020.

Common photograph of mouse strain 000664, by author: Jennifer L. Torrance, The Jackson Laboratory.

Printed by Chalmers Digitaltryck and Reproservice
Gothenburg, Sweden 2023

Microbial Synthetic Biology, Systems Metabolic Engineering and Enzyme Engineering for Advanced Microbial Biodiesel Production with *Saccharomyces cerevisiae*

JUAN VALLE-RODRIGUEZ
Department of Life Sciences
Chalmers University of Technology

Abstract

The continuous requirement of transportation biofuels has brought the necessity to establish alternatives permitting low-cost production of biodiesel while being environmentally friendly. Biodiesel production was achieved utilizing *Saccharomyces cerevisiae* by employing respective enzymes that catalyze the synthesis of fatty acid ethyl esters (FAEEs) based on fatty acyl-CoA molecules and ethanol. Five acyltransferases/wax ester synthases were tested and heterologously introduced in yeast by expressing their codon-optimized gene for expression in a yeast host under the strong promoter TEF1p using plasmid pSP-GM2. In conclusion MhWS2 from oil bacteria *Marinobacter hydrocarbonoclasticus* was the highliest active with 8.1 pmol/(mg protein•min).

Through Metabolically Engineering, metabolism was widely modified for increasing biodiesel production by eliminating fatty acid-consumption competitive pathways, therefore augmenting the fatty acid pool. This was achieved by deleting genes *ARE1, ARE2, LRO1*, *DGA1* and *POX1*, which conferred a 5-fold increase of FAEEs formation (17.2 mg/L). Right after, MhWS2 was overproduced in yeast by chromosomal integration of its codon-optimized *ws2*. Then gene copy number was enhanced by integrating it in δ-regions, conferring 7.5-fold higher biodiesel production in a gradually evolved strain tolerant to 20 mg/mL antibiotic G418.

Furthermore, Protein Engineering of two natural catalysts (MhWS2 and α/β-hydrolase Eeb1p homolog to yeast *Saccharomyces cerevisiae*) was addressed. In these subprojects, directed evolution of these two enzymes was achieved for favoring the synthesis of biodiesel by augmenting their efficiency and altering selectivity towards biocatalyzing FAEEs of desired chain length (C16 and C18, either saturated or monounsaturated). Starting with random mutagenesis of the respective codifying genes (*ws2* and *EEB1*) allowed libraries of random point-mutations. Then library screening was conducted for reducing the CFU (colony formation unit) number; since lipotoxicity was employed as screening method due the condition of the yeast mutants, modifying to a weaker promoter was needed: KEX2p was then further applied. Ultimate selection of the best evolved variants of these enzymes was performed: variants MhWS2-v11 (65.3%

increment when compared to natural MhWS2) and Eeb1p-v04 (45.7% increment). MhWS2-v11 possesses five residue substitutions, while Eeb1p-v04 has 19 residue substitutions. In this case of scientific and technological studies, an advanced biofuel of an upcoming generation has been produced.

Keywords: advanced biofuel, wax ester synthase, hydrolase, fatty acid metabolism, chromosomal delta-integration, directed evolution, protein rational engineering

List of Publications

This doctorate thesis is based on the following publications and manuscripts:

I. "Prospects for microbial biodiesel production."
Shi S, Valle-Rodríguez JO, *et al. Biotechnology Journal* 2011; 6:277-285

II. "Functional expression and characterization of five wax ester synthases in *Saccharomyces cerevisiae* and their utility for biodiesel production."

Shi S, Valle-Rodríguez JO, *et al. Biotechnology for Biofuels* 2012; 5:7-16

III. "Metabolic engineering of Sa*ccharomyces cerevisiae* for production of fatty acid ethyl esters, an advanced biofuel, by eliminating non-essential fatty acid utilization pathways."

Valle-Rodríguez JO, Shi S, *et al. Applied Energy* 2014; 115:226-232

IV. "Engineering of chromosomal wax ester synthase integrated *Saccharomyces cerevisiae* mutants for improved biosynthesis of fatty acid ethyl esters."

Shi S, Valle-Rodríguez JO, *et al*. *Biotechnology and Bioengineering* 2014; 111(9):1740-1747

V. "Directed evolution of a wax ester synthase for production of fatty acid ethyl esters in *Saccharomyces cerevisiae."*

Valle-Rodríguez JO, *et al*. *Applied Microbiology and Biotechnology* 2023; 107:2921-2932

https://doi.org/10.1007/s00253-023-12466-8

VI. "Protein engineering of homologous Eeb1p from *Saccharomyces cerevisiae* for altering its substrate specificity."

Valle-Rodríguez JO, *et al*. in preparation

VII. "Rational engineering of MhWS2 from *Marinobacter hydrocarbonoclasticus* for altering its alcohol substrate specificity towards short-chain alcohols."

Valle-Rodríguez JO, *et al*. in preparation

Contribution summary

I.
I co-planned content of the review.
I co-searched literature for references.
I co-wrote text of the manuscript.
I generated table 1 and co-drew and co-generated figures.
I edited the paper for adjusting it to journal guidelines.
I fixed the paper for compiling after comments and observations from reviewers.

II.
I performed experiments: physiological studies and enzyme activities.
I co-analyzed and co-discussed results.
I co-searched literature for references.
I co-wrote text of the manuscript.
I generated figure 1 and co-drew and co-generated figures.
I co-edited the paper for adjusting it to journal guidelines: text, figures, tables and references.
I co-fixed the paper for compiling after comments and observations from reviewers.

III.
I co-planned the innovations, aims, experimental designs and methodologies of the subproject.
I performed experiments on cell molecular biology: oligonucleotide primer design, DNA cloning and constructing mutants. I performed cultivations, fermentations, physiological studies and enzyme activities.
I co-analyzed and co-discussed results.
I co-searched literature for references.
I co-wrote text of the manuscript.
I co-generated tables, I co-drew and co-generated figures.
I edited the paper for adjusting it to journal guidelines: text, figures, tables and references.
I co-fixed the paper for compiling after comments and observations from reviewers.

IV.

I co-planned experimental designs and methodologies of the subproject.
I co-performed experiments: cultivations, physiological studies, enzyme activities and determinations of gene copy numbers and expression levels.
I co-analyzed and co-discussed results.
I co-searched literature for references.
I co-wrote text of the manuscript.
I co-generated tables.
I co-edited the paper for adjusting it to journal guidelines: text, figures, tables and references.
I co-fixed the paper for compiling after comments and observations from reviewers.

V.

I co-planned the innovations, aims, experimental designs and methodologies of the subproject.
I performed experiments on cell molecular biology: oligonucleotide primer design, DNA cloning, constructing mutants and library generation. I performed screening, selection, analysis and assessment of selected variants, including selectivity; by performing cultivations, fermentations, physiological studies and enzyme activities. Besides modeling of enzyme and variants.
I co-searched literature for references.
I co-wrote text of the manuscript.
I generated tables 2 and 3; I generated figures 1, 3 and 4.
I edited the paper for adjusting it to journal guidelines: text, figures, tables and references.

VI.

I co-planned the innovations, aims, experimental designs and methodologies of the subproject.
I performed experiments on cell molecular biology: oligonucleotide primer design, DNA cloning, constructing mutants and library

generation. I performed screening, selection, analysis and assessment of selected variants, including selectivity; by performing cultivations, fermentations, physiological studies and enzyme activities. Besides modeling of enzyme and variants.
I co-searched literature for references.
I co-wrote text of the manuscript.
I generated all tables, I drew and generated all figures.
I edit the paper for adjusting it to journal guidelines: text, figures, tables and references.

VII.
I planned the innovations, aims, experimental designs and methodologies of the subproject.
I performed experiments on cell molecular biology: oligonucleotide primer design, DNA cloning, constructing mutants and variants. I performed analysis and assessment of created variants, including selectivity; by performing cultivations, fermentations, physiological studies and enzyme activities. Besides modeling of enzyme and variants.
I searched literature for references.
I wrote text of the manuscript.
I generated all tables, I drew and generated all figures.
I edit the paper for adjusting it to journal guidelines: text, figures, tables and references.

Preface

This scientific research thesis handbook is presented as partial fulfillment of the fore ultimate requirement to obtain the degree of Doctor of Philosophy (PhD) in Biosciences at the Department of Life Sciences, Chalmers University of Technology, in Gothenburg, Sweden. The doctorate studies were carried out at the division/group of Systems and Synthetic Biology (SysBio) under the supervision of Prof. Dr. techn. Jens Nielsen (Quantitative Systems Biology Discipline), co-supervision of Docent Res. Prof. Dr. Verena Siewers (Microbial Synthetic Biology Discipline), assistance of Dr. Shuobo Shi; examination of Prof. Dr. Christer Larsson (Microbial Physiology Discipline) and Prof. Dr. Ivan Mijakovic (Bacterial Systems Biology Discipline); final approval of Docent Assoc. Prof. Dr. Elin Esbjörner Winters.

/*Juan Valle-Rodríguez*
August 2023
Gothenburg, Sweden

TABLE OF CONTENTS

Abbreviations

ANOVA	Analysis Of Variance
ARE1	Acyl-CoA:sterol acyltransferase 1 gene
ARE2	Acyl-CoA:sterol acyltransferase 2 gene
ASAT	Acyl-CoA:sterol acyltransferase
BO	Betaβ-Oxidation
CB2I0.2	Yeast strain resistant to 0.2 mg/ml G418
CoA	Coenzyme-A
Cn	n carbon atoms in the molecule
CFU	Colony Forming Unit
DCW	Dry Cell Weight
*DGA*1	Acyl-CoA:diacylglycerol acyltransferase gene
DGAT	Diacylglycerol Acyltransferase
Eeb1p	Ethyl ester biosynthesizing protein 1
E10	Ethanol 10%, blend
E85	Ethanol 85%, blend
Erg10	Acetyl-CoA C-acyltransferase gene
FA	Fatty Acid
FAEE	Fatty Acid Ethyl Ester
G418	Geneticin aminoglycoside antibiotic
HEK	Human Embryonic Kidney
HHV	Higher Heating Value
JV03	Yeast strain, storage-lipid free and unable of betaβ-oxidation
LCAT	Lecithin:cholesterol acyltransferase
LRO1	Lecithin:cholesterol acyltransferase gene
MAG	MonoAcylGlycerol
MhWS2	*Marinobacter hydrocarbonoclasticus* Wax Synthase 2
neo^r	neomycin resistance gene
OA	Oleic Acid
OD	Optical Density
PCR	Polymerase Chain Reaction
POX1	Peroxisomal acyl-coA Oxidase 1 gene
r_{max}	Maximum specific rate
SE	Steryl Ester
SL	Storage Lipid
TAG	Triacylglycerol
TLC	Thin Layer Chromatography
TEF1	Translational Elongation Factor 1-alpha
TT	betaβ-Turn in protein structure
WS	Wax ester Synthase
WT	Wild-Type

Acknowledgements

Foremost, the Chalmers Foundation is acknowledged for hosting as residing address where most of the work was performed and developed by the doctor program conducted at Chalmers University of Technology, Graduate School of Biosciences.

The Knut and Alice Wallenberg Foundation (as Wallenberg Scholar), the European Research Council (project no. 247013), the Swedish Scientific Research Council (Vetenskapsrådet), the Mexican National Council of Science and Technology (CONACYT) are also acknowledged for their economic support to the scientific projects presented herein. Additionally, the 7th European Framework, UNICELLSYS, SYSINBIO, FORMAS Research Council for Sustainable Development (Categories: Environment and Agricultural Sciences; Metabolic Engineering), Futura Foundations, Ångpanneföreningens Forskningsstiftelse (ÅForsk) Research Foundation, NordForsk (sec. Green Growth) and the USA Department of Energy also supported the work presented in this thesis.

Moreover, all staff, crew, personnel, organizers, administrators, logistic staff and related people are also acknowledged due to their direct or indirect contribution to the scientific and technological progress derived from the work presented in this thesis. Specially the Novo Nordisk Foundation Center for Biosustainability, Chalmers House of Chemistry Studies and Chemical Engineering (K-KT hus), and House of Biology Studies and Biological Engineering (B-BT hus), House of Microtechnology and Nanosciences (MC2 hus); Chalmers libraries, specially Chalmers Main Library (Huvudbiblioteket) and Science Communication and Learning Center. Besides the IT-group and technical support of the fore mentioned houses and buildings. Also: Innovation and Patent Group, Mathematical Sciences, Pedagogy Faculty of the Division of Engineering Education Research (Areas: Science Communication and Presentation Techniques), Chalmers Konferens (Academic and Scientific Event Venues) and Development Support.

In publication aspects, all co-authors, journal editors-in-chief, assistants, editorial board members and reviewers are acknowledged for contributing and making possible digitally publishing the papers through the publications platforms: Wiley, BioMed Central, Elsevier, Science Direct and Springer Nature.

Besides, Chalmers Reproservice, Digitaltryck and Printing Group are acknowledged for reproducing and delivering units of this handbook.

The Gothenburg Student Foundation (SGS) is also acknowledged for housing and practicalities needed for the development of these studies. The Chalmers Student Union, Doctor Student Guild and Human Resources Department are also acknowledged for their active support. Aiding Chalmers kårskyrkan, International Student Group and the City Educational Council through the libraries and literature houses.

The deepest thankfulness and gratitude to all institutions, foundations, councils and people through which and whom these scientific and technological projects were developed and this thesis handbook exists!

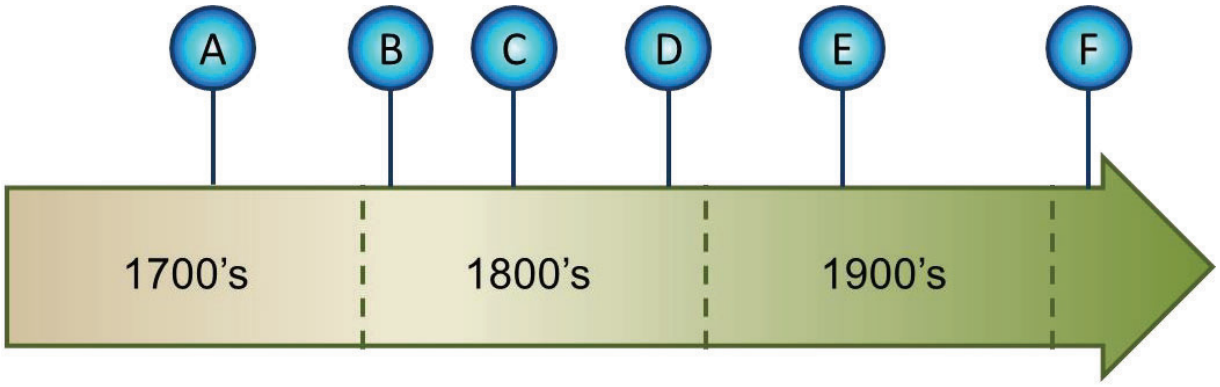
CHAPTER 1

Introduction

1.1 The necessity of novel methods for the synthesis of new generation biofuels

INTRODUCTION

The energy needs of society have increased dramatically since the start of the industrial revolution in England in the eighteenth century (Figure 1). Industry and transportation technology have been evolving rapidly during this period, with a direct link to requirements for incomes and fuels. Ever since it began in 1846, petrochemistry has helped shaping the development of industry and society. However, the impending depletion of the world's oil supplies has created the need to develop new and sustainable alternatives, such as biofuels for transportation and bioplastics. This change has gradually moved us towards a biobased economy. By employing a fermentation process that involves microorganisms, we can now obtain bioethanol and biobutanol as a replacement for gasoline to drive cars and other small vehicles, while biodiesel is replacing diesel in larger vehicles, such as trucks, and alkanes, alkenes, and isoprenoids are being used to fly airplanes.



A 1763	The industrial revolution begins. James Watt invents the water steam engine.
B 1807	Niépce patents the first internal combustion engine.
C 1846	Petroleum industry starts with the first distillation to produce kerosene, by Abraham Gessner.
D 1890	Akroyd Stuart patents the first modern diesel engine.
E 1938	Start of Biofuel production applied to transportation. A passenger bus moved with palm oil ethyl ester in Belgium (Knothe, 2005).
F 2010	First project proposal for a factory of advanced bioethanol produced through yeast fermentation.

Figure 1. Timeline of the technological era linked to the production of fuels, biofuels and advanced biofuels.

Biodiesel is a fuel that consists of mono-alkyl esters obtained from a biological carbon source, mostly vegetable oil or animal fat. It is the second largest biofuel produced globally and its production has been increasing steadily since 2005 reaching a maximum of 18.7 millions of tonnes produced worldwide in 2012 (F.O. Licht, 2013) and it has improved properties compared with petroleum diesel, gasoline and ethanol. Nowadays, biodiesel is mainly produced from soybean oil and methanol. It possesses improved properties compared to the major worldwide fuels that are utilized for terrestrial vehicles (Table 1). For a start, its average chemical formula is larger than those of diesel, gasoline, and ethanol, which means that its molecular weight is higher (almost twice that of diesel and much higher than those of gasoline and ethanol). The density of biodiesel is only 5% greater than that of diesel, but more than 10% greater than those of the other abovementioned fuels. In terms of energy content, biodiesel possesses 13% lower heat of combustion value than diesel and gasoline, but it can perform a more complete combustion, making the engines that employ it more efficient in mass terms; this compensates for the higher energy content of diesel (UNH Biodiesel Group, 2010). Because of its origin and production method, biodiesel represents no sulfur dioxide or monoxide emissions, which reduces its environmental impact (MacDonald *et al.*, 2011).

Table 1. Main properties of ethyl palmitate and common transportation fuels currently utilized worldwide.

Property	Biodiesel*	Diesel†	Gasoline	Ethanol
Chemical formula	$C_{19}H_{35}O_2$	$C_{12}H_{26}$	C_8H_{18}	C_2H_6O
Molecular weight, g/mol	309	170	114	46
Density (20°C), kg/L	0.88	0.83	0.74	0.79
HHV, MJ/kg	37.3[a]	43.1	43.2	29.7

*Common biodiesel produced from soybean oil and methanol
†Petroleum based diesel (petrodiesel)
Excepting ethanol, properties of all fuels are a representative average

Currently, the most common method for biodiesel production is transesterification, by which the oil reacts with a short chain alcohol, yielding esters and glycerol (Figure 2). However, this process has several disadvantages, including large energy intensiveness, having to compete with food supply, and carrying toxic waste water.

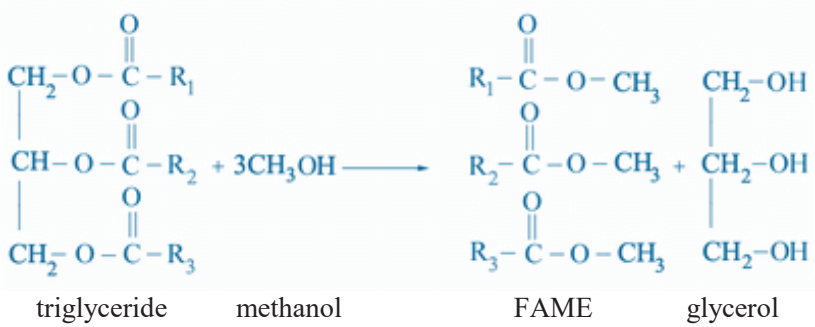


Figure 2. Transesterification reaction to achieve biodiesel production through a chemical or enzymatic reaction of triglyceride present in plant oils and methanol to form FAMEs and glycerol as a by-product.

Novel methods for the production of biodiesel are chemical synthesis or enzyme catalysis of lipids obtained from microorganisms such as microalgae, bacteria, filamentous fungi, and yeasts. In the last four years, other alternatives have started to be investigated through synthetic biology and metabolic engineering of microorganisms to carry on *in vivo* production of the biodiesel esters. Examples include *E. coli* (Steen *et al.*, 2010) and *S. cerevisiae* developed first by (Kalscheuer *et al.*, 2004) and also in the present PhD project.

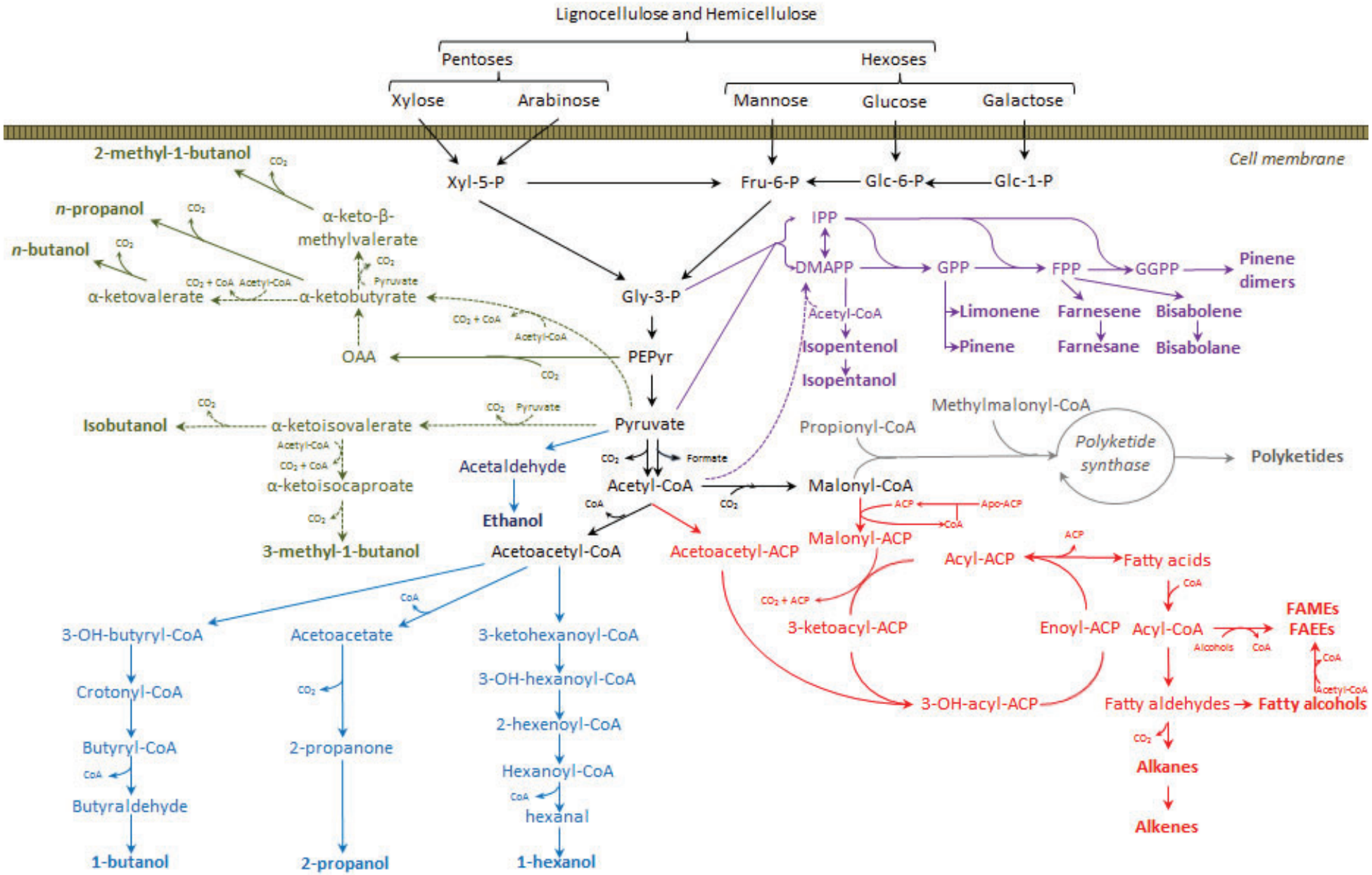
1.2 Advanced biofuels

1.2.1 Alcohols

Some alcohols can be mixed with gasoline where they can function as oxygenates, and some of them can fully replace gasoline (Lee *et al.*, 2008). Currently, flexible-fuel vehicles in North America and Europe are using **ethanol** blends of up to 85% (E85) and up to 100% (E100) in Brazil (Renewable Fuels Association, 2011). It is naturally produced by *S. cerevisiae.* **Butanol**, on the other hand, has a limited miscibility with water, but is fully miscible with gasoline (Lee *et al.*, 2008). These properties, together with the generation of a greater heat of combustion (HHV), make butanol a better fuel candidate than ethanol (Jin *et al.*, 2011). **n-Butanol** has been produced in *Clostridium* bacteria (Qureshi *et al.*, 1999), and its biosynthesis pathway has been introduced in *S. cerevisiae* (Krivoruchko *et al.*, 2013; Steen *et al.*, 2008). On 2013, titers of 16.3 mg/L of production in yeast were internationally reported; in this project n-butanol was synthesized by the conversion of acetyl-CoA into acetoacetyl-CoA through the catalysis of thiolase Erg10, followed by the reactions propitiated by heterologous enzymes: 3-hydroxybutyryl-CoA dehydrogenase (Hbd), crotonase

(Crt), *trans*-enoyl-CoA reductase (Ter) and bifunctional butyraldehyde/butanol dehydrogenase (AdhE2) (Krivoruchko *et al.*, 2013). It can also be produced through the catabolism of glycine, being its direct precursor α-ketovalerate (Branduardi *et al.*, 2013), or through the catabolism of L-threonine, being its direct precursor butyl aldehyde, where the deletion of the *adh1* gene conferred yeast to produce more than 120 mg/L; later by overexpression of pathway enzymes and elimination of competing pathways, yeast was able to produce 242.8 mg/L (Si *et al.*, 2014). **Isobutanol** can also be applied as a fuel. It possesses a very similar heat of combustion to n-butanol, but with a higher octane number due to the carbon atom branches in its molecular spatial distribution, making it more applicable for industrial purposes. Isobutanol can be produced in *S. cerevisiae* from pyruvate through α-isoketovalerate (Chen *et al.*, 2011; Branduardi *et al.*, 2013; Kondo *et al.*, 2012); its production was 5-fold increased through the valine biosynthesis metabolic pathway in anaerobic fermentation, in which the overexpression of the genes *ILV2*, *ILV3* and *ILV5* allowed this increase. Isobutanol yield was then doubled by the additional overexpression of the gene *BAT2*, encoding the cytoplasmic branched-chain amino acid aminotransferase. Under aerobic conditions, its yield was 3.86 and 0.28 mg per g glucose, for the genetically engineered and the wild-type strain, respectively. Isobutanol can also be biosynthesized using several yeasts, through the Ehrlich pathway (Yan *et al.*, 2009). This biofuel is a good candidate for being produced in baker's yeast, due to the high tolerance that this yeast exhibits towards it (Knoshaug *et al.*, 2009).

Figure 3. ***(on next page)*** Metabolic pathways (natural and synthetic) for production of diverse advanced biofuels. Carbon source utilization biochemical pathways and glycolysis are depicted in black. Keto-acid mediated metabolism through the Ehrlich pathway is shown in green. Biosyntheses of short chain alcohols are presented in blue. Isoprenoid synthesis through the mevalonate and methylerythritol pathway is reported in purple. The cycle for polyketide biosynthesis is shown in gray. Fatty acid metabolism for fatty acid methyl/ethyl esters (FAMEs/FAEEs), fatty alcohols and log chain alkanes and alkenes are presented in red. The dashed lines represent multiple biochemical reactions. Biofuel products are in bold font. Abbreviations: ACP, acyl carrier protein; CoA, Coenzyme A; CO_2, carbon dioxide; Xyl-5-P, xylose-5-phosphate; Fru-6-P, fructose-6-phosphate; Glc-6-P, glucose-6-phosphate; Glc-1-P, glucose-1-phosphate; Gly-3-P, glycerol-3-phosphate; PEPyr, phosphoenolpyruvate; OAA, oxaloacetate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl diphosphate; GGPP, geranyl-geranyl pyrophosphate; FAMEs, fatty acid methyl esters; FAEEs, fatty acid ethyl esters


Lignocellulose and Hemicellulose
Pentoses
Hexoses
Xylose
Arabinose
Mannose
Glucose
Galactose
Cell membrane
2-methyl-1-butanol
n-propanol
n-butanol
α-keto-β-methylvalerate
α-ketovalerate
α-ketobutyrate
Xyl-5-P
Fru-6-P
Glc-6-P
Glc-1-P
Gly-3-P
PEPyr
OAA
Pyruvate
IPP
DMAPP
GPP
FPP
GGPP
Pinene dimers
Limonene
Pinene
Farnesene
Farnesane
Bisabolene
Bisabolane
Acetyl-CoA
Isopentenol
Isopentanol
Isobutanol
α-ketoisovalerate
α-ketoisocaproate
3-methyl-1-butanol
Acetaldehyde
Ethanol
Formate
Methylmalonyl-CoA
Propionyl-CoA
Polyketide synthase
Polyketides
Malonyl-CoA
Malonyl-ACP
Acetoacetyl-CoA
Acetoacetyl-ACP
Acyl-ACP
Fatty acids
3-ketoacyl-ACP
Enoyl-ACP
3-OH-acyl-ACP
Acyl-CoA
FAMEs
FAEEs
Fatty aldehydes
Fatty alcohols
Alkanes
Alkenes
3-OH-butyryl-CoA
Crotonyl-CoA
Butyryl-CoA
Butyraldehyde
1-butanol
Acetoacetate
2-propanone
2-propanol
3-ketohexanoyl-CoA
3-OH-hexanoyl-CoA
2-hexenoyl-CoA
Hexanoyl-CoA
hexanal
1-hexanol

1.2.2 Isoprenoids

Due to the methyl branches and cyclic structures of their molecules, isoprenoids (or terpenoids) can serve as high octane number fuels (Lee *et al.*, 2008; Peralta-Yahya *et al.*, 2011), and are also employed as precursors of aroma compounds and pharmaceuticals (Kirby *et al.*, 2009). They are produced from two or three units of isoprene, derived from isopentenyl diphosphate. This compound is a product of both the mevalonate pathway and the 2-methylerythritol 4-phosphate pathway. Isoprenoids are naturally produced in plants and algae, but in low quantities, and inefficiently (Banerjee *et al.*, 2002; Tsukahara *et al.*, 1999). In order to improve total yields and reduce production times, they have been heterologously produced in both bacteria and yeasts (Martin *et al.*, 2003; Westfall *et al.*, 2012). **Farnesane** can be produced from the hydrogenation of farnesene, and is currently being produced for the transportation market by the company Amyris, using *S. cerevisiae*. Besides this **farnesene** has also been successfully biosynthesized in this yeast through a conversion of FPP through the mevalonate (MVA) pathway or the 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway rendering IPP and DMAPP, respectively (Chen *et al.*, 2011); farnesene has also been homologously overproduced from glucose at more than half the theoretical yield (Chandran *et al.*, 2011) and more recently Amyris has developed the production of β-farnesene, where its formation has been optimized via the central carbon metabolism. The heterologous expression of farnesene synthase (FS), X5P-using phosphoketolase (xPK), phosphotransacetylase (PTA) and aldehyde dehydrogenase acylase (ADA) permitted a lower ATP need for acetyl-CoA production, shortening carbon waste therefore improving cofactor consumption. These mutant yeast cells allowed 25% increased farnesene production when compared with control cells grown in same sugar intake, while they consumed 75% less dissolved oxygen. They maintained a stable yield for two weeks in cell culture that reached more than 15% (v/v) farnesene (Meadows *et al.*, 2016). **Bisabolane** is formed through hydrogenation of **bisabolene**. The biosyntheses of α-alpha-bisabolene, β-beta-bisabolene and γ-gamma-bisabolene have recently been introduced in *Yarrowia lipolytica*, through the heterologous expression of α-alpha-bisabolene synthase of *Abies grandis*, β-beta-bisabolene synthase of *Zingiber officinale* and γ-gamma-bisabolene synthase of *Helianthus annuus* respectively. Later, production was improved by overexpressing the endogenous mevalonate pathway genes and introducing two multidrug efflux transporters: AcrB, a resistance-nodualtion-cell division (RND) family efflux pump of *Escherichia coli* and ABC-G1 of *Grosmania clavigera*, a member of the ATP-binding cassette (ABC) transporter superfamily. Moreover, cultivation conditions on glucose as

substrate were optimized for increasing production where titers were 100 mg/L of α-alpha-bisabolene, 5.8 mg/L of β-beta-bisabolene and 3.6 mg/L of γ-gamma-bisabolene; besides confirming the potential of this oleaginous yeast to produce these bisabolene molecules from 67 mg/L of β-beta-bisabolene and 21 mg/L of γ-gamma-bisabolene (Zhao, 2021). Bisabolane and bisabolene have cetane numbers 58 and 52 respectively. These are similar to those of diesel, which are in the range 40 to 60. **Limonene** is produced by converting geranyl pyrophosphate (GPP), its formation has been achieved through the catalytic activity of (−)-limonene synthase from *Perilla frutescens* and a (+)-limonene synthase from *Citrus limon*; in both cases by utilizing dodecane as extractive solvent allowed a recovery of 0.028 mg/L (−)-limonene and 0.060 mg/L (+)-limonene (Jongedijk *et al.*, 2014). However, limonene is highly toxic to *S. cerevisiae*, its minimum inhibitory concentration (MIC) was determined as 0.44 mM (Brennan *et al.*, 2012). A study showed that the yeast cell wall gets affected by this isoprenoid (Brennan *et al*., 2013), therefore an evolutionary engineering approach conferred *S. cerevisiae* a higher tolerance towards it (Brennan *et al*., 2015). Its formation in *E. coli* has been confirmed through the methylerythritol 4-phospahte (MEP) pathway, while it has also been proven through the mevalonate pathway (Alonso-Gutierrez *et al*., 2013; Jongedijk *et al*., 2016). Besides being present in lemon and citric fruits, limonene is present in a large number of plants (Jongedijk *et al*., 2016) and some bacterium species (Effmert *et al.*, 2012; Hung *et al.*, 2013; Heddergott *et al.*, 2014; Guneser *et al.*, 2017). *d*-Limonene has been synthesized in yeast by the enzymatic catalysis of the protein encoded by t*LS* of *Citrus limon* (Hu *et al.*, 2020).

Limonene can act as a direct precursor of **pinene**. α-Alpha-Pinene biosynthesis in *E. coli* has been achieved (Yang *et al.*, 2013; Sarria *et al.*, 2014). By directed evolution of the pinene-synthase Pt1 from *Pinus taeda* hosted in cyanobacteria, the variant Q457L was isolated and its mutated gene expressed in *E. coli* and cyanobacteria. Then, its coexpression with pathway involved enzymes and GGPS from *Abies grandis* conferred a higher GPP consumption yielding 140 mg/L of pinene (Tashiro *et al.*, 2016). Later, through employing tolerance, directed evolution and modular co-culture engineering with two *E. coli* strains permitted a higher 166.5 mg/L production (Niu *et al.*, 2018). Furthermore, genomic and transcriptional studies combining with CRISPR activation and CRISPR interference revealed that its tolerance and overproduction may be related to a) the mutations of the DXP pathway genes, the *rpoA* and some membrane protein genes, by their up regulated transcription; and b) the mutations of *ydiJ*, *yjbQ*, *prpR* and *cedA* by their down regulated transcription

(Niu *et al.*, 2019). A synthetic feedback loop has been designed for optimizing its tolerance in *E. coli* (Siu *et al.*, 2018).
More recently, its production has been introduced in *Yarrowia lipolytica* (Wei *et al.*, 2021), in purple non-sulfur photosynthetic bacterium *Rhodobacter sphaeroides* (Wu *et al.*, 2021) and in the hyperextremophile bacterium *Deinococcus radiodurans* (Helalat *et al.*, 2021).
Comparing all fuels, **pinene dimers** possess the highest energy content in their molecules due to their embedded ring structures (Harvey *et al.*, 2009). Their chemical synthesis and characterization has been recently patented (Harvey *et al.*, 2012).

1.2.3 Polyketides

Until recently, polyketides have been used as antibiotics and insecticides (Yuzawa *et al.,* 2012). Their properties as fuels have not yet been tested, but reduced polyketides are a good source of hydrocarbons that can produce biofuels. Their special property is that they can achieve a higher degree of polymerization by way of a polyketide synthase catalysis of propionyl-CoA and methyl malonyl-CoA (Kirby *et al.*, 2009). This biochemical pathway is highly flexible for producing hydrocarbons with variated structures (Yuzawa *et al.,* 2016). Designing and controlling the structure are results as the choice of domains and modules in the modular polyketide synthases (Cai and Zhang, 2018); therefore the production of short-chain hydrocarbons can be predicted by a rational combination of these domains and modules (Yuzawa *et al.,* 2016). For instance, polyketide synthases can be engineered to form 1-butene, 1-hexene, 1-pentanol and 1-hexanol. These chemical compounds can later be oligomerized to produce diesel and jet fuels (Liu *et al.*, 2015). Ramifications in the polyketide molecule can be obtained either by incorporating the methyl malonyl-CoA as substrate or by methylating with *S*-adenosyl methionine (Poust *et al.*, 2015). The small ketone butanone has been synthesized in *E. coli* cells engineered to express a promiscuous β-ketothiolase from *Cupriavidus necator*, a CoA transferase and an acetoacetate decarboxylase from *Clostridium acetobutylicum* (Srirangan *et al.*, 2016). Lately, the bacterium *Streptomyces albus* has been rationally engineered to produce short-chain ketones (methyl- and ethylketones: 3-methyl-2-butanone, 2-methyl-3-pentanone and 3-methyl-2-pentanone) at a large production (> 1 g/L) based on plant biomass hydrolysates (Yuzawa *et al.,* 2018). Despite that polyketide synthases can be engineered to make ultimately desired biofuels with almost perfect properties, the internationally reported values for polyketide-derived biofuels are considerably lower than those for the fatty acid- or isoprenoid-derived fuels, so nowadays it

is a big challenge to produce economically feasible polyketide-derived biofuels (Zargar *et al.*, 2017). Scientific research is being developed for establishing and validating metabolic pathways that could give a higher flexibility for producing tightly reduced polyketide molecules intended to be used as biofuels. For example, the effect of TtgABC efflux pump transporter from *Pseudomonas putida* on the production of short-chain alcohols (Basler *et al.*, 2018).

1.2.4 Fatty-acid derivatives

Fatty acids can not be used as fuels, but their hydrophobic acyl chains can be used as precursors of fatty acid esters, fatty alcohols, alkanes, and alkenes. **Fatty alcohols** are non-branched-chain primary alcohols derived from a biological source. They range from butyryl alcohol to hexadocanol. The ones most commonly produced nowadays are lauryl, stearyl and oleyl alcohols (Noweck *et al.*, 2000). They have been produced in microorganisms through the expression of an acyl-coA synthase for activating fatty acids, followed by the catalysis of an acyl-CoA reductase, which reduces the activated fatty acid to a fatty aldehyde and the subsequent hydrogenation rendering a fatty alcohol (Steen *et al.*, 2010). Long-chain fatty alcohols (LCFOHs, C14 to C20) and Medium-chain fatty alcohols (MCFOHs, C6 to C12) are proven to be biologically synthesized. *E. coli* has been successfully engineered to produce 1-dodecanol and 1-tetradecanol through the balanced expression of acyl-ACP thioesterase (BTE), an acyl-CoA ligase (FadD), and an acyl-CoA/aldehyde reductase (MAACR). By employing a fed-batch cultivation and adding dodecane, a production of 1.6 g/L of these fatty alcohols was achieved at a yield higher than 0.13 g/g, glucose was used as carbon source (Youngquist *et al.*, 2013). More recently, a plasmid-free strain was engineered in this bacterium, by integrating in its chromosome the full pathway for producing the MCFOH: 1-octanol, where a C8 specific thioesterase was employed, the ultimate titer obtained was 1.3 g/L and a greater than 90% specificity from glycerol (Hernandez Lozada *et al.* 2020). In *S. cerevisiae*, MCFOHs production of 252 mg/L was obtained by the strain YH28 *tpo1*Δ RF1+303 hosting an engineered carboxylic acid reductase from *Mycobacterium marinum* (MmCAR) through protein engineering, where the selected variant M150 produced 2.8-fold more MCFOHs (Hu *et al.*, 2020). Moreover, by expressing the human transporter FATP1 fatty alcohol export was approx. 5-fold increased while the acyl-CoA synthase catalytic activity of this transporter augmented cell growth (Hu *et al.*, 2018). In another research study, an ultimate production of 1.2 g/L was obtained through increasing FAR expression; deleting competing reactions encoded by *DGA1*, *HFD1*, and *ADH6*; overexpressing a mutant acetyl-CoA carboxylase; limiting NADPH and carbon

usage by the glutamate dehydrogenase encoded by *GDH1*; and overexpressing the Δ9-desaturase encoded by *OLE1.* Besides production based on lignocellulosic feedstocks allowed 0.7 g/L, while 6.0 g/L were produced in fed-batch cultivation (d'Espaux *et al.*, 2017).

Since they are constituents of diesel, **alkanes** obviously comprise a fuel. Species of cyanobacteria can naturally synthesize long-chain alkanes in low titers (Tan *et al*., 2011). These include: *Shewanella oneidensis,* which produces them by the condensation of fatty acids (Sukovich *et al.*, 2010); then alkanes were heterologously obtained in *E. coli*, through the expression of an acyl-acyl carrier protein reductase and an aldehyde decarbonylase from cyanobacteria for rendering fatty acids through fatty aldehydes and finally produce C13 to C17 alkanes (Schirmer *et al.*, 2010). More recently, the biosynthesis of long-chain alkanes has been heterologously achieved in *S. cerevisiae* through the deformylation of fatty aldehydes by the enzyme fatty aldehyde deformylating oxygenase (FADO) from *Synechococcus elongates* (Buijs *et al.*, 2014). Moreover, an algal fatty acid photodecarboxylase (FAP) driven by light was recently found to be responsible for the conversion of fatty acids to their corresponding alkanes, so this enzyme was introduced in *E. coli* to produce hydrocarbons in the presence of light at the visible range (Sorigué *et al.*, 2017). The biosynthesis of alkanes can also occur using fatty aldehydes as substrate by the catalysis of aldehyde deformylating oxygenase (ADO) or aldehyde decarbonylase (AD). ADO is classified as a non-heme di-iron oxygenase requiring molecular oxygen (O_2) and an external reducing system to provide four electrons, resulting in hydrogen peroxide (H_2O_2) and formate (HCOOH) as by-products of the catalytic reaction (Herman, 2016). AD decarbonylates fatty aldehydes to form alkanes. The most studied ADs are CYP4G1, a P450 oxidative decarbonylase from *Drosophila melanogaster* and ECERFERUM1 (CER1) from *Arabidopsis* with the associated protein ECERFERUM3 (CER3); they can naturally catalyze the synthesis of very-long-chain alkanes in insects (Bernard *et al.*, 2012) and plants (Qiu *et al.*, 2012), respectively. They have been expressed in *S. cerevisiae* to generate alkanes with one less carbon atom per molecule from their respective fatty aldehyde (Kang *et al.*, 2017). Alkane biosynthesis in yeast cells has largely progressed in the last years. Despite this, the low efficiency of involved enzymes aimed by the strong competition of fatty alcohols as products are considered to be the main impediments of alkane formation improvement utilizing this yeast (Zhou *et al.*, 2016; Buijs *et al.*, 2015). Because of this, compartmentalization in yeast peroxisome has been validated as a good solution to the problem, by isolating the enzymes involved in alkane formation; so that cytosolic competing pathways remain excluded. As

organelles, peroxisomes are suitable for alkane biosynthesis due the absence of ALRs and ADHs, aimed to the broader NADPH availability from the peroxisomal NADP-dependent isocitrate dehydrogenase Idp3 (Rottensteiner *et al.*, 2006). Lately, an ADO from cyanobacterium *Synechococcus elongatus* (*Se*ADO) together with a CAR from *Mycobaterium marium* (*Mm*CAR) were targeted to the peroxisomes in yeast providing approximately 0.12 mg/L alkanes, representing 90% more alkanes than the production through the cytosolic pathway (Zhou *et al.,* 2016). This strategy was followed in a related project, there medium-chain alkane (C7-C13) production in yeast was achieved (Zhu *et al.,* 2017). In a further study where an increment in the precursor supply and knocking out the genes *ADH5* and *SFA1*, encoding cytosolic ALR/ADH, the alkane titer was 1.2 mg/L with a predominantly smaller fatty alcohol accumulation (Zhu *et al.,* 2017).

All the fore mentioned strategies can be also employed for synthesizing a respective **alkene** molecule, where two unsaturations associated to a double bond and two less hydrogen atoms in the molecule referring to the regarding alkane, are particular of these fatty-acid derivative.

Moreover, alkenes can be formed directly from free fatty acids as precursor through several biochemical reactions: 1) one of them is the catalytic reaction propitiated by the enzyme OleTJE P450 fatty acid decarboxylase from *Jeotgalicoccus* sp. ATCC 8456 with the need of hydrogen peroxide and producing carbon dioxide as by-product (Liu *et al.,* 2014); 2) another reaction is by the non-heme iron oxidase UndA from *Pseudomonas fluorescens* Pf-5 (Rui *et al.,* 2014); 3) another alternative is by the catalysis of UndB from *Pseudomonas mendocina* ymp (Rui *et al.,* 2015); 4) alkenes can also be synthesized by the catalytic reaction of UndB from *Pseudomonas fluorescens* Pf-5*;* 5) a 1-alkene molecule can be considered as an olefin molecule where the radical group is an aliphatic hydrocarbon chain, therefore 1-alkenes can be produced directly from free fatty acids by a PKS also consuming malonyl-CoA rendering carbon dioxide as by-product, too (Zhu *et al., 2017*); 6) alkenes can also be produced with fatty acyl-CoA as precursor through the reaction catalyzed by the thiolase-based system OleABCD from *Micrococcus luteus* having as by-products CoA-SH and carbon dioxide (Beller *et al.,* 2010); 7) very-long-chain alkenes can be synthesized by the catalysis of the modular polyketide synthase Ols from *Synechococcus* sp. PCC 7002 utilizing acyl-ACP as substrate (Mendez-Perez *et al.,* 2011).

Alkenes can also be naturally produced in microorganisms such as the marine bacteria *Shewanella* spp that possess the ability to synthesize **1-alkenes**; moreover cyanobacteria such as *Jeotgalicoccus* sp., and *Nostoc punctiforme*, can also synthesize them directly (Schirmer *et al*., 2010). If the starting acyl molecules in the biochemical reactions are unsaturated, then microbes would eventually produce alkenes.

1.3 Advanced microbial biodiesel production

Novel methods for the production of biodiesel are being applied by utilizing alternative sources of triacylglycerols derived from oils extracted from microbes. This method can be used to manufacture an indirect advanced biodiesel product. One of the substrates, methanol, is the most common alcohol employed, and is obtained from natural gas (Figure 3). A big disadvantage of this scenario is that natural gas is a non-renewable hydrocarbon, and is also toxic and hazardous.

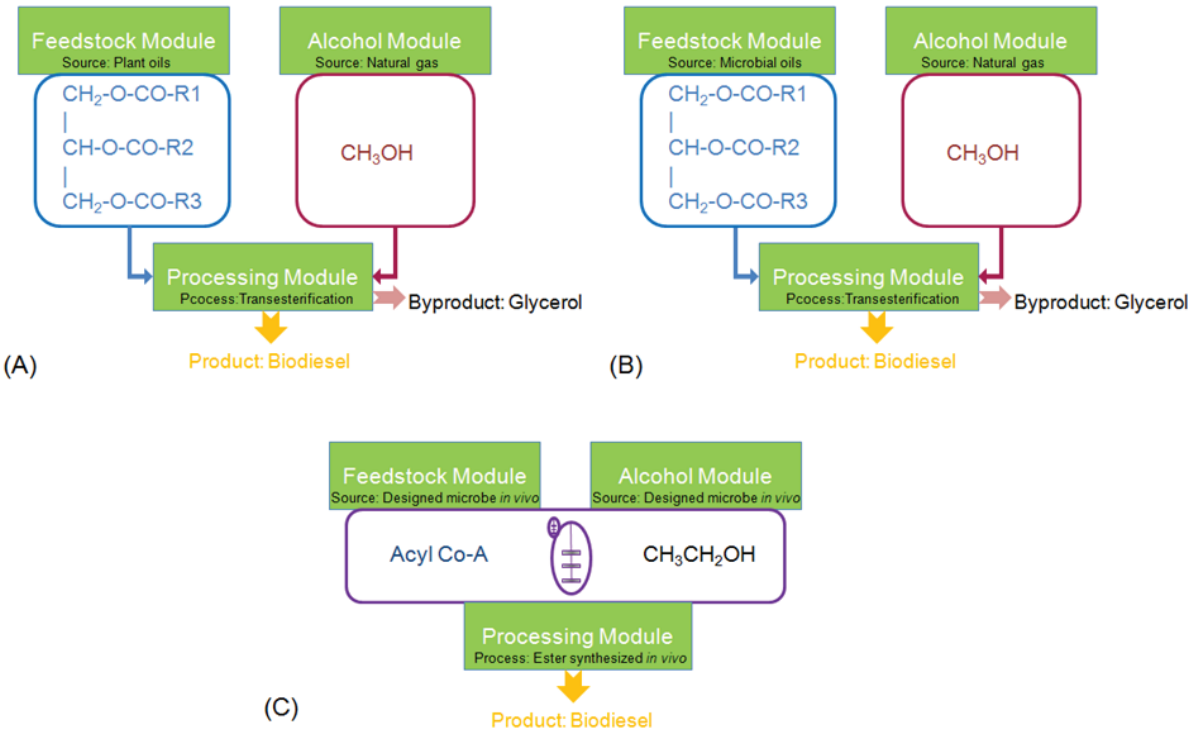


Figure 3. Production of biodiesel through A) chemical or enzymatic transesterification reaction utilizing vegetable oils, B) chemical or enzymatic transesterification reaction utilizing microbial oils, and C) direct *in vivo* synthesis using genetically engineered cell factories (taken from Shi *et al*., 2011).

Microorganisms that can be utilized as a source of oils are microalgae, bacteria and fungi, including filamentous fungi and yeasts. Of these, **microalgae** are currently being widely used. Species of microalgae that can achieve the highest content of lipids (between 25 and 77% by dry weight) are *Schizochytrium sp*. (Khozin-Goldberg *et al*., 2002), *Botryococcus braunii* (Illman *et al*., 2000) and *Nannochloropsis sp*. (Khozin-Goldberg *et al*., 2002). Microalgae produce complex oils, mainly composed of mono- and poly-unsaturated fatty acids, such as: palmitoleic (C16:1), oleic (C18:1), linoleic (C18:2) and linolenic (C18:3) acids. Saturated fatty acids such as palmitic (C16:0) and stearic (C18:0) acids are also present, but in low concentrations (Meng et al., 2009). In certain species, higher polyunsaturated fatty acids can be synthesized (Thomas *et al*., 1984). These compounds oxidize at a faster rate than hydrocarbons of petroleum diesel, hence they form sediments that affect the performance of the combustion engine. Another disadvantage of the use of these microorganisms is that, due to photosynthetic activity, they require large superficial areas. Also, they grow at a slower rate than oleaginous bacteria and yeasts (Meng et al., 2009).

Only a few species of **bacteria** can produce lipids composed of saturated triacylglycerols (TAGs) with one or two unsaturations, that can be used for prducing esters that can substitute the currently employed biodiesel (Shaojin *et al*., 2006). *Rhodococcus opacus* accumulates up to 87% dry weight of TAGs (Alvarez *et al*., 2000), mainly synthesized from hexadecanoic (C16:0) and octadecenoic (C18:1) acids (Walther et al., 2009). *E. coli* has been metabolically engineered in order to increase its fatty acid content at 2.5 g/L by eliminating the *fadD* gene encoding the fatty acyl-CoA synthetase, and by over-expressing the genes encoding ACC and thioesterase (Lu *et al.*, 2008).

Another source of suitable lipids are **filamentous fungi,** such as *Humicola lanuginosa* (75% dry weight) (Meng *et al*., 2009), while in *Aspergillus oryzae* a lower level of lipids was found (57% dry weight) (Meng *et al*., 2009), and in *Mucor mucedo* a lipid content of 62% dry weight was obtained (Ratledge *et al.*, 2002).

Yeasts that can produce more than 20% dry weight in lipids are classified as **oleaginous yeasts**. *Lipomyces starkeyi* can contain 68% dry weight of lipids (Vicente *et al*., 2010), and *Rhodosporidium toruloides* produces 58% dry weight of lipids (Liu et al., 2007). This yeast can be cultivated in a fed-batch mode to achieve a higher productivity of lipids (0.54 g/(Lh)), when compared to batch fermentations (Li et al., 2007).

1.3.1 Direct advanced biodiesel production from genetically engineered cell factories

The activated fatty acids (FAs) found in major proportion in *S. cerevisiae* are FAs with chain lengths of 16 and 18 carbon atoms, saturated and monounsaturated with a very low amount (less than 2%) of myristic acid (C14:0) (Valle-Rodríguez *et al*., 2014). In an *in vivo* reaction process utilizing these FAs and yeast fermentation ethanol, fatty acid ethyl esters can be produced, which represent advanced microbial biodiesel (Table 2).

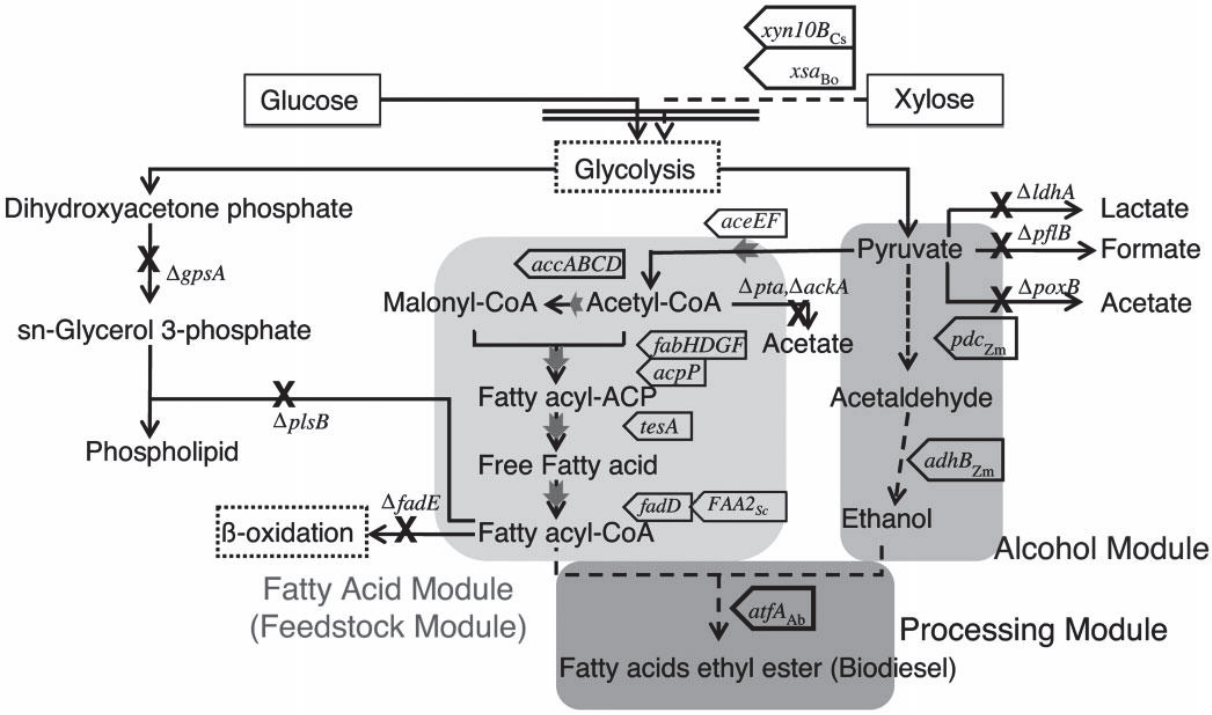
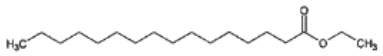
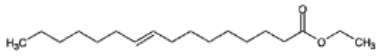
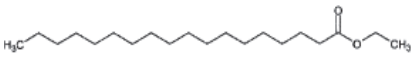
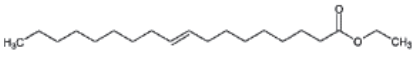


Figure 4. Overview of the engineered biochemical pathways for producing advanced biodiesel from hemicelluloses or glucose in recombinant *E. coli*. Overexpressed genes or operons are indicated in thick arrows; deleted or attenuated genes are indicated by crosses; heterologous pathways are highlighted in dashed lines; introduced heterologous genes are *xyn10B* (*C. stercorarium*), *xsa* (*B. ovatus*), *pdc* and *adhB* (*Z. mobilis*), *atfA* (*A. baylyi*) and *FAA2* (*S. cerevisiae*). (taken from Shi *et al.*, 2011).

Table 2. Composition of advanced microbial biodiesel produced in *S. cerevisiae* strains CEN.PK.

Common name	**Semicondensed formula**	**Developed formula**	**Composition** (% of total FAEEs)*
Ethyl palmitate	C16:0-C2		19.8 ± 3.1
Ethyl palmitoleate	C16:1-C2		55.1 ± 4.9
Ethyl stearate	C18:0-C2		7.1 ± 1.1
Ethyl oleate	C18:1-C2		15.5 ± 0.8

*Valle-Rodríguez *et al*., 2014

CHAPTER 2

2. Construction and validation of the microbial cell factories

2.1 Wax ester synthase selection

In order to be able to produce advanced microbial biodiesel in yeast, the first step taken was applying metabolic engineering through the introduction of the biochemical reaction in which the fatty acids could react with ethanol produced from alcoholic fermentation yielding fatty acid ethyl esters (FAEEs) (Figure 5).

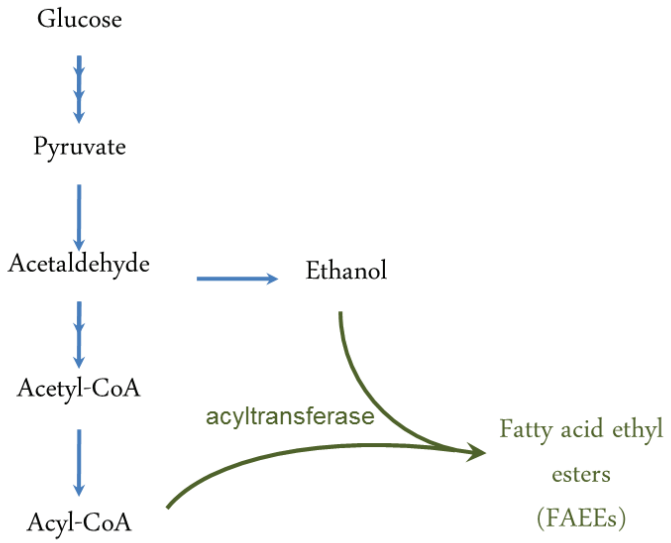


Figure 5. General scheme of the biochemical reactions occurring in mutant *S. cerevisiae* cells for the production of advanced microbial biodiesel (FAEEs). Homologous pathways stated with blue reaction arrows and the heterologous reaction stated in green.

CoA-dependent *O*-acyltransferases can most efficiently carry on this reaction. From these enzymes, wax ester synthases (WSs) are the ones that have been most well studied and characterized. They can synthesize wax esters from alcohols and acyl-CoA, some of them being promiscuous and others being specific for long-chain versions of these two substrates (Stöveken *et al*., 2005), they are widely applied in the industry of lubricants, printing inks, candles and cosmetics. Due to their very long chain, they have very high fusion points (Jetter *et al*., 2008) but those ones produced from long-chain fatty acids and ethanol possess good properties, so that they can replace diesel as they match american and asian biodiesel standards (Westfall *et al*., 2011).

WSs are classified in three unrelated subfamilies each one found in bacteria, plants and mammals (Jetter *et al*., 2008). In this initial approach, we tested four different WSs from bacteria and one from a mammal (Table 3), since a

previously identified WS in *Arabidopsis* did not present activity when introduced and expressed in *S. cerevisiae* (Lardizabal *et al*., 2000). One of the WSs considered was a WS/DGAT from *Acinetobacter baylyi* ADP1 that has been expressed in *E. coli* for ester production (Duan *et al*., 2011; Kalscheuer *et al*., 2006). WS homologs are frequently found in the genomes of actinomycetes such as *Rhodococcus* (Wältermann *et al*., 2007) or in the genome databases of several marine bacteria like *Marinobacter* (Holtzapple *et al*., 2007) and *Psychrobacter* (Wältermann *et al*., 2007). Few reports are available about WSs from mammals but we based the mammalian WS selection on a report of the isolation and characterization of a WS from *Mus musculus* C57BL/6, which was expressed in human embryonic kidney (HEK) 293 cells (Cheng *et al*., 2004). The genes of these enzymes were codon optimized for yeast and cloned in a plasmid under a strong constitutive promoter *TEF1* and expressed in *S. cerevisiae* (Table 3).

Table 3. List of strains generated for the WS selection, their genotype and plasmid contained (Shi *et al*, 2012).

Strain	**Genotype**	**Plasmid**
CEN.PK 113-5D	*MAT*a *MAL2-8*c *SUC2 ura3-52*	--
CB0	CEN.PK 113-5D	pSP-GM2
CB1	CEN.PK 113-5D	pSP-GM2 carrying WS gene from *A. baylyi* ADP1
CB2	CEN.PK 113-5D	pSP-GM2 carrying WS gene from *M. hydrocarbonoclasticus* DSM 8798
CB3	CEN.PK 113-5D	pSP-GM2 carrying WS gene from *R. opacus* PD630
CB4	CEN.PK 113-5D	pSP-GM2 carrying WS gene from *M. musculus* C57BL/6
CB5	CEN.PK 113-5D	pSP-GM2 carrying WS gene from *P. arcticus* 273-4

CB2A	CEN.PK 113-5D	pSP-GM2 carrying WS gene from *M. hydrocarbonoclasticus* DSM 8798 and *ACC1*

The following step was the evaluation of the alcohol specificity through the measurement of the activity of these WSs when catalyzing *in vitro* reactions (Table 4). The WS showing the highest activity from ethanol (shortest chain alcohol produced in yeast) and up to 1-octadecanol, was the WS of *M. hydrocarbonoclasticus* WS2 (Wältermann *et al*., 2007). While other enzymes of *P. arcticus* 273-4 and *A. baylyi* ADP1 were the following ones having a lower activity when degrading ethanol. Hence, the FAEE titer obtained when extracting them from cultivated cells containing the specific WS, was the highest from all mutant strains, while there was no statistical difference in the ethanol concentration nor in its specific rate of production (Table 5).

Table 4. Comparison of acyl acceptor specificities of different WSs using palmitoyl-CoA as acyl donor (Shi *et al*, 2012).

Acyl acceptor	**Wax ester synthase activity (pmol · [mg cell extract · min]$^{-1}$)**			
	CB0	**CB1**	**CB2**	**CB3**
Ethanol	0.67 ± 0.15	4.6 ± 0.55	8.1 ± 1.87	2.7 ± 0.37
Butanol	0.42 ± 0.21	10.8 ± 1.60	14.6 ± 1.75	6.8 ± 0.82
1-Hexanol	0.72 ± 0.20	17.3 ± 2.04	33.8 ± 3.77	16.1 ± 2.29
1-Octanol	0.83 ± 0.19	23.0 ± 2.39	45.7 ± 4.51	32.3 ± 3.84
1-Decanol	0.75 ± 0.21	19.7 ± 3.11	41.1 ± 4.13	37.3 ± 3.90
1-Dodecanol	0.78 ± 0.17	31.8 ± 3.48	48.4 ± 4.56	36.7 ± 3.78
1-Tetradecanol	0.81 ± 0.27	45.0 ± 4.72	49.7 ± 4.38	33.5 ± 3.66
1-Hexadecanol	0.90 ± 0.20	41.6 ± 2.21	49.0 ± 3.65	28.9 ± 3.29
1-Octadecanol	0.77 ± 0.21	43.6 ± 2.21	40.1 ± 3.77	30.9 ± 3.12

Table 5. Kinetic parameters and productions in the WS-expressing and reference strains (Shi *et al*, 2012).

	CB0	CB1	CB2	CB3	CB4	CB5
Specific growth rate (/h)	0.44 ± 0.01	0.32 ± 0.01	0.36 ± 0.02	0.34 ± 0.01	0.37 ± 0.01	0.33 ± 0.02
$q_{ethanol}$ (mmol/g/h)	43 ± 3	28 ± 3	40 ± 4	38 ± 3	36 ± 4	38 ± 4
Ethanol (g/L)[a]	6.9 ± 0.2	5.9 ± 0.2	6.6 ± 0.3	6.5 ± 0.2	6.2 ± 0.3	4.9 ± 0.2
FAEEs (mg/L)	nd	5.0 ± 0.8	6.3 ± 1.2	2.1 ± 0.3	1.3 ± 0.2	2.3 ± 0.4

CHAPTER 3

3. Strategies for increasing the FAEE production

3.1 Metabolic engineering by abolishing non-essential fatty acid utilization pathways

By increasing the availability of free FAs in the cytosol of yeasts, a theoretical increment in FAEE titers was planned. In *S. cerevisiae*, FAs are used to synthesize phospholipids and storage neutral lipids. Phospholipids constitute the cell membrane, which is essential to the cell; storage neutral lipids, such as triacylglycerols (TAGs) and steryl esters (SEs), are the main FA reserves (Murphy *et al*., 1999), they are found in 50-50 proportion and can constitute up to 97% of the storage lipid content of the cell (Leber *et al*., 1994), and are not essential to the cell (Sandager *et al*., 2002). The syntheses of TAGs and SEs are widespread among eukaryotic organisms, whereas they do not occur in *E. coli*. In *S. cerevisiae*, TAGs can be synthesized from FAs by two different enzymes, acyl-CoA:diacylglycerol acyltransferase (DGAT, encoded by *DGA1*) or lecithin:cholesterol acyltransferase (LCAT, encoded by *LRO1*); SEs are formed from FAs through the action of an acyl-CoA:sterol acyltransferase (ASAT, encoded by *ARE1* and *ARE2*) (Figure 6). Syntheses of TAGs and SEs would compete with a microbial biodiesel forming pathway for utilizing the same substrate, FAs. It is supposed that deleting these genes (i.e. *DGA1*, *LRO1*, *ARE1* and *ARE2*) to block the storage neutral lipid forming pathways would be beneficial for the accumulation of FAs and in turn increase the production of microbial biodiesel. Previous studies have shown that storage neutral lipid synthesis is non-essential in yeast and the quadruple mutant, in which *DGA1*, *LRO1*, *ARE1* and *ARE2* were disrupted, showed a 2.5-fold increase in FAs (Sandager *et al*., 2002). This approach included: (a) deletion of *DGA1*, *LRO1*, *ARE1* and *ARE2*, to block storage neutral lipid formation; (b) deletion of *POX1* to avoid FA degradation; and (c) a combination of both strategies by deleting all five genes (Figure 6).

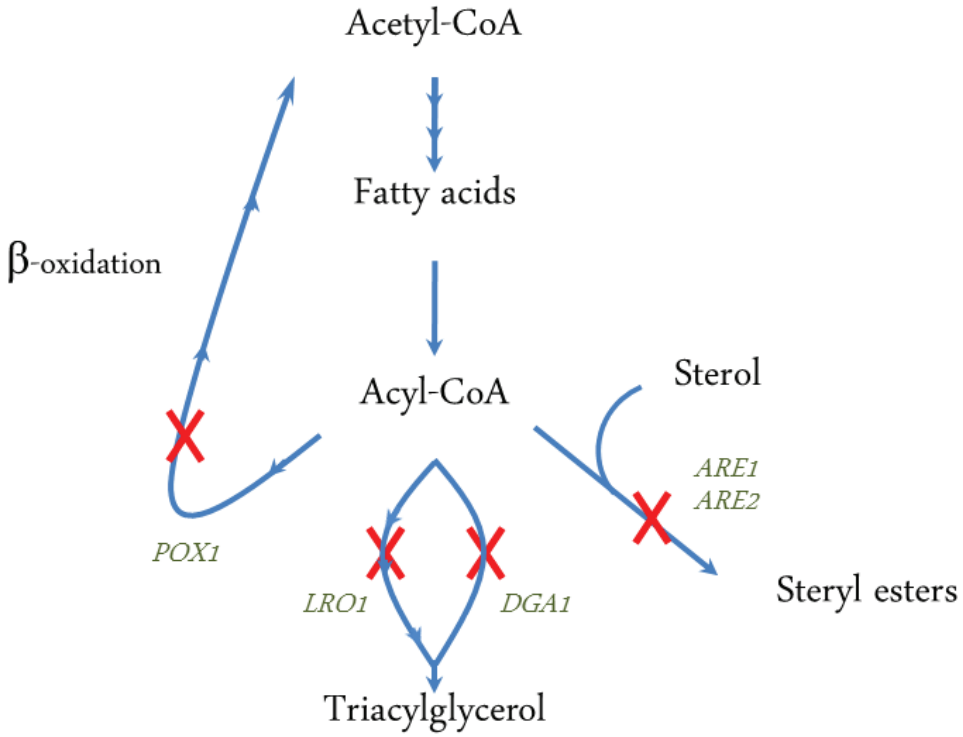


Figure 6. Abolished biochemical pathways in these yeast strains. Genes that were deleted in the course of this work are stated.

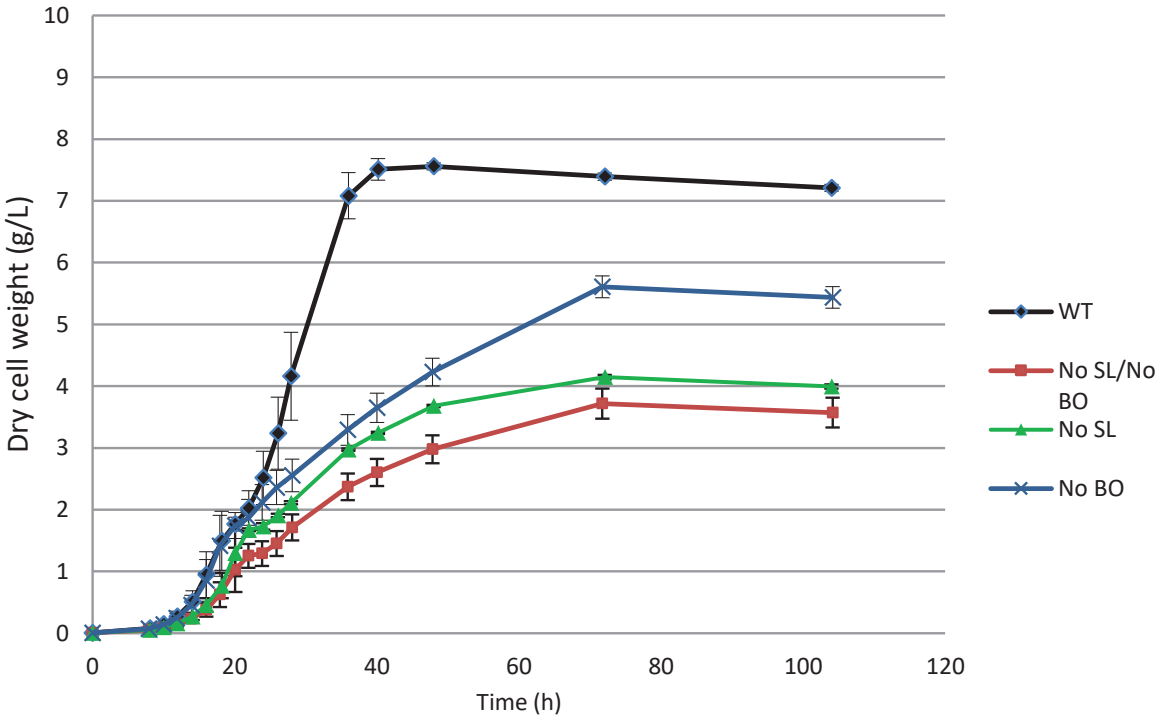


Figure 7. Biomass growth curves of the mutant strains obtained in aerobic (1 vvm air flow) batch bioreactor fermentations started with an inoculation of an initial OD^{600} of 0.01.

The mutant strains of gene knock-out exhibited a lower growth rate and a reduction on biomass formation when compared to the wild-type strain (Figure 7). The final dry cell weight and specific growth rate decreased as increasing the number of gene deletions (Table 6). An exception was the strain NoBO were no statistical differences were found in comparison to WT. In the case of NoSL, a 15% reduction in specific growth rate was achieved, while NoSL/NoBO presented a 20% decrease. The mutant strains exhibited a higher maximum specific glucose uptake rate when compared to the wild-type, and lower maximum ethanol production rates, suggesting that the mutant strains performed a more respiration metabolism than fermentative. The total production of ethanol and yield of ethanol formation were lower in the mutant strains when compared to the wild-type. The lowest values achieved were on strain NoSL, when compared to the wild-type strain where no genetic modification was performed (Table 5).

Table 6. Kinetic parameters, yields and concentrations of cell dry weight (DCW), ethanol (Eth) and glycerol (Gly) obtained in aerobic (1 vvm air flow) batch bioreactor fermentations.

	Specific rates			Yields			Concentrations (in glucose depletion phase)		
Strain	**μ_{max} (h^{-1})**	**$r_{s\,max}$ ($g_s\ g_{DW}^{-1}\ h^{-1}$)**	**$r_{e\,max}$ ($g_e\ g_{DW}^{-1}\ h^{-1}$)**	**Y_{sx} (g/g)**	**Y_{se} (g/g)**	**Y_{sg} (g/g)**	**DCW (g/L)**	**Eth (g/L)**	**Gly (g/L)**
NoSL *(are1Δ are2Δ dga1Δ lro1Δ)*	0.258 ± 0.013	3.14 ± 0.112	0.764 ± 0.002	0.082 ± 0.007	0.243 ± 0.008	0.091 ± 0.008	1.67 ± 0.13	4.86 ± 0.16	1.82 ± 0.20
NoSL/NoBO *(are1Δ are2Δ dga1Δ lro1Δ pox1Δ)*	0.245 ± 0.016	3.60 ± 0.032	0.927 ± 0.012	0.068 ± 0.004	0.258 ± 0.001	0.125 ± 0.007	1.46 ± 0.09	5.16 ± 0.14	2.50 ± 0.60
NoBO *(pox1Δ)*	0.297 ± 0.023	3.45 ± 0.079	1.060 ± 0.015	0.086 ± 0.009	0.307 ± 0.003	0.097 ± 0.007	1.72 ± 0.17	6.14 ± 0.06	1.94 ± 0.34
Wild-Type CEN.PK 113-5D	0.304 ± 0.026	3.14 ± 0.086	1.060 ± 0.017	0.097 ± 0.011	0.338 ± 0.004	0.068 ± 0.001	1.76 ± 0.19	6.76 ± 0.02	1.36 ± 0.60

Due to their inability to store fatty acids or reduce their chain length through the β-beta-oxidation pathway, these mutant strains presented a liposensitivity condition (Table 7). The experiment that was conducted was the addition of oleic acid to the growth medium, since it was determined previously that this fatty acid was the one that represented a higher sensitivity to these cells (determined previously). The wild-type strain did not exhibit a decrease in its specific growth rate when oleic acid was added to the cultivation medium. The strain that was less liposensitive and hence had no big affectation from oleic acid addition, was NoBO where a 7% reduction in specific growth rate was found when compared to this same strain without addition of oleic acid. In the meantime, in the storage neutral lipid strains the effect was more prominent: in NoSL a decrease of 13% was encountered when compared to no addition of oleic acid, while in NoSL/NoBO the specific growth rate was reduced in 15%.

Table 7. Specific growth rate of the strains studied in normal medium and in medium supplemented with oleic acid (OA, 0.01%).

Strain	**μ (h^{-1})**
NoSL	*0,340*
NoSL-NoB	*0,317*
NoBO	*0,391*
WT	*0,389*
NoSL+OA	*0,297*
NoSL-NoBO+OA	*0,270*
NoBO+OA	*0,362*
WT+OA	*0,390*

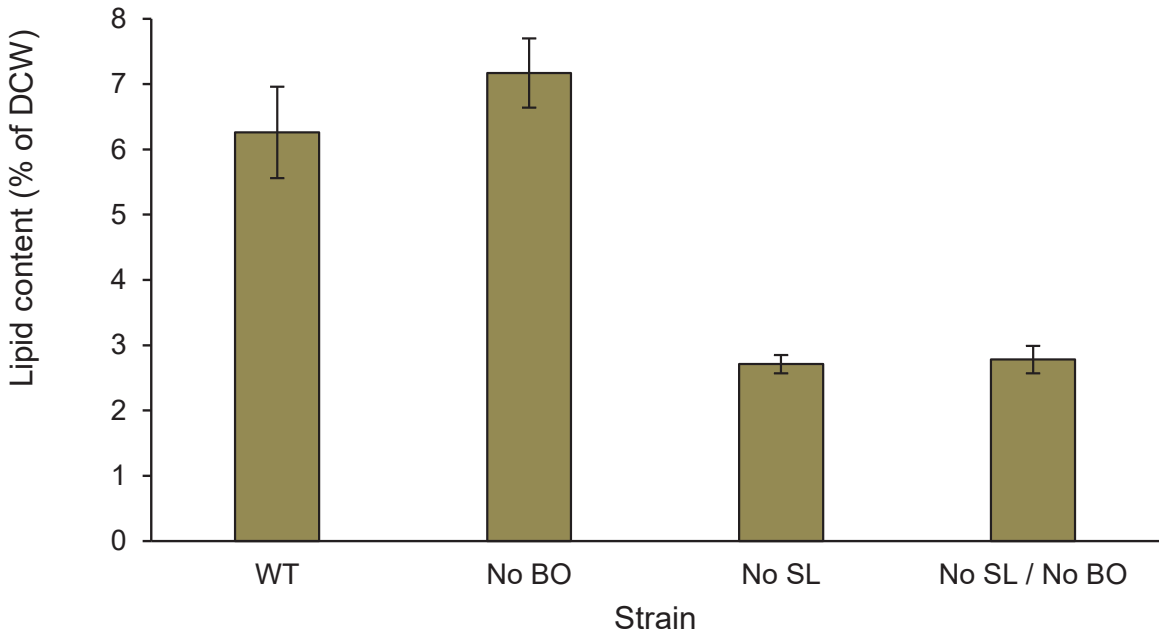


Figure 8. Total lipid analysis (A) of mutant strains and wild-type. The total lipid contents are expressed as percentage of biomass dry weight. The FA compositions are expressed as percentage of total FAs in cells. Strains were incubated in SD-CSM medium until early stationary phase. Abbreviations: DCW, dry cell weight; The reported results are the average of three replicate experiments and error bars show the standard deviation. Asterisks (*) indicate significant difference (Fisher's LSD test) compared to CEN.PK 113-5D strain as control; P-values are: *<0.05 and ** <0.01. No statistically differences between strains were found in the FA fractions of C16:0 and C16:1 (ANOVA test).

Figure 8 shows the total lipid content of these four strains: WT (CEN.PK 113-5D), NoSL, NoSL/NoBO, and NoBO. We found that the total lipid content is strongly affected in the knock-out strains. Because TAGs and SEs constitute a major part of the lipids in stationary phase cells, NoSL and NoSL/NoBO, which are devoid of both TAGs and SEs, both have a significant reduction in total lipid, which is about 50% of wild type. Results are similar to a previous report on a strain unable to produce TAGs and SEs (Sandager *et al*., 2002). Interestingly, there is a slight increase (16%) of total lipid content in NoBO, which is incapable of β–oxidation. The increase could be explained by accumulation of FAs derived from degradation of membrane lipids (phospholipids) that would otherwise enter the β-oxidation pathway (Hernawan *et al*., 1995). The positive effect on lipid accumulation caused by blocking β–oxidation is reported for the first time in *S. cerevisiae*, but has been described previously for *Yarrowia lipolytica*, in which a 4-fold increase was found (Beopoulos *et al*., 2008) and for the plant *Arabidopsis* with an increase in the amount of TAG (Slocombe *et al*., 2009).

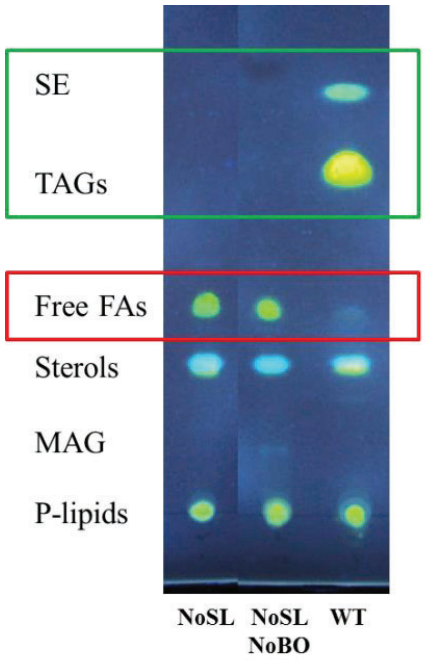


Figure 9. Photograph of TLC run for separating the lipidic classes synthesized in the mutant strains. Abbreviations are SE: steryl esters, TAGs: triacylglycerols, FAs: fatty acids and MAG: monoacylglycerol.

Saturated and monounsaturated FAs with C_{16} and C_{18} straight carbon chains (palmitic, palmitoleic, stearic, and oleic acids) accumulate in major proportion in yeast. The natural chain length of FAs is also reflected in the composition of FA derivatives produced in yeast, such as FAEEs and alkanes (Schirmer *et al*., 2010). To see whether the gene deletions would affect the balance between different FAs, we proceeded to analyze the FA composition in the different strains (Figure 11). There was no significant difference in the FA composition between CEN.PK 113-5D and NoBO. In contrast, the fraction of FAs with a carbon chain length of 18 in the storage neutral lipid free strain (NoSL and NoSL/NoBO) was around 36%, which is higher than the value of 22% in WT.

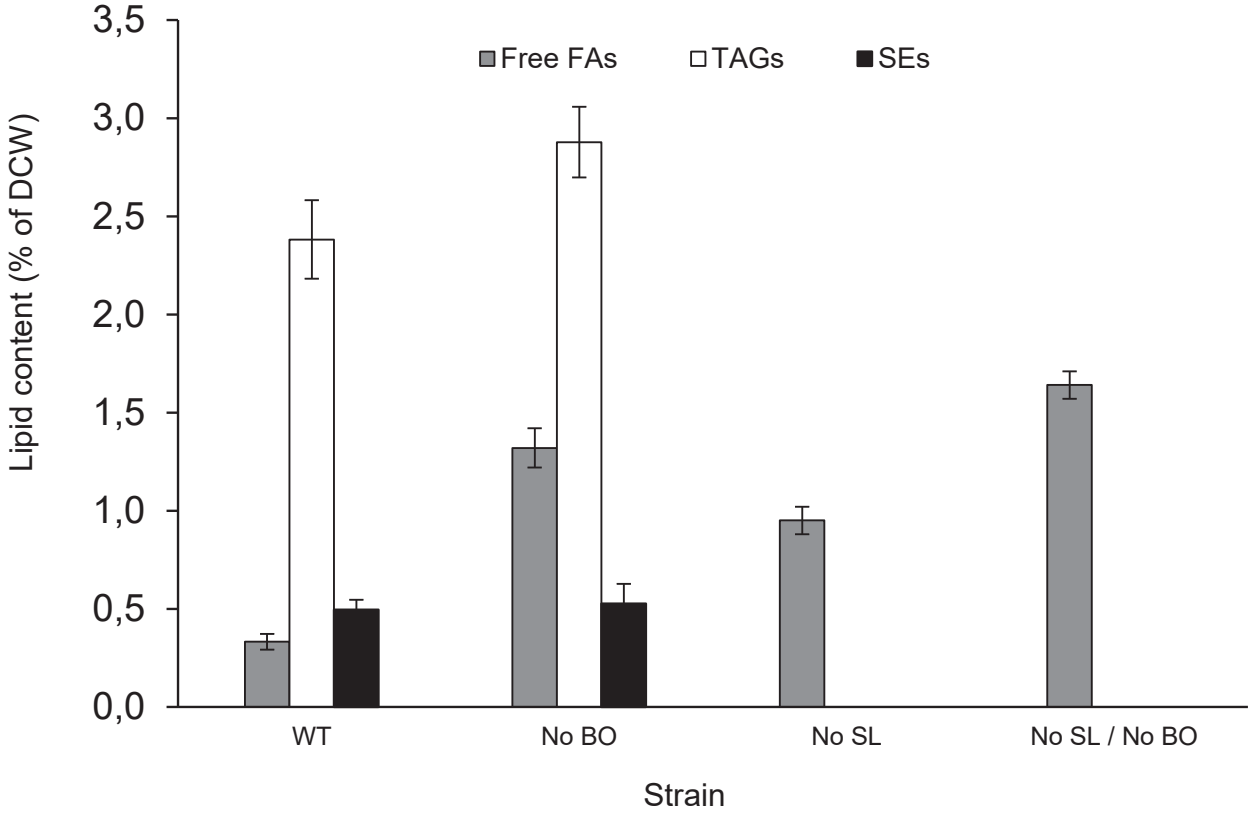


Figure 10. Fraction of FAs, TAGs and SEs as percentage of total lipids for each deletion strain and the wild-type as control strain. Cells were cultivated in SD-CSM medium in shake flasks and harvested in early stationary phase. DCW: dry cell weight. The reported results are the average of three replicate experiments and error bars show the standard deviation. Asterisks (*) indicate significant difference (Fisher's LSD test) compared to CEN.PK 113-5D strain as control; P-values are: *<0.05 and **<0.01.

The biosynthesis of storage neutral lipids in mutant strains NoSL and NoSL/NoBO did not occur (Figure 9), as it can be compared to the wild-type strain. Meanwhile, FA concentrations were visibly higher in these mutant strains. In figure 10, it can be observed a 3-fold increase of FAs in the strain NoSL, compared to WT. In a previous study, a similar result was found, where elimination of TAGs and SEs syntheses resulted in a yeast strain with a 2.5-fold increase in the FA content of *S. cerevisiae* (Sandager *et al*., 2002). The abolishment of the β–oxidation pathway (*POX1* deletion, strain NoBO), carried a 4-fold increase in accumulation of FAs. By combining the disruption of these three pathways of β–oxidation and syntheses of TAGs and SEs (strain NoSL/NoBO), FA production was enhanced in a 5-fold increase. Aimed to avoid storage neutral lipids, the

β–oxidation deletion increased the FAs availability by 73%. This indicates the important role of β–oxidation for accumulation of FAs and its derivatives in *S. cerevisiae*. It has also been observed in *E. coli* (Lu *et al*., 2008; Steen *et al*., 2010) and *Yarrowia lipolytica* (Dmochowska *et al*., 1990) that abolishing the β–oxidation pathway leads to a positive effect on FA accumulation.

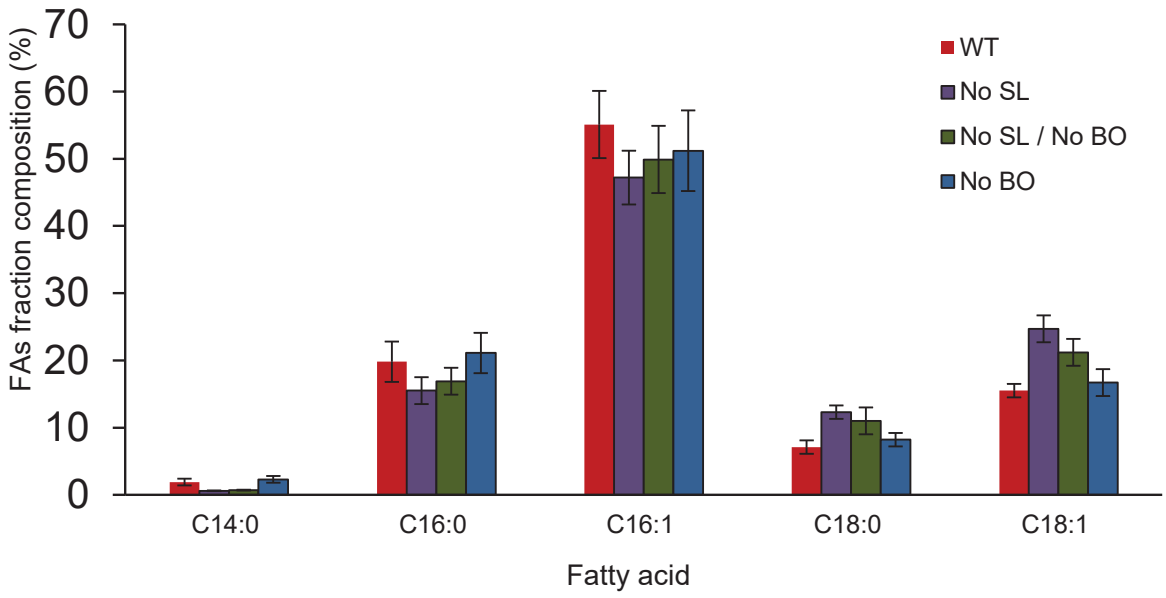


Figure 11. FA composition analysis of mutant strains and wild-type. The total lipid contents are expressed as percentage of biomass dry weight. The FA compositions are expressed as percentage of total FAs in cells. Strains were incubated in SD-CSM medium until early stationary phase. Abbreviations: DCW, dry cell weight; C12:0, lauric acid; C14:0, myristic acid; C16:0, palmitic acid; C16:1, palmitoleic acid; C18:0, stearic acid; C18:1, oleic acid; C20:0, arachidic acid. The reported results are the average of three replicate experiments and error bars show the standard deviation. Asterisks (*) indicate significant difference (Fisher's LSD test) compared to CEN.PK 113-5D strain as control; P-values are: *<0.05 and **<0.01. No statistically differences between strains were found in the FA fractions of C16:0 and C16:1 (ANOVA test).

3.2 Integrating the wax ester synthase gene and increasing its copy number

The next strategy on the improvement of biosynthesis of FAEEs and the stability of the host cell as microbial cell factory was the genomic integration and augment of copy number of the *ws2* gene in *S. cerevisiae* (Figure 12). By achieving the replacement of the plasmid-based mutants with a genomic integration strain supported by a chromosome engineering method, stability of the microbial biodiesel production process could be obtained.

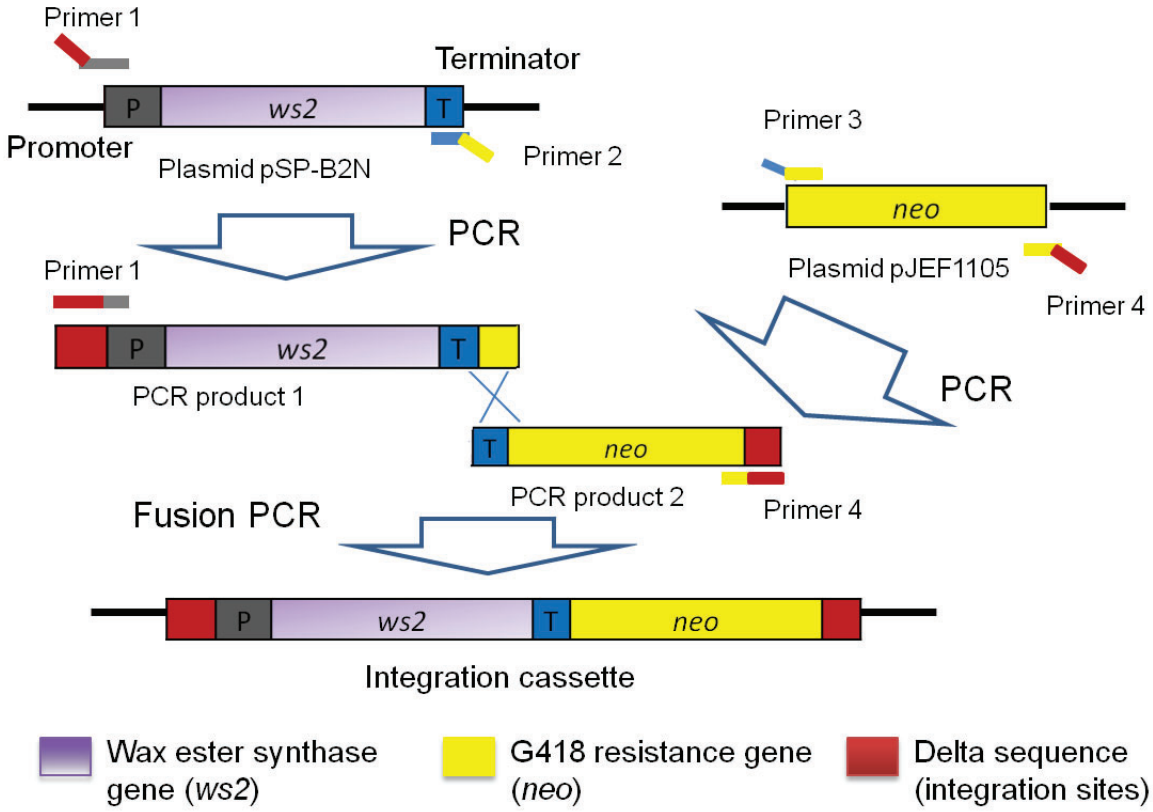


Figure 12. Scheme of assembling the chromosomal integration cassette, which is obtained by fusion PCR and contains the wax ester synthase gene (*ws2*) controlled by the *PGK1* promoter and a selectable marker (*neo*^{r}). This cassette is flanked with homologous sequences to delta(δ)-sites for facilitating genome integrations through homologous recombination (from Shi *et al*, 2014).

The primordial requirement for this subproject was an integration cassette containing the gene of interest (wax ester synthase gene, *ws2*) and the gene encoding the antibiotic marker (*neo*^{r}), flanked on both sides by identical short regions of the delta(δ) sequence that exists in multiple copies in the *S. cerevisiae* genome (Lesage *et al*., 2005). The approach takes advantage of multiple copies of delta (δ) regions in the genome and the efficient recombination in yeast under the pressure of antibiotics. This integration

cassette was transformed and inserted into the chromosome of strain CEN.PK 113-5D (Figure 12).

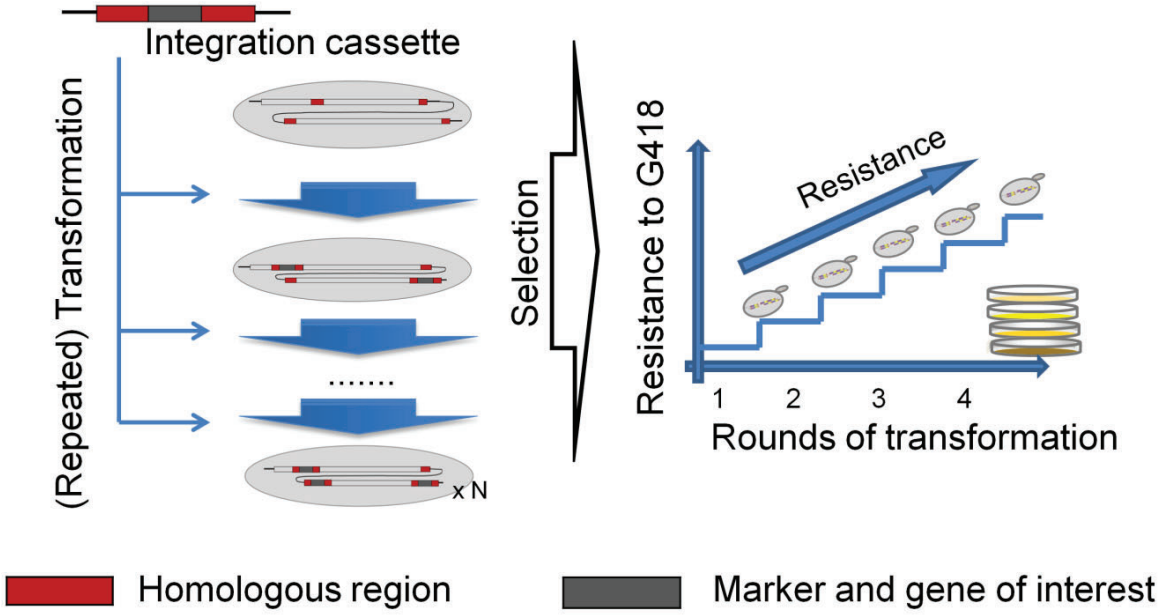


Figure 13. Schematic representation of the construction of delta (δ)-integration strains. Multiple gene copies are obtained by selection on plates with increasing antibiotics concentration after repeated transformation with integration cassette that is flanked with homologous sequences to delta (δ)-sites. Each integration event leads to integration of one antibiotic resistance gene and one gene of interest. Therefore, strains with desired gene copy numbers ($\times N$) are generated through plating on certain antibiotic concentrations (from Shi *et al*, 2014).

As shown in Figure 13, during the evolution process the increasing G418 stress was applied to augment the number of interest gene copies and to prevent their loss due to possible homologous recombination, which had been reported previously in delta (δ)-integration strains (Lee *et al*., 1997; Wang *et al*., 1996).

The production of FAEEs increased remarkably in strains that were selected from higher G418 concentrations (Figure 14). In strain CB2I20 able to grow at 20 mg/ml G418, the accumulated FAEE concentration (34 mg/l) was around 4.5-fold higher compared with strain CB2I0.2, firstly evolved and selected from a plate containing 0.2 mg/ml G418.

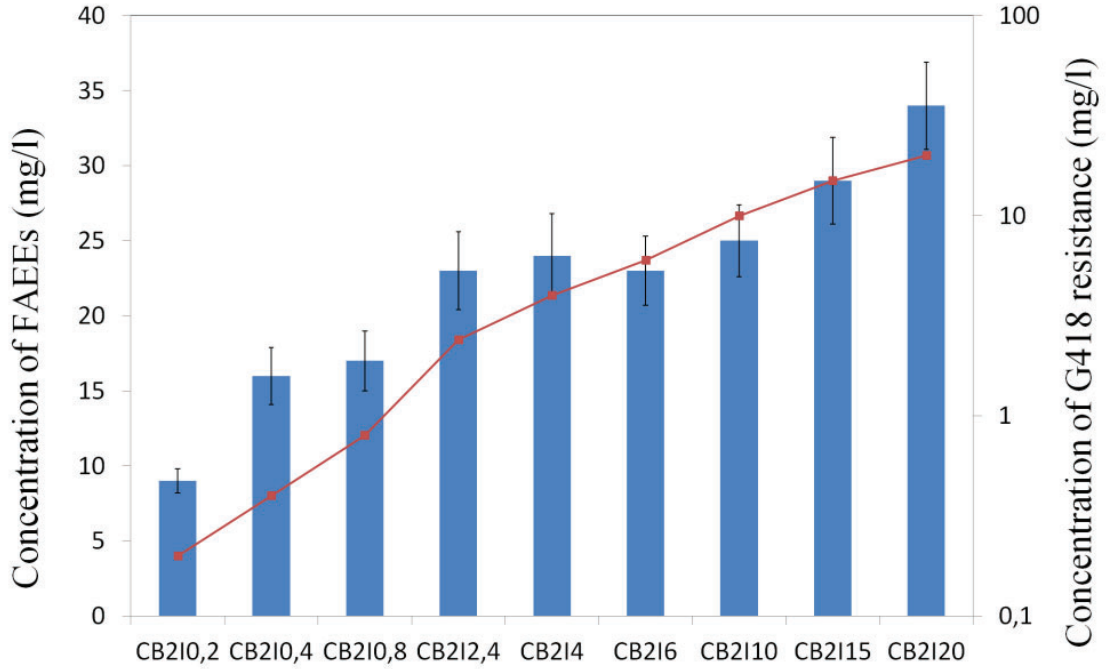


Figure 14. Relationship of FAEE production and the concentration of G418 used for strain selection. CB2I0.2 represents a strain isolated from 0.2 mg/ml G418. The meaning is the same for the other strains. CB2N represents the strain carrying the *ws2* gene on a 2 μm based plasmid. All strains were cultured in shake flasks in SD medium containing 2% glucose supplemented with the corresponding concentration of G418. Values are the mean of three independent assays ± standard deviation.

As expected, strains that can grow at higher concentrations of G418 contained a higher chromosomal copy number of the integration construct (*ws2* and *neo*^{r} gene). The copy number of *ws2* in CB2I20 was approx. 5.4 per cell (Table 8), which is similar to the copy number of certain 2 μm based plasmids (Chen *et al.*, 2012). The ratio of *ws2* and *neo*^{r} copy number is close to 1 in most of the cases. Compared to the change of chromosomal copy number (*ws2* and *neo*^{r} genes), the gene expression changed to a much larger extent. For example, compared to strain CB2I0.2, there is only a 3-fold increase in the copy number of *ws2* in CB2I20, but there is a 24-fold up-regulation of *ws2* transcription. A 10-fold up-regulation of WS activity was observed in strain CB2I20, which contributes to a 5-fold increase in FAEE titers and 7.5-fold increase in FAEE yield (Figure 14).

Table 8. Gene copy numbers, relative gene expression levels and activity of wax ester synthase during exponential phase for a selection of integration transformants isolated from different G418 concentrations. The average values with strain CB2I0.2 were set as unit for gene expression. Values are the mean of three independent assays ± standard deviation.

Strain	Gene	Gene copy no.	Relative gene expression	Enzyme activity [pmol/(mg extract•min)]
CB2I0.2	*ws2*	1.84 ± 0.11	1 ± 0.12	6.4 ± 0.71
	neo^r	1.72 ± 0.30	1 ± 0.15	-
CB2I4	*ws2*	2.61 ± 0.13	5.66 ± 0.6	31.8 ± 4.0
	neo^r	3.83 ± 0.21	5.55 ± 0.6	-
CB2I20	*ws2*	5.39 ± 0.55	24.6 ± 2.8	62.0 ± 0.8
	neo^r	5.51 ± 0.46	22.87 ± 2.5	-

CHAPTER 4

Protein Engineering

Cases of study:
MhWS2 and Eeb1p

Directed evolution
and
Rational mutagenesis

4.1 Directed evolution of MhWS2 (wax ester synthase)

4.1.2 Lipotoxicity as fundament for selective method

Oleic acid toxicity was applied for developing a high-throughput selection system with the goal of isolating MhWS2 variants with higher catalytic activity for FAEE synthesis (Figure 15). In the first step, the parental ws2 gene was mutagenized and the yeast strain JV03 (storage-lipid free and unable of β-beta-oxidation) was transformed with the resulting library. The transformants harboring the mutant library were plated on oleic acid containing plates at a concentration above 10 mM, at which oleate represents a strong toxicity.

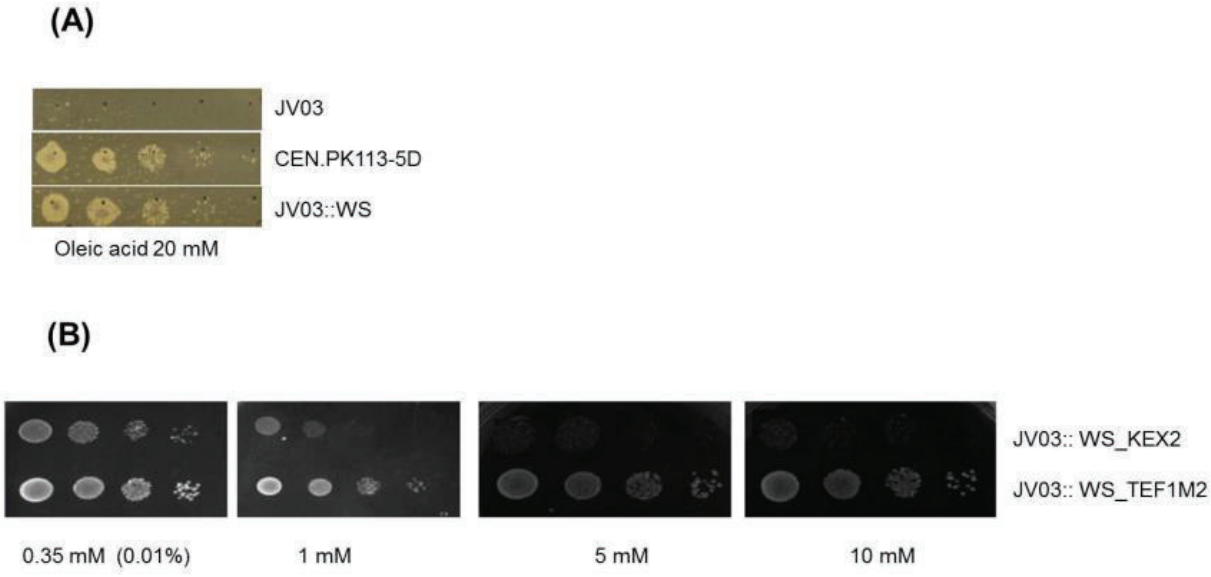


Figure 15. **Assessment of oleic acid toxicity**. **(A)** Growth of storage-lipid free strain JV03, parental strain CEN.PK 113-5D, and JV03 expressing *ws2* under control of the *PGK1* promoter (JV03::WS) was assessed on SD medium supplemented with 20 mM oleic acid. **(B)** Growth of strains expressing *ws2* under control of the *KEX2* or the *TEF1M2* promoter was assessed on SD medium supplemented with oleic acid at concentrations ranging from 0.35 mM to 10 mM. Cells were cultivated overnight in liquid media, adjusted to an OD_{600} of 0.5 and serially diluted in 1:10 steps, and 2 µl of the respective dilutions were spotted onto plates.

4.1.3 Screening and selection

By increasing oleic acid concentration in agar plates, a screening procedure was developed with the goal of selecting those successful variants with the assumption that beneficial residue substitutions in MhWS2 would make the

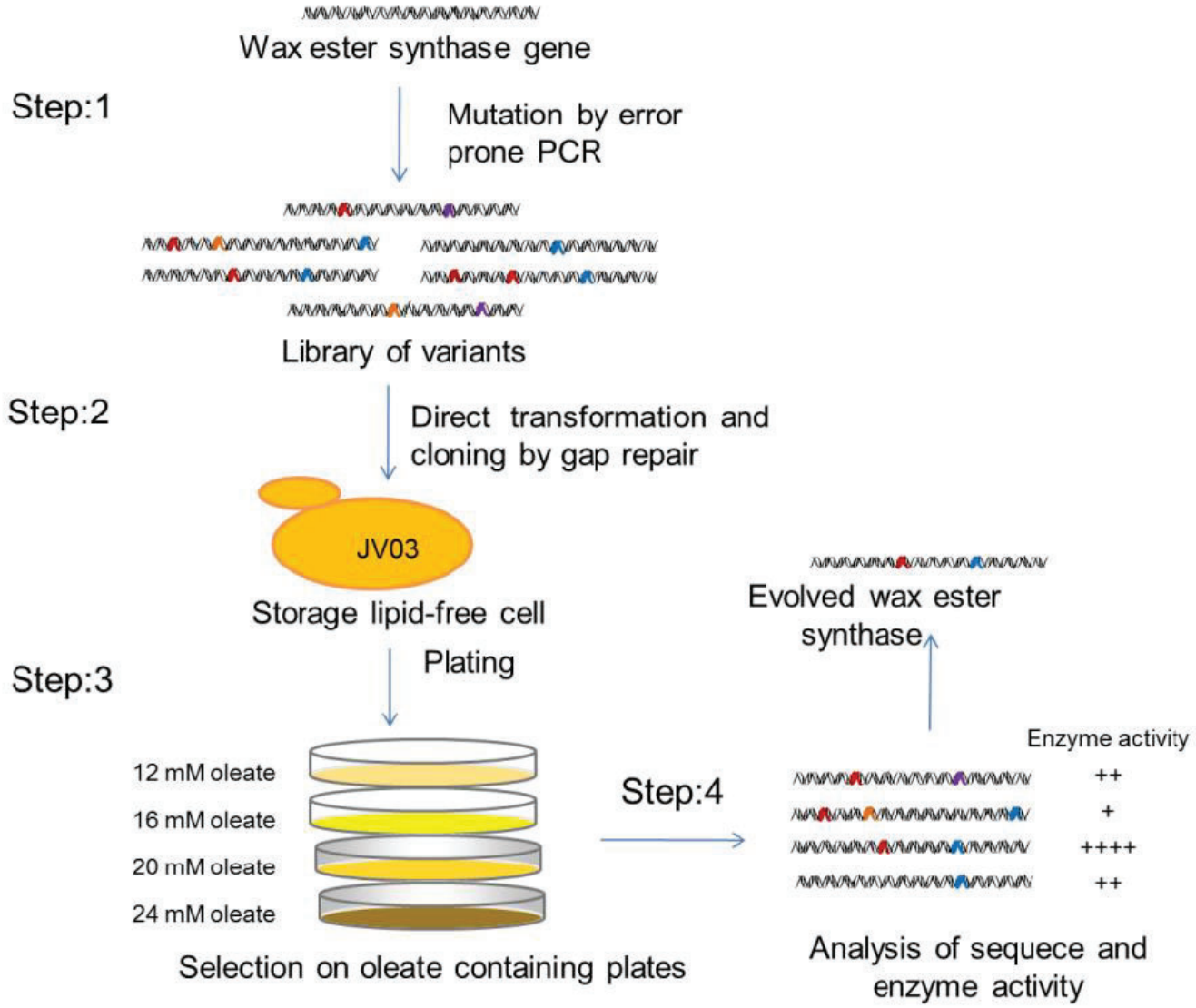
harboring cell more tolerant to oleic acid. Concentrations of oleic acid increased by steps of 4 mM from an initial value of 12 mM until a final value of 24 mM (Table 9). Then, the enzyme activities and sequences of selected variants were measured and analyzed. Figure 16 contains a process flow illustrating these procedures.

Table 9. Number of transformants formed in oleic acid containing plates after transforming JV03 with the *ws2* mutant library.

Oleic acid (mM)	Colony number
12	190
16	108
20	65
24	12

Oleic acid concentrations shown are the concentrations in SD-*ura* agar plates.

Figure 16. *(on next page)* **Schematic outline of selection of MhWS2 variants**. The mutation process was initiated with the parental *ws2* gene from *M. hydrocarbonoclasticus*. This parental gene was randomly mutagenized for generation of the library of mutants using error-prone PCR. The mutagenized library was used to transform yeast strain JV03 (storage-lipid-free and unable of β-beta-oxidation) and plated on different oleate concentrations. Mutants that were able to grow on plates containing higher concentration of oleic acid were selected; their enzyme activities and the sequences of the mutated genes were determined. This procedure can be repeated until the evolved protein exhibits the desired level of activity.



Twelve variants were selected (Table 10). They were chosen since they could cope with the highest pressure that the lipotoxic screening exerted on them. Out of this group, we then selected and sequenced the genes of four MhWS2 variants that showed statistical significance in the improvement of enzyme activity. We then analyzed their amino acid sequences, and we could identify five mutations (A35V, R132G, N165S, A344T, I462V) in WS2-v11, four mutations (A35V, L128V, A344T, I462V) in WS2-v6, two mutations (P352R, A344T) in WS2-v7 and we only identified one mutation (A344T) in WS2-v2. Residue positions were based on the numbering convention for the *M. hydrocarbonoclasticus* DSM 8798 WS2 wild-type enzyme (MhWS2), NCBI reference sequence ABO21021. Among these mutations, A35V and I462V were common mutations for WS2-v11 and WS2-v6; while A344T was a common residue substitution for all four identified variants.

Table 10. **Comparison of enzyme activities using palmitoyl-CoA and ethanol as substrates.** Activity increments in strains expressing MhWS2 variants are shown in relation to the wild-type enzyme.

Strains	Enzyme specific activity (pmol product/mg total protein/min)	Increment (%)
JV03:WS_KEX2	8.02 ± 0.53	0
WS2-v1	8.11 ± 1.60	1.01
WS2-v2	11.09 ± 1.94	38.3
WS2-v3	10.64 ± 4.08	32.7
WS2-v4	8.76 ± 0.27	9.2
WS2-v5	8.84 ± 0.33	10.2
WS2-v6	12.44 ± 0.74	55.1
WS2-v7	11.78 ± 2.41	46.9
WS2-v8	9.09 ± 4.86	13.3
WS2-v9	10.55 ± 0.80	31.6
WS2-v10	8.25 ± 1.14	2.9
WS2-v11	13.26 ± 2.88	65.3
WS2-v12	8.90 ± 0.53	10.9

4.1.4 Alcohol selectivity of selected variants.

The enzyme activity was once more determined, in order to further investigate whether the identified point mutations altered selectivity towards other alcohols

(Figure 17). The activity analysis indicated that short-chain alcohol substrates (such as ethanol and 1-butanol) were still poorer substrates compared to longer-chain alcohols (e.g. 1-hexanol, 1-octanol, 1-decanol, 1-dodecanol, 1-tetradecanol, 1-hexadecanol and 1-octadecanol). However, there was a clear shift in the preference towards alcohol as substrate. All selected variants (WS2-v11, WS2-v6, WS2-v7 and WS2-v2) presented an enhanced activity in the catalysis of shorter chain alcohols (C2 ethanol and C4 butanol) compared to the wild-type MhWS2 enzyme. For instance, the variant WS2-v2 had 38% enhancement in its catalytic activity of esterifying palmitoyl-CoA with ethanol (when compared to the wild-type), and the specific activity with butanol as substrate was increased 22%. There was a shift in preferred activity towards medium-chain alcohols (C6 hexanol to C14 tetradecanol) in all variants. However, the absolute values did not exceed the corresponding activity of wild-type MhWS2 for the specific alcohol. There were no statistical differences in catalytic activity when using 1-hexanol (C6) or 1-octanol (C8) for all variants. When utilizing long chain alcohols (C16 to C18) however, the enzyme activities of all variants were reduced significantly compared to wild-type MhWS2.

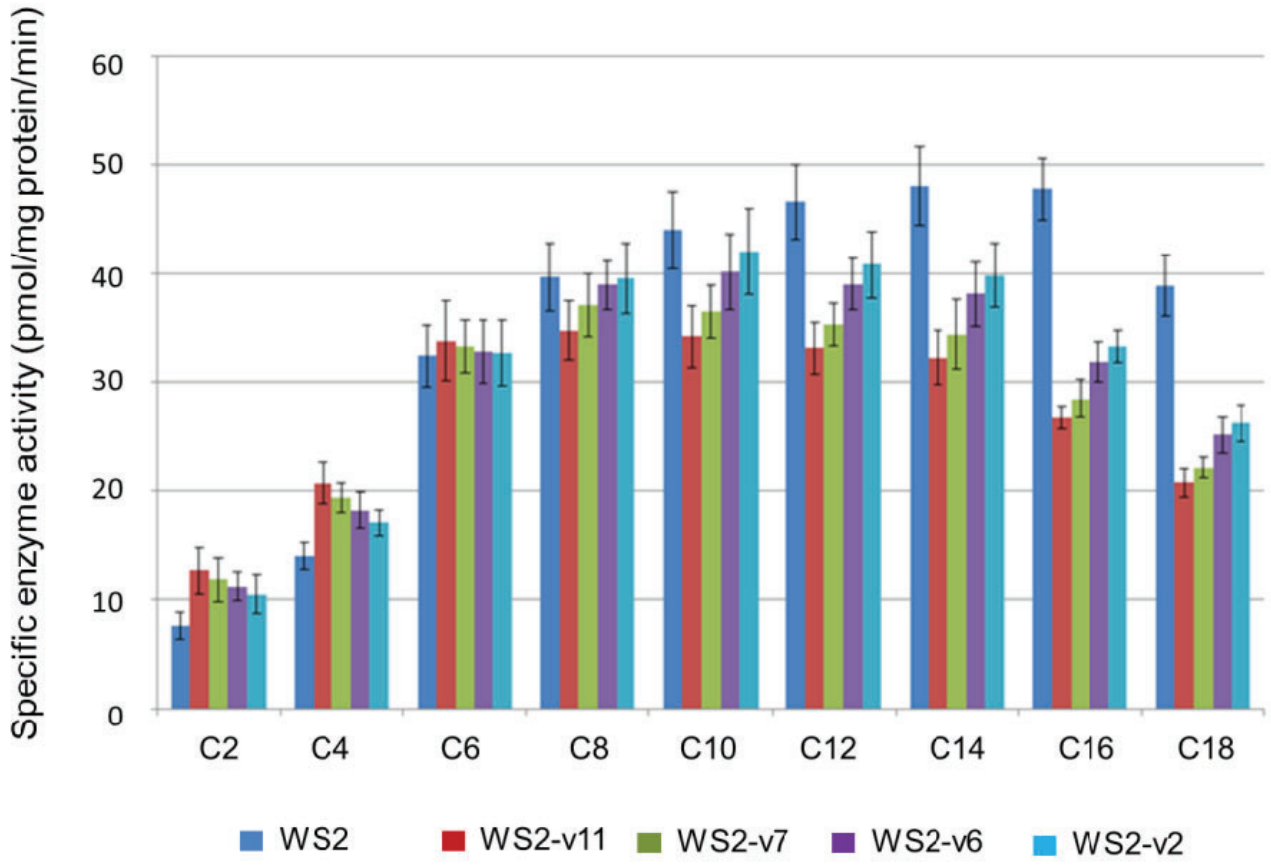


Figure 17. **Alcohol selectivity of selected variants of MhWS2**. Specific enzyme activities with

various alcohols as substrate were assessed in extracts derived from yeast cells expressing the respective variant. Results are depicted as the formation of product (fatty acid alkyl ester) per amount of total protein in the extract and time. Abbreviations: C2, ethanol; C4, 1-butanol; C6, 1-hexanol; C8, 1-octanol; C10, 1-decanol; C12, 1-dodecanol; C14, 1-butadecanol; C16, 1-hexadecanol; and C18, 1-octadecanol.

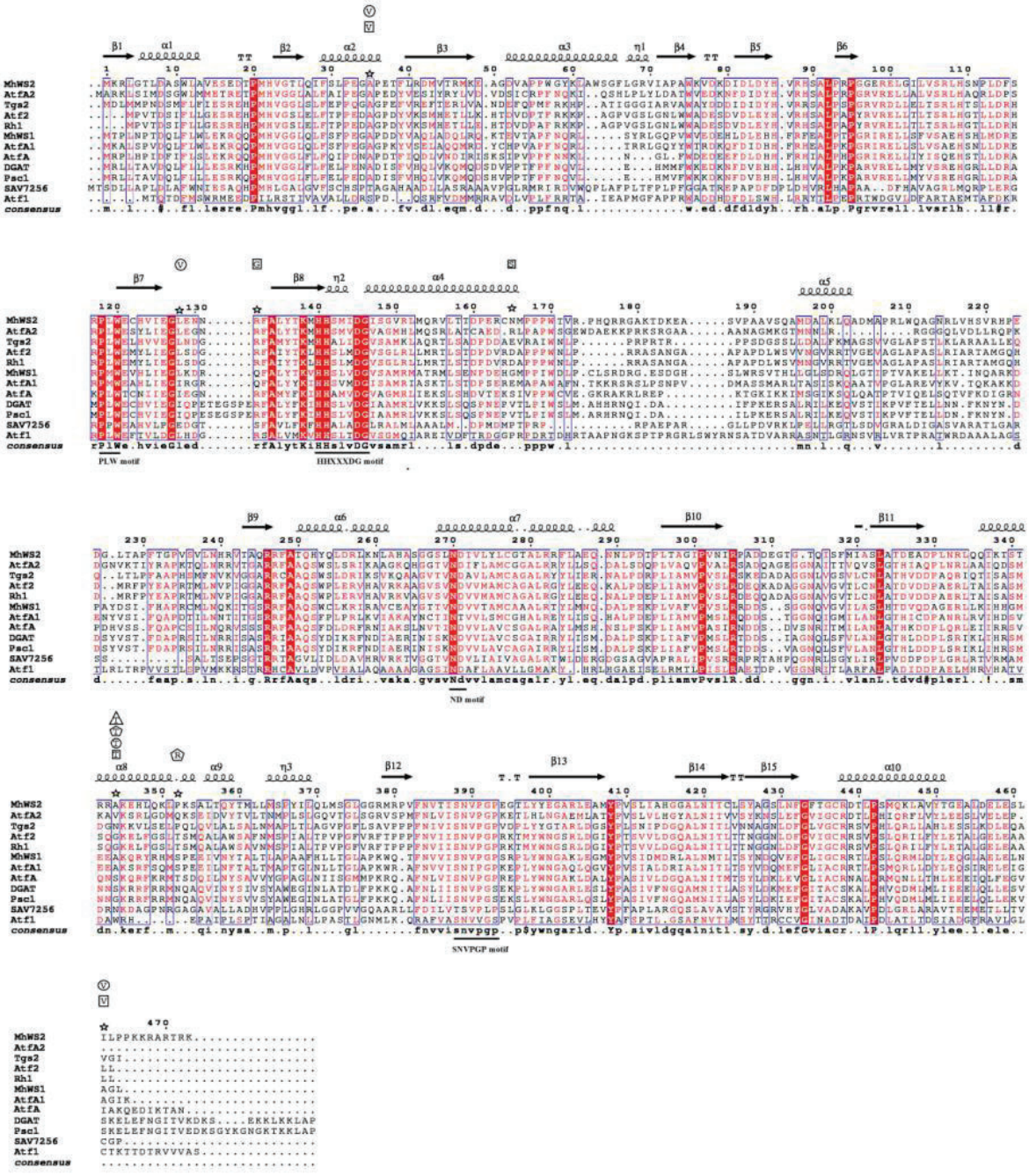
4.1.5 Effect of the residue substitutions on the structure of the enzyme *(in silico* analysis)

Three separate groups (Röttig and Steinbüchel, 2013a; Shi *et al*., 2012): mammalian WSs, plant WSs and bacterial WS/DGAT enzymes were identified through BLAST-searches of a wide distribution of MhWS2 homologous enzymes. WS/DGAT is a bifunctional enzyme exhibiting wax ester synthase (WS) as well as acyl-CoA:diacylglycerol acyltransferase (DGAT) activity (Information on all reported proteins showing WS activity in bacteria from the available literature, can be found in Table S3). In order to disclose the potentially conserved sequences in these WSs and predict the secondary structure of MhWS2, a multiple-sequence alignment of various reported WSs was performed (Figure 18). In analogy to previous studies, several reported conserved motifs could be found in these aligned enzymes, such as the HHXXXDG^{146} motif (Stöveken *et al*., 2009), the PLW^{120} motif (Villa *et al*., 2014), the ND^{271} motif (Villa *et al*., 2014), and the SNVPGP^{393} motif (Röttig and Steinbüchel, 2013b). It had been also reported that point mutations in L356 and M405 could alter small and medium-chain alcohol selectivity in the wax ester synthase (Ma1) from *M. aquaeolei* VT8 (Barney *et al*., 2015). In our case, these two residue positions were corresponding to M364 and A414 of MhWS2. In this current study, none of the identified mutations (A35V, L128V, R132G, N165S, A344T, P352R and I462V) were present in these motifs or the demonstrated active site; which let us confirm that the motifs and active site in the variants were still conserved since no alteration was found in the corresponding amino acids.

Figure 18. ***In silico* analysis of WSs and selected variants of MhWS2**. **(A)** Multiple sequence alignment of representative WS proteins from different organisms as described in Table S3, created using MultAlin and ESPript. Predicted secondary structural motifs of MhWS2 are schematically displayed above the alignment. Wax ester synthases provided are from *M. hydrocarbonoclasticus* DSM 8798 (MhWS1 and MhWS2), *Alcanivorax borkumensis* SK2 (AtfA1 and AtfA2), *M. tuberculosis* H37Rv (Tgs2), *Rhodococcus opacus* PD630 (Atf1 and Atf2), *Rhodococcus jostii* RHA1 (Rh1), *A. baylyi* ADP1 (AtfA), *Psychrobacter arcticus* 273-4 (DGAT), *Psychrobacter cryohalolentis* K5 (Psc1), and *Streptomyces avermitilis* MA-4680 (SAV7256). Squiggles represent helices; arrows indicate beta strands; TT letters show beta turns. Gaps are denoted by a dot. A residue that is weakly conserved is shown in red and as a lowercase letter in the consensus line. A residue that is highly conserved is highlighted in red and as an uppercase

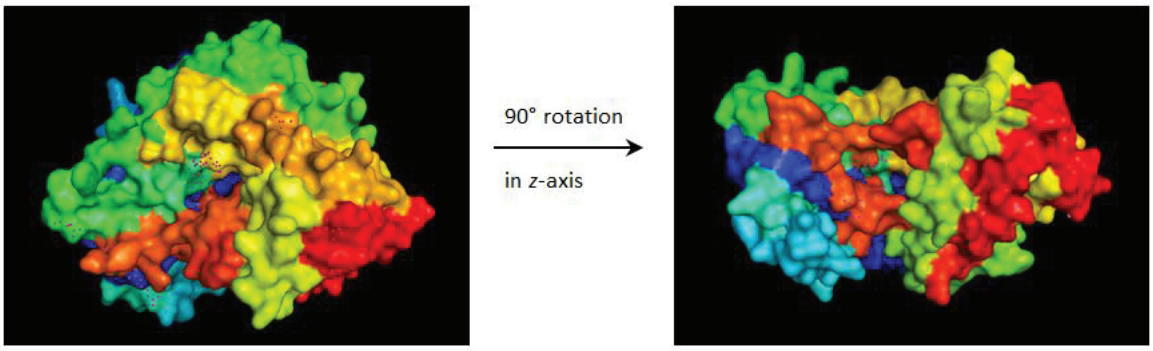
letter in the consensus line. Residues of >70% consensus level are indicated by a blue rectangle. Positions where mutations were detected in this study are marked with "☆", and the corresponding mutation is shown above the aligned sequences. In particular, residues marked with "□" are mutations found in WS2-v11; residues marked with "○" are mutations found in WS2-v6; residues marked with "⌂" are mutations found in WS2-v11; and the residue marked with "Δ" represents the mutation found in WS2-v2. **(B)** Surface structure prediction of the WS2 wild-type and selected variant WS2-v11 (Kelley and Sternberg, 2009). Sequences are presented in rainbow spectrum from protein head by N-terminus (dark blue) to protein tail by C-terminus (dark red). Residues that constitute the active site (HHXXXDG, H140, H141, D145 and G146) are indicated as gray dots.

A)

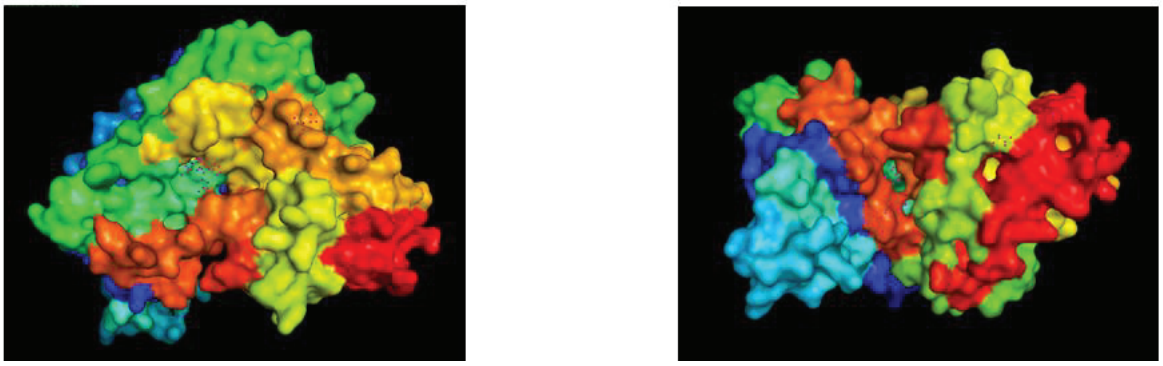


B)

MhWS2-WT



MhWS2-v11



All selected variants of MhWS2 possessed the point mutation A344T, and the variant WS2-v2 contained only this point mutation. This mutation alone increased 38% the specific enzyme activity when using ethanol as substrate and changed the substrate preference to medium-chain alcohols, while the catalytic activity when utilizing long-chain was pronouncedly affected. The other three WS mutants, WS2-v6, WS2-v7 and WS2-v11, presented a further increase in the catalytic activity using ethanol, while compared to the wild-type (Table 11). For example, in variant WS2-v11 65% increase in enzyme activity when utilizing ethanol was obtained. However, we also noticed that compared to WS2-v2, the improvement in ethanol selectivity for WS2-v6, WS2-v7 and WS2-v11 was not significant. Thus, we reached the premise that the point mutation A344T was the main and most important residue alteration in terms of changing the enzyme activity when employing ethanol. Therefore, we considered that the additional mutations (A35V, L128V, R132G, N165S, P352R and I462V) found in these variants were not as relevant as A344T.

The simulated structure models of MhWS2 and WS2-v11 were built employing

the Phyre2 server with the enzyme PapA5 as template (Kelley and Sternberg, 2009). PapA5 was used since its crystal structure had been successfully used to probe the active site of a WS (Barney *et al*., 2015; Barney *et al*., 2013). Similar to the two-domain structure of PapA5, the predicted structural model suggested that MhWS2 should have two domains connected by a helical linker (located between the residues at positions 170-230 of MhWS2). The existence of two domains had already been demonstrated in a WS from *M. aquaeolei* VT8 (Ma2) (Villa *et al*., 2014); besides that the protein sequence of Ma2 (WP_011786509) is nearly the same as the protein sequence of MhWS2 (ABO21021) except for position 395, which is glycine in MhWS2 but aspartate in Ma2. Thus, these findings increased the confidence in our model.

4.2 Directed evolution of Eeb1p (hydrolase)

4.2.1 Lipotoxicity

A previously constructed yeast strain JV03 free of storage lipids and unable of β-beta-oxidation (Valle-Rodríguez *et al*., 2014) and its parental strain CEN.PK113-5D were cultured in SD agar plates containing 20 mM oleic acid. Similar to previous reports (Petschnigg *et al*., 2009; Siloto *et al*., 2009), the storage-lipid free yeast JV03 was sensitive to oleic acid and not able to grow in the presence of 20 mM oleic acid due its lack of coping with long-chain fatty acids. On the other hand, the wild-type strain CEN.PK113-5D could grow at the same condition as the formation of storage lipids could detoxify the oleic acid (Petschnigg *et al*., 2009; Siloto *et al*., 2009).

4.2.2 Screening and selection

Error prone PCR was performed utilizing the GeneMorph® II Random Mutagenesis kit (Stratagene, La Jolla CA, USA), in order to create the library of Eeb1p variants *EEB1*^{L}, followed by transformation and expression in yeast JV03 utilizing the gap repair method. This method allows joining the mutated linear DNA fragments with the cut linear backbone plasmid pSS-E through an overlapping DNA section (*Not*I, *Sac*I), then the varied plasmids were recircularized. Yeast transformation was achieved by electroporation following a previously published protocol (Becker and Guarente, 1991). These variants possessed a higher activity which conferred their hosting JV03 cell the ability to survive and therefore, select the ones with the highest enzyme activity.

Solid agar plates consisting of SD-*ura* medium (6.9 g/l yeast nitrogen base without amino acids nor uracil) containing 100 mM of oleic acid (SD-*ura*+OA) were used for cultivating the mutant cells. A reduction (14%) on the transformation efficiency was obtained from growth on SD-*ura* plates with no oleic acid to growth on SD-*ura*+OA plates (Table 12). Then, the fatty acid concentration increased step-wise to 130, 160 and 200 mM as indicator in the evolution screening procedure (Table 13).

Table 12. Transformation efficiencies in average SD-*ura* plates containing mutant cells plated after gap repair and yeast transformation.

Randomly mutated gene	Oleic acid (mM)	Colonies (CFU) per plate
$EEB1^L$	0	2624
	100	356

Oleic acid concentrations shown are the concentrations in SD-*ura*+OA agar plates.

Table 13. Evolution screening of the libraries of variants by increasing concentration of the selective marker through its lipotoxic effect.

Randomly mutated gene	Oleic acid (mM)	Total no. of colonies (variants)
$EEB1^L$	100	152
	130	87
	160	49
	200	6

Oleic acid concentrations shown are the concentrations in SD-*ura*+OA agar plates.

A selection of the desired mutant strains (each one of them expressing a different hypothetical variant) was effectuated by choosing those mutants that could cope with the growth pressure through the screening procedure. In order to rate their performance, specific enzyme activity was evaluated utilizing as substrates palmitoleyl-CoA and ethanol (Table 14). Four strains that had over 20% increase in their specific enzyme activity when compared to the wild-type Eeb1p, were finally selected. These highest activities could be explained either

by a larger copy number of the plasmid or by an augment of the enzyme transcription level, or by a combination of both possible consequences. With the objective of answering this query, the recovered plasmids were transformed once again in clean cells of the host strain JV03. Each plasmid containing its corresponding theoretical variant of this hydrolase (Table 14). No significant statistical differences were found between activity values obtained utilizing the mutants directly chosen from the screening procedure and the retransformed strains (Table 13).

Table 14. Specific enzyme activities and increments (in relation to the wild-type enzyme) for the screened libraries derived from the lipotoxicity test.

Strain (variant)	**Specific enzyme activity (pmol/[mg cell extract ·min])**	**Increment (%) relative to wild-type**
Library of Eeb1p variants (hydrolase homolog to *S. cerevisiae*)		
Eeb1p-WT	1,057 ± 0,17	0
Eeb1p-v1	1,381 ± 0,58	30,7
Eeb1p-v2	1,289 ± 0,27	22,0
Eeb1p-v3	1,272 ± 0,47	20,4
Eeb1p-v4	1,530 ± 0,26	45,7
Eeb1p-v5	1,188 ± 0,32	12,4
Eeb1p-v6	1,468 ± 0,16	38,9

Abbreviations: WT, wild-type; v, variant

4.2.3 Sequence analysis of Eeb1p variants

Out of the whole library derived from mutagenized versions of EEB1, four variants were selected. These showed the highest enzyme activities. Their corresponding sequence was analyzed (Eurofins MWG Operon). Results of functional mutations are presented on Table 15 and Figure 19. There was no residue substitution in the catalytic triad of the active site (S251, D399 and H425) for any of the selected variants. This triad consists of three amino acids (Figure 19 and Figure 20) fully conserved among alpha/beta-hydrolases of the same family (4, PF00561) and it has to be totally preserved in Eeb1p for the catalytic mechanism to occur (Holmquist, 2000).

Table 15. Residue substitutions derived from point mutations in the selected variants obtained from the screening procedure.

E-v4	**E-v6**	**E-v1**	**E-v2**
R3P	N18S	**Q319K**	P364T
H14P	T30N	L432F	T373S
N18T	T69P		T397P
A34P	K84I		
R36K	Q87K		
F47L			
V48L			
L66F			
G101K			
A105S			
D106Y			
W107C			
S124T			
K135R			
H146Y			
S189Y			
G196S			
R218H			
Q319K			

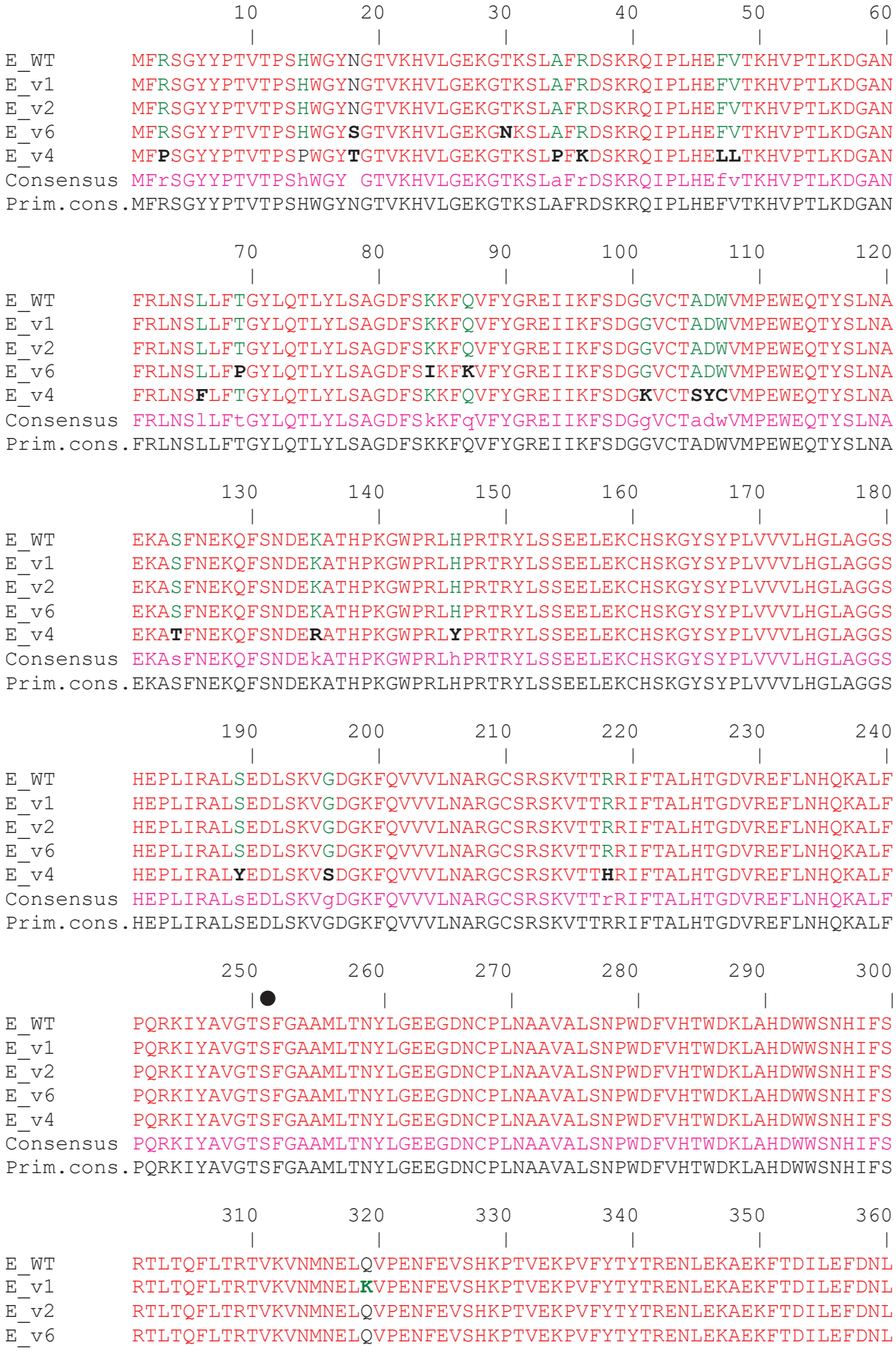

10 20 30 40 50 60
E_WT MFRSGYYPTVTPSHWGYNGTVKHVLGEKGTKSLAFRDSKRQIPLHEFVTKHVPTLKDGAN
E_v1 MFRSGYYPTVTPSHWGYNGTVKHVLGEKGTKSLAFRDSKRQIPLHEFVTKHVPTLKDGAN
E_v2 MFRSGYYPTVTPSHWGYNGTVKHVLGEKGTKSLAFRDSKRQIPLHEFVTKHVPTLKDGAN
E_v6 MFRSGYYPTVTPSHWGYSGTVKHVLGEKGNKSLAFRDSKRQIPLHEFVTKHVPTLKDGAN
E_v4 MFPSGYYPTVTPSPWGYTGTVKHVLGEKGTKSLPFKDSKRQIPLHELLTKHVPTLKDGAN
Consensus MFrSGYYPTVTPShWGY GTVKHVLGEKGTKSLaFrDSKRQIPLHEfvTKHVPTLKDGAN
Prim.cons.MFRSGYYPTVTPSHWGYNGTVKHVLGEKGTKSLAFRDSKRQIPLHEFVTKHVPTLKDGAN
70 80 90 100 110 120
E_WT FRLNSLLFTGYLQTLYLSAGDFSKKFQVFYGREIIKFSDGGVCTADWVMPEWEQTYSLNA
E_v1 FRLNSLLFTGYLQTLYLSAGDFSKKFQVFYGREIIKFSDGGVCTADWVMPEWEQTYSLNA
E_v2 FRLNSLLFTGYLQTLYLSAGDFSKKFQVFYGREIIKFSDGGVCTADWVMPEWEQTYSLNA
E_v6 FRLNSLLFPGYLQTLYLSAGDFSIKFKVFYGREIIKFSDGGVCTADWVMPEWEQTYSLNA
E_v4 FRLNSFLFTGYLQTLYLSAGDFSKKFQVFYGREIIKFSDGKVCTSYCVMPEWEQTYSLNA
Consensus FRLNSlLFtGYLQTLYLSAGDFSkKFqVFYGREIIKFSDGgVCTadwVMPEWEQTYSLNA
Prim.cons.FRLNSLLFTGYLQTLYLSAGDFSKKFQVFYGREIIKFSDGGVCTADWVMPEWEQTYSLNA
130 140 150 160 170 180
E_WT EKASFNEKQFSNDEKATHPKGWPRLHPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
E_v1 EKASFNEKQFSNDEKATHPKGWPRLHPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
E_v2 EKASFNEKQFSNDEKATHPKGWPRLHPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
E_v6 EKASFNEKQFSNDEKATHPKGWPRLHPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
E_v4 EKATFNEKQFSNDERATHPKGWPRLYPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
Consensus EKAsFNEKQFSNDEkATHPKGWPRLhPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
Prim.cons.EKASFNEKQFSNDEKATHPKGWPRLHPRTRYLSSEELEKCHSKGYSYPLVVVLHGLAGGS
190 200 210 220 230 240
E_WT HEPLIRALSEDLSKVGDGKFQVVVLNARGCSRSKVTTRRIFTALHTGDVREFLNHQKALF
E_v1 HEPLIRALSEDLSKVGDGKFQVVVLNARGCSRSKVTTRRIFTALHTGDVREFLNHQKALF
E_v2 HEPLIRALSEDLSKVGDGKFQVVVLNARGCSRSKVTTRRIFTALHTGDVREFLNHQKALF
E_v6 HEPLIRALSEDLSKVGDGKFQVVVLNARGCSRSKVTTRRIFTALHTGDVREFLNHQKALF
E_v4 HEPLIRALYEDLSKVSDGKFQVVVLNARGCSRSKVTTHRIFTALHTGDVREFLNHQKALF
Consensus HEPLIRALsEDLSKVgDGKFQVVVLNARGCSRSKVTTrRIFTALHTGDVREFLNHQKALF
Prim.cons.HEPLIRALSEDLSKVGDGKFQVVVLNARGCSRSKVTTRRIFTALHTGDVREFLNHQKALF
250 260 270 280 290 300
E_WT PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
E_v1 PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
E_v2 PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
E_v6 PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
E_v4 PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
Consensus PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
Prim.cons.PQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSNHIFS
310 320 330 340 350 360
E_WT RTLTQFLTRTVKVNMNELQVPENFEVSHKPTVEKPVFYTYTRENLEKAEKFTDILEFDNL
E_v1 RTLTQFLTRTVKVNMNELKVPENFEVSHKPTVEKPVFYTYTRENLEKAEKFTDILEFDNL
E_v2 RTLTQFLTRTVKVNMNELQVPENFEVSHKPTVEKPVFYTYTRENLEKAEKFTDILEFDNL
E_v6 RTLTQFLTRTVKVNMNELQVPENFEVSHKPTVEKPVFYTYTRENLEKAEKFTDILEFDNL

Figure 19. Multiple sequence alignment of wild-type hydrolase Eeb1p and the selected variants. Conserved residues (Ser-251, Asp-399 and His-428) that are essential for catalytic activity are marked with a big dot. Residue substitutions in bold font and enclosed in a box. Alignments generated with MultAlin (Corpet 1988).

There were 19 residue substitutions in the variant (E-v4) with highest enzyme activity (46% increase regarding wild-type), while the second variant (E-v6) with 39% increase presented five. The third variant (E-v1, 31%) had two and the fourth (E-v2, 21%) presented three residue substitutions.

In variants E-v4 and E-v1, one amino acid substitution (Q319K) was found in common. Aimed to this, the residue substitution at position N18 was also found in E-v4 and E-v6, but in this occasion to amino acids of polar chemical nature and containing a hydroxyl functional group: threonine (N18T) and serine (N18S), respectively.

4.2.4 Novel structure-function features of variants

The residue substitution T397P (in variant E-v2) is positioned two residues distant from aspartate catalytical residue D399, which plays a functional role in the biocatalysis of formation of esters, T397P most probably causes a subtle change at this protein region very close to the critical catalytical site favouring the bioreaction mechanism thermodynamically and/or kinetically for easing

and/or speeding the esterification process. The exposed nitrogen atom of the proline functional group provides a partial positive charge for which it can perform as an acid and coordinate a reactant molecule by orienting, in this case one of the polar atoms of the acyl-coA molecule or the oxygen atom of the alcohol molecule.

The residue substitution L432F found in E-v1 is present at very high proximity to the catalytical histidine (H428) at just 4 residues distant from it. The chemical nature of the amino acid did not change remaining as hydrophobic, but the single branched-chain of leucine was altered for the phenyl group contained in phenylalanine. This substitution must have a very low probability on catalytic performance matters, but could still bring a steric or conformational change in the structure very close to the active site.

No leader sequence was found at the first section (residues 1-167) of wild-type Eeb1p, but a signal peptide has been theoretically assumed at the N-terminal from residue M1 until a cleavage site at N18. In variant E-v4, three residue substitutions (R3P, H14P, N18T) are present at this initial sequence. This decisive modification in primary structure strongly suggests that this variant does not have a signal peptide, since the positively charged start was suppressed by the residue substitution R3P. This represents that E-v4 does not get localized to any cellular component or organelle, but rather gets freely dispersed in cytoplasm reacting and forming FAEEs.

This initial sequence with a defined secondary structure in E-v4, got adapted to the rest of the protein (see Figure 20), while a target peptide would be cleaved and these residues would be lost (as in E-WT). This finding most probably means that in order to be more catalytically active at the conditions of the screening procedure, this enzyme would require a larger structure that would provide a higher stability and robustness to perform as a biocatalyst under the stress of predominant presence of oleyl-CoA and ethanol.

E-WT E-v4

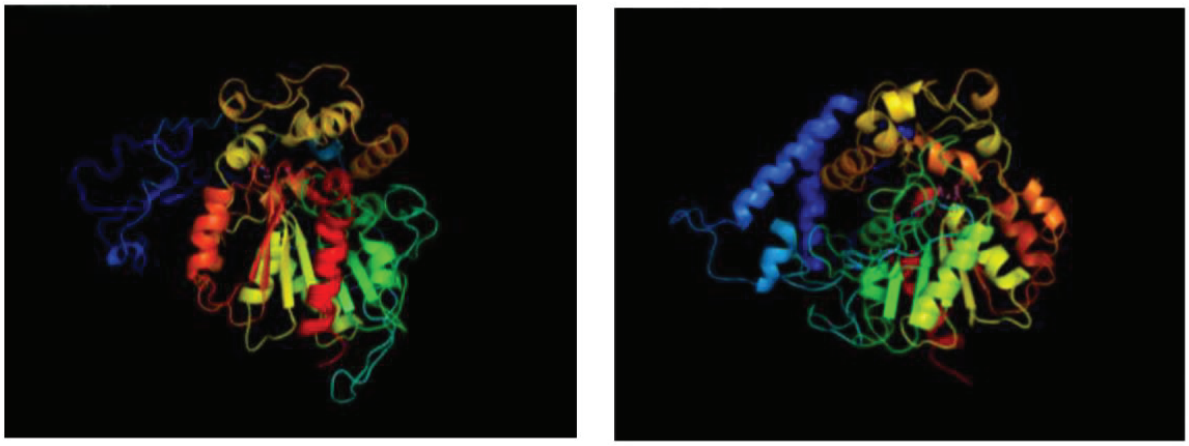


Figure 20. Prediction of tertiary structure of the Eeb1p wild-type and selected variant (E-v4). Sequences are presented in rainbow spectrum from protein head by N-terminus (dark blue) to protein tail by C-terminus (dark red). The active site is marked with gray spots of the residues of the catalytic triad (S251, D399 and H428). Reference: Kelley and Sternberg, 2009.

At this first section, the EK motif was found as an amino acid repeat 5 times, while the KHV motif repeated itself only once for the native enzyme. In E-v6, no residue substitutions were found at the initial peptidic sequence corresponding to the signal peptide, suggesting that this variant still gets localized in yeast cell; despite that all residue substitutions in this variant were present at this first section of the protein. The chemical nature (affinity of functional group) of the first two residues remained unmodified (N18S, T30N). However, N18S propitiated a more defined secondary structure at the sequence it directly precedes, where homology modeling forecasts an extended beta-strand of 7 residues (from G19 to L25) compared to a 6 residue shorter beta-strand in wild-type Eeb1p with probability of a disordered region at the end of it. This feature sets E-v6 with an evolved structure at this first section of the protein.

The later residue substitution T69P (polar to non-polar) preceded the formation of a more defined structure with a longer alpha-helix of 9 residues (Y71 to A79) when compared to 8-residue helix (L72 to A79) in wild-type, while substitutions K84I and Q87K forecast a coil or turn (as in the wild-type) but there is a higher probability of a small disordered region (D81 to K85) in wild-type Eeb1p.

Part of this first section of the protein could represent a transmembrane sequence where hydrophobic residues usually prevail, as it has been found in Dga1p, Yim2p, as well as in Are1p and its paralog Are2p. Variants E-v1 and E-v2 had no residue substitution at the first section of their protein sequence,

remaining with a theoretical signal peptide as the wild-type Eeb1p.

Directed evolution behind E-v4 brought the loss of the signal peptide for this variant, while it is not diagnosed. Therefore, this suggests that this variant does not localize in any cellular component, organelle nor lipid particle, while most probably being dispersed in the cytoplasm directly reacting with acyl-CoA and ethanol forming ethyl esters. *In vitro* experiments have clarified that the wild-type Eeb1p localizes at lipidic membranes (Knight *et al.*, 2014). Moreover, it has been found that its paralog Eht1p locates in mitochondria and lipid particles (Saerens *et al.*, 2000). Eeb1p most probably arose from Eht1p due to the whole genome duplication. Both signal peptides of Eeb1p and Eht1p highly conserve the chemical nature of their residues (67%), where 8 residues are totally identical in both signal sequences. Given the fore two premises, it can be assumed that wild-type Eeb1p must localize at mitochondria and lipid particles also. Other proteins have been proven to be localized at lipid particles, for instance: Tgl3p lipase localizes at TAG droplets (Athenstaedt and Daum, 2003), Yju3p at lipid particles and cellular membranes (Heier *et al.*, 2010; Zahedi *et al.*, 2006), Tgl1p hydrolase at lipid particle membranes (Köffel *et al.*, 2005); while Mgl2p is localized at the plasma membrane and dispersed in cytosol (Selvaraju *et al.*, 2016), Rog1p at nucleus and cytosol too (Varthini and Selvaraju, 2015). Erg1p is localized at the endoplasmic reticulum (Leber *et al.*, 1998).

In E-v4, most of the residue substitutions occurred at the first section of the protein. The chemical nature of the residues was not altered from R36K to L66F, but the nature pronouncedly changed at later residue susbstitutions. Three substitutions occurred in series with a predominant modification in their chemical nature: A105S, D106Y and W107C. Moreover, the cap-subdomain extends and starts at residue E127; here the nature remained as positively charged at K135R but severely altered at H146Y and S189Y. In the wild-type E-WT, the cap-subdomain is present at the ABHD domain. Hence, it can be assumed that most of these residue substitutions in E-v4 conferred this variant with the highly evolved feature of allowing this primary structure to get integrated into the cap-subdomain that covers the protein.

Variants E-v1 and E-v2 have no residue substitutions at the signal peptide section, therefore these variants should still get localized. Their residue substitutions occurred at positions very close to the C-terminal of their ABHD domain; while for E-v1 only Q319K happened at the cap subdomain. Variant E-v2 features a more evolved structure of a longer alpha-helix (G371 to K377) than wild-type Eeb1p, with an alpha-helix of 6 residues (G371 to R376).

A cap-subdomain must be present, with the main function of covering this protein and protecting it from the surroundings. This should contain its own structural set (or set of structures). There was a definitive modification on the presence of this subdomain from the wild-type (residues Y168 to S327) to selected variant E-v4 (residues E127 to H284). I assume that this cap subdomain contains a NC-loop, a cap-loop and a back-loop, as proven for other alpha/beta-hydrolases of the same superfamily where Eeb1p belongs (Mitusinska *et al.*, 2022; Bauer *et al.*, 2019; Barth *et al.*, 2004).

Based on alignments (Figure 21), the NC-loop of wild-type Eeb1p and variants should consist of 23 residues, which is in accordance with previous reports of 16-57 residues (Barth *et al.*, 2004). This structure performs as a hinge where it changes the acute angle between its connecting amino acid chains that could implicate a spatial modification on the protein. It has also been found in other hydrolases that this structure contributes in substrate binding by defining the binding pocket and therefore regulating the income of reactants to the active site (Kaneko *et al* 2010). It is predicted to be composed of a beta-sheet and an alpha-helix. In the selected variant E-v4, two residue substitutions were found at this structure (K135R and H146Y).

A cap-loop of 41 residues is predicted in E-WT and variants E-v6 and E-v1, while 37 residues are assumed at this subdomain structure in E-v4 and E-v2. This cap-loop is composed of 5 to 59 residues in similar proteins. An alpha-helix is theoretically thought at this subdomain part. This must be the most flexible region of this protein and related variants (Mitusinska *et al.*, 2022; Bzowka *et al.*, 2022). In E-v4, three residue substitutions should be present at its cap-loop (S189Y, G196S, R218H). This supports the theory of this section playing an important role of flexibility, since these residue substitutions contributed on positively altering the catalytic performance of this selected variant. Moreover, cap-loops of eukaryotes have been found to be longer and participate in fatty acid metabolism (Zeldin *et al.*, 1995, Fretland and Omiecinski, 2000), therefore involved in the alterations prognosticated for this variant.

A back-loop must be present where 29 residues are predicted there for the wild-type and all variants. This set of structures also acts as a hinge provoking a spatial modification at its proximity. It is formed of 15 to 33 residues in other hydrolases (Mitusinska *et al.*, 2022). An alpha-helix followed by a beta-sheet are hypothesized at this back-loop. Fungal and mammalian hydrolases possess longer back-loops than those of plants and bacteria (Mitusinska *et al.*, 2022). This subdomain part was totally conserved for all variants where no residue

substitution were found, excepting Q319K in E-v1.

```
                    10        20        30        40        50        60
                     |         |         |         |         |         |
E_v4xx0      MFPSGYYPTVTPSPWGYT----GTVKHVLGEKGTKSLPFKDSKRQIPLHELLTKHVPTLK
MsEHxx1      ----------DVSHGYVTVKPGIRLHFVEMGSGP---ALCLCHGFPESWFSWRYQIPAL
StEHxx3      ---------MEKIEHKMVAVN-GLNMHLAELGEGP---TILFIHGFPELWYSWRHQMVYL
TrEHxx2      MDTSKLKPNDPRVKYETKQIRGKTYSYILGEPQGPKLETVVLVHGWPDMAFGWRHQIPYL
BmEHxx4      -----------MSKQYINVN-GVNLHYISKGQG---ELMLFLHGFPDFSHIWRHQIDEF
Consensus    m  s   p        y     gt  h lge  g  l      hg P     wrHqi
Prim.cons.   M22S222P2334VSHGYV4VN2GTNLH5L2EGQGP2LET5L25HGFP25WFSWRHQIPYL

                    70        80        90       100       110       120
                     |         |         |         |         |         |
E_v4xx0      DGANFRLNSFLFTGYLQTLYLSAGDFSKKFQVFYGREIIKFSDGKVCTSYCVMPEWEQTY
MsEHxx1      AQAGFRVLAIDMKGYGDSSSPP--EIEEYAMELLCKEMVTFLDKLGIPQ--AVFIGHDWA
StEHxx3      AERGYRAVAPDLRGYGDTTGAPLNDPSKFSILHLVGDVVALLEAIAPNEEKVFVVAHDWG
TrEHxx2      MSLGFQVVAPNMLGYAGTDAPR--DLSQFTLKSVSADIAELARSFVGQDGQIVLGGHDWG
BmEHxx4      SN-DFHTVALDLRGYNLSEKPS--GLESYEIDVLVEDIRQVIEGLGYSS--CTLVVHDWG
Consensus        Fr va    GY  t  ps  d s         #!     g v  s       h#wg
Prim.cons.   A5AGFRVVAPD2RGYGDT55P222DLSK25I55LV5DIV52L2GL255S33VVLVGHDWG

                   130       140       150       160       170       180
                     |         |         |         |         |         |
E_v4xx0      SLNAEKATFNEKQFSNDERATHPKGWPRLYPRTRYLSSEELEKCHSKGYSYPLVVVLHGL
MsEHxx1      GVMVWNMALFYPERVRAVASLNTPFMPPDPDVSPMKVIRSI-PVFNYQLYFQEPGVAEAE
StEHxx3      ALIAWHLCLFRPDKVKALVNLSVHFSKRNPKMNVVEGLKAIYGEDHYISRFQVPGEIEAE
TrEHxx2      GAVVWRTAYYHPELVKAVFSVCTPLHPLSAEYKPLEDIVAAGHMLNFKYQLQLKGP-DVE
BmEHxx4      AGIGWTFAYRYPEYVQKLIAFNGP-----------------HPYTFMR----------E
Consensus        w  a   p# v    a   p  p                    y   l      e
Prim.cons.   2LI2W55A2FYPE5VKA252LNTPF4PR4P444P4E4I4AI3H55N25Y4FQLPGV3EAE

                   190       200       210       220       230       240
                     |         |         |         |         |         |
E_v4xx0      AGGSHEPLIRALYEDLSKVSDGKFQVVVLNARGCSRSKVTTHRIFTALHTGDVREFLNHQ
```

```
MsEHxx1      LEKNMSRTFKSFFRA-SDETGFIAVHKATEIG ILVNTPEDPNLSKITTEEEIEFYIQQF
StEHxx3      FAPIGAKSVLKKILT-YRDPAPFYFPKGKGLEAI---PDAPVALSSW SEEELDYYANKF
TrEHxx2      ARIQGKDMLRRFFRAMFGGRGPNGEAGFSTSD VHFDVLDKIGAPPL DEQELEYYVEQY
BmEHxx4      LRTNKNQQKASEYMKWFQKQEVQDYMERDNFSGLRKLVIDPGVKKGY TADDVQAYMNSW
Consensus    a        r  %               n  G             L   #v  % #
Prim.cons.   2R5NG5555RSF2RA32555GP5555K55N55GI444V5DP55LS55LTEEE2EYY5NQF

                      250       260       270       280       290       300
                       |         |         |         |         |         |
E_v4xx0      KALFPQRKIYAVGTSFGAAMLTNYLGEEGDNCPLNAAVALSNPWDFVHTWDKLAHDWWSN
MsEHxx1      ---KKTGFRGPLNWYRNTERNWK----WSCKGLGRK--ILVPALMVTAEKDIVLRPEMSK
StEHxx3      ---EQTGFTGAVNYYRALPINWE----LTAPWTGAQ--VKVPTKFIVGEFDLVYHIPGAK
TrEHxx2      ALQEAPELRGPLNWYRTRELNAKDEMDRAKNGPPLR--FEMPALFVAASKDNALPPAMSK
BmEHxx4      ---ENGSVLSMLSYYRNLKIFTEEDLRRKSLFPLEEEVLNIPVQIIWGNQDPTFMPENLD
Consensus       e       l  yr     t        n pl   v   p        D    p  s
Prim.cons.   222E5TGFRG2LN2YRNLEIN223333R55NGP2552V5LVPALF2V2EKD5VLHPEMSK

                      310       320       330       340       350       360
                       |         |         |         |         |         |
E_v4xx0      HIFSRTLTQFLTRTVKVNMNELKVPENFEVSHKPTVEKPVFYTYTRENLEKAEKFTDILE
MsEHxx1      N------MEKWIPFLKRGHI-EDC-GHWTQIEKPTEVNQILIKWLQTE------------
StEHxx3      EYIHNGGFKKDVPLLEEVVVLEGA-AHFVSQERPHEISKHIYDFIQKF------------
TrEHxx2      ------GMDAFYKDLTRAEV--DA-THWALTQAGDEVNRVIGEWLNKALNGATKAAL---
BmEHxx4      ------GIEEYVPNISVHRL-AEA-SHAPQHEKPQEVNNVMWNFLNK-------------
Consensus          g #       v      a  h     kp evn !    l k l  a k
Prim.cons.   322222GMEK5VP5LK255V2EDAP5H25Q5EKPTEVN5VIY52L2K4L22A2K222ILE

                      370       380       390       400       410       420
                       |         |         |         |         |         |
E_v4xx0      FDNLFTAPSMGLPDGLTYYRKASSINRLPNIKIPTLIINATDDPVTGENVIPYKQARENP
MsEHxx1      ------------------------------------------------------------
StEHxx3      ------------------------------------------------------------
TrEHxx2      ------------------------------------------------------------
BmEHxx4      ------------------------------------------------------------
```

```
Consensus
Prim.cons.    FDNLFTAPSMGLPDGLTYYRKASSINRLPNIKIPTLIINATDDPVTGENVIPYKQARENP

                     430       440       450       460
                       |         |         |         |
E_v4xx0       CVLLCETDLGGHLAYLDNESNSWLTKQAAEFLGSFDELVL
MsEHxx1       ----------------------------------------
StEHxx3       ----------------------------------------
TrEHxx2       ----------------------------------------
BmEHxx4       ----------------------------------------
Consensus
Prim.cons.    CVLLCETDLGGHLAYLDNESNSWLTKQAAEFLGSFDELVL
```

Figure 21. Multiple sequence alignment of Eeb1p-v4 (selected hydrolase variant) and reference hydrolases. Residue substitutions highlighted in yellow. Cap-subdomain structures: NC-loop in purple, cap-loop in red and back-loop in blue. E_v4: Eeb1p-v4, MsEH: *Mus musculus* soluble epoxide hydrolase 1cqz, StEH: *Solanum tuberosum* soluble epoxide hydrolase 2cjp, TrEH: *Trichoderma reesei* soluble epoxide hydrolase 5uro, BmEH: *Bacillus megaterium* soluble epoxide hydrolase 4nzz. Alignments generated with MultAlin (Corpet 1988).

It has been shown for other hydrolases that the main ABHD domain regions and regions between the fore presented set of structures of their cap subdomain displayed a high level of structural similarity, whereas the NC-loop, cap-loop, and back-loop regions displayed dissimilarity (Mitusinska *et al.*, 2022). This is in accordance with the theory that these set of structures are good targets for creating changes in the functionality, while being more susceptible for modifications, as found in variants E-v4 and E-v1.

Only four residue substitutions were found at the homology identified ABHD domain of E-v4 (residues 168-434, Ollis *et al*., 1992). Meanwhile, all residue substitutions in E-v6 were present at the first section of this variant, while no residue substitutions were found at its ABHD domain (see Table 15 and Figure 19). This confirms the assumption that this enzyme needs to preserve the folding pattern of its related alpha/beta-hydrolase domain since residues were more conserved there and if any substitutions occurred, the chemical nature of the residues was overall preserved, excepting Q319K (see Figure 19).

All residue substitutions in E-v6 happened at the first section of the protein. This

section must act as a lid covering the ABHD domain and forming channels for accessing the active site region from the surroundings of the enzyme (Bzowka *et al.*, 202; Bauer *et al.*, 2019). α/β hydrolase folds of this superfamily domain (residues 168-434) that sustain the active site (catalytic triad at residues S251, D399 and H428; Rauwerdink and Kazlauskas, 2015) are predominantly conserved between wild-type and variants while no severe mutations occurred in this domain (excepting Q319K), which suggests that it has to be totally preserved in order to maintain the functionality of the enzyme. It has been shown that structural elements in this domain play important roles for providing orientation zones, e.g. oxyanion by N-H or O-H groups (Yang and Wong, 2013), acid and nucleophile zones; of substrates and products involved during the biocatalysis (Denesyuk *et al.*, 2020).

4.2.5 Analysis of chemical nature in residue substitutions

The chemical nature of the amino acids involved in the substitutions was mostly preserved. In variant E-v4, it changed from positively charged to non-polar near the N-terminus (R3P and H14P), while the chemical nature was absolutely conserved from residues N18T to L66F. But definitive modifications happened at approximately the middle of the first section of this protein variant (G101K, A105S and D106Y), while the chemical nature of those residues closer to the ABHD domain was preserved (W107C, S124T and K135R).

In E-v6, the chemical nature of the residues as polar was also conserved close to the N-terminus (N18S and T30N), but the chemical nature absolutely changed from polar to non-polar at T69P, while it was also modified at residues K84I (from positively charged to hydrophobic) and Q87K (from polar to positively charged).

Two residue substitutions were found in E-v1: Q319K and L432F, in which the chemical nature remained as hydrophobic (neutral). While in E-v2, the residue nature was altered between non-polar and polar at substitutions P364T and T397P, but the chemical nature of T373S remained as polar. All these were present at the ABHD domain.

By applying directed evolution, a large number of residue substitutions were effectuated in Eeb1p, where pronounced improvements on structure-function were achieved: 1) the signal peptide sequence was severely altered and therefore integrated as structure; 2) the initial section of this protein (theoretically assumed as a lid) was highly susceptible for residue substitutions; 3) a number of residue alterations were found at its homology identified cap

subdomain with main implications on covering properties at the surface of this biocatalyst variants; and 4) two residue substitutions (T397P and L432F) happened at high proximity to the active site.

It can be summarized that the main target of Eeb1p as a biocatalyzer by directly evolving towards the biocatalysis of acyl-CoA and ethanol for producing FAEEs, is achieved by a) extending the protein size to get structurally integrated, b) improving its structural features with more defined secondary structures and c) altering its residues at the covering cap subdomain and the lid formed by its initial section.

4.2.6 Alcohol selectivity

A different biocatalytic performance for the selected Eeb1p variants was found when assessing enzyme specific activity in order to determine its alcohol selectivity as substrate, while still employing palmitoyl-CoA for achieving FAEE formation (Figure 22). These variants were more selective towards primary alcohols of very short-chain length, as their selectivity while utilizing ethanol was further pronounced. Compared to the wild-type, E-v4 presented the most predominant effect where its catalysis of formation of palmitoyl ethyl ester through esterifying palmitoyl-CoA with ethanol increased 46%, despite that its catalysis of 1-butanol was not modified. The specific activity when reacting with 1-hexanol was reduced to 43%, while in the case of 1-octanol reduced to 38%, the biocatalysis was also diminished when utilizing 1-decanol (48%) and 1-dodecanol (68%) but when utilizing longer-chain alcohols (C14 to C18) its enzyme specific activity remained unmodified and very low. All other variants (E-v1, E-v2 and E-v6) altered their alcohol preference toward ethanol, they presented no statistically significant modification by employing 1-butanol and a decrease on their activity when utilizing 1-hexanol and longer-chain alcohols. No statistical differences were found among the enzyme specific activities of these variants by utilizing ethanol, but still a consequent reduction on mean values (from variants) when utilizing longer-chain alcohols was present.

These results lead to assess that the selected variants possess an alteration in their alcohol preference, which is successful due to their higher selectivity toward ethanol, one of the substrates requested to produce microbial biodiesel (FAEEs).

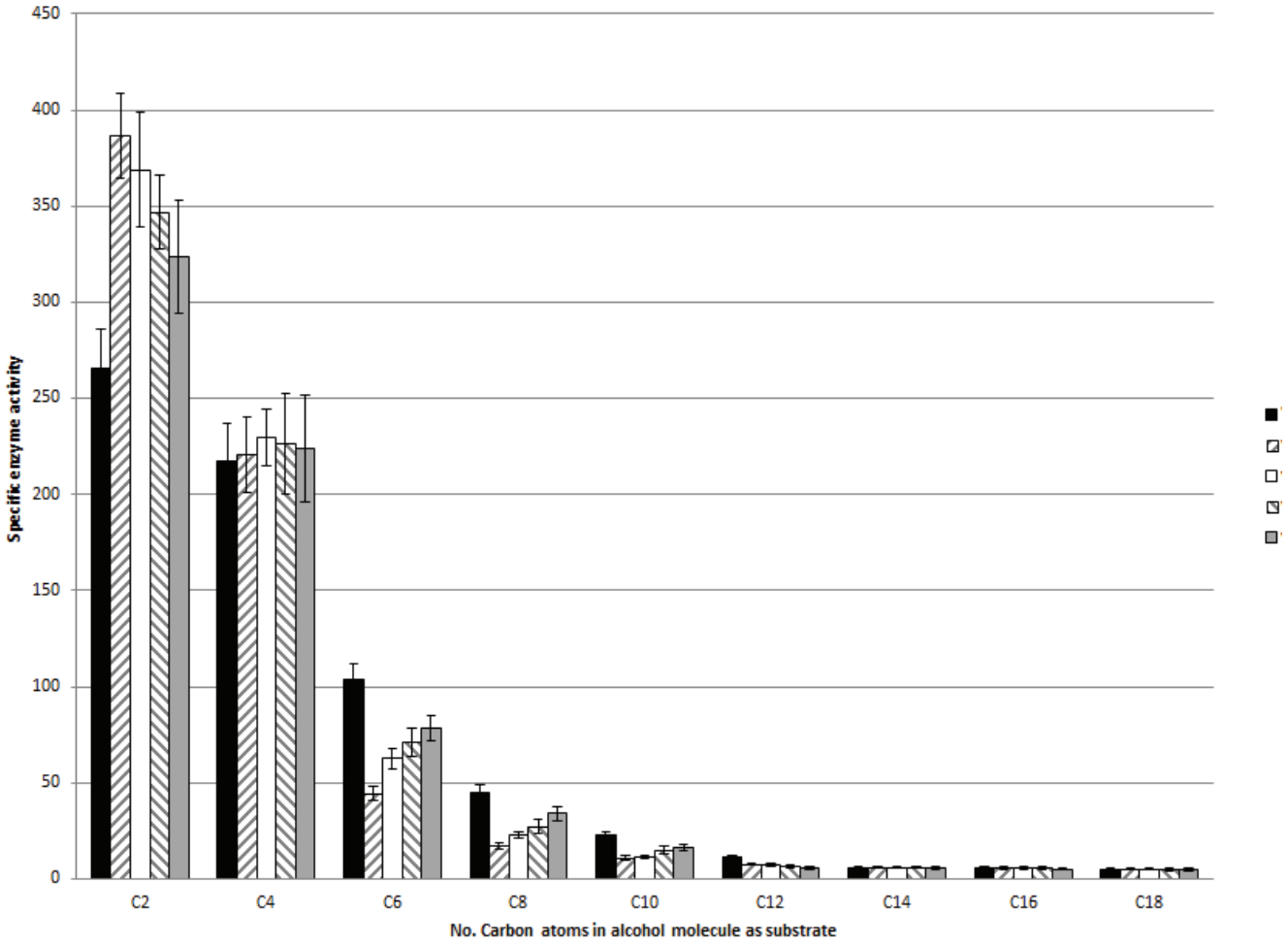


Figure 22. Selectivity of Eeb1p variants towards alcohol as substrate. Specific enzyme activities with various alcohols as substrate, abbreviations are C2: ethanol, C4: 1-butanol, C6: 1-hexanol, C8: 1-octanol, C10: 1-decanol, C12: 1-dodecanol, C14: 1-butadecanol, C16: 1-hexadecanol and C18: 1-octadecanol.

4.2.7 Acyl-CoA selectivity

The enzyme performance was also altered while testing preference towards the activated fatty acid (FA-CoA), as second substrate and using ethanol as first substrate for the FAEE formation (Figure 23). This can be stated due the results obtained from enzyme specific activity tests. The wild-type Eeb1p possessed a round preference for n-octanoyl-CoA as activated fatty acid, where a specific activity of 46.1 nmoles/(mg cell extract • min) was found. By utilizing shorter or longer activated fatty acids, specific activity is reduced to less than one fifth of the activity with n-octanoyl-CoA. When wild-type Eeb1p is exposed to n-hexanoyl-CoA, its activity diminishes 16% that one with n-octanoyl-CoA, while with butyril-CoA it is lowered to 6%. In the presence of n-decanoyl-CoA, hydrolase activity drops 14% of the wild-type activity, while with lauroyl-CoA it is reduced 4% and with higher activated fatty acids, it is lower than 1%.

The enzyme variant that had the most pronounced modification on its activated fatty acid specificity was E-v4, which is the variant that also showed the highest catalytic activity with ethanol and palmitoyl-CoA. This variant preferred utilizing lauroyl-CoA with a specific activity of 6.9 nmoles/(mg cell extract • min), with shorter and longer activated fatty acids its activity was much lower. The variant E-v6 showed a clear preference for n-decanoyl-CoA with a specific activity of 21.3 nmoles/(mg cell extract • min). Both variants E-v1 and E-v2 presented the highest activity with n-octanoyl-CoA, 28.6 and 35.5 nmoles/(mg cell extract • min), respectively.

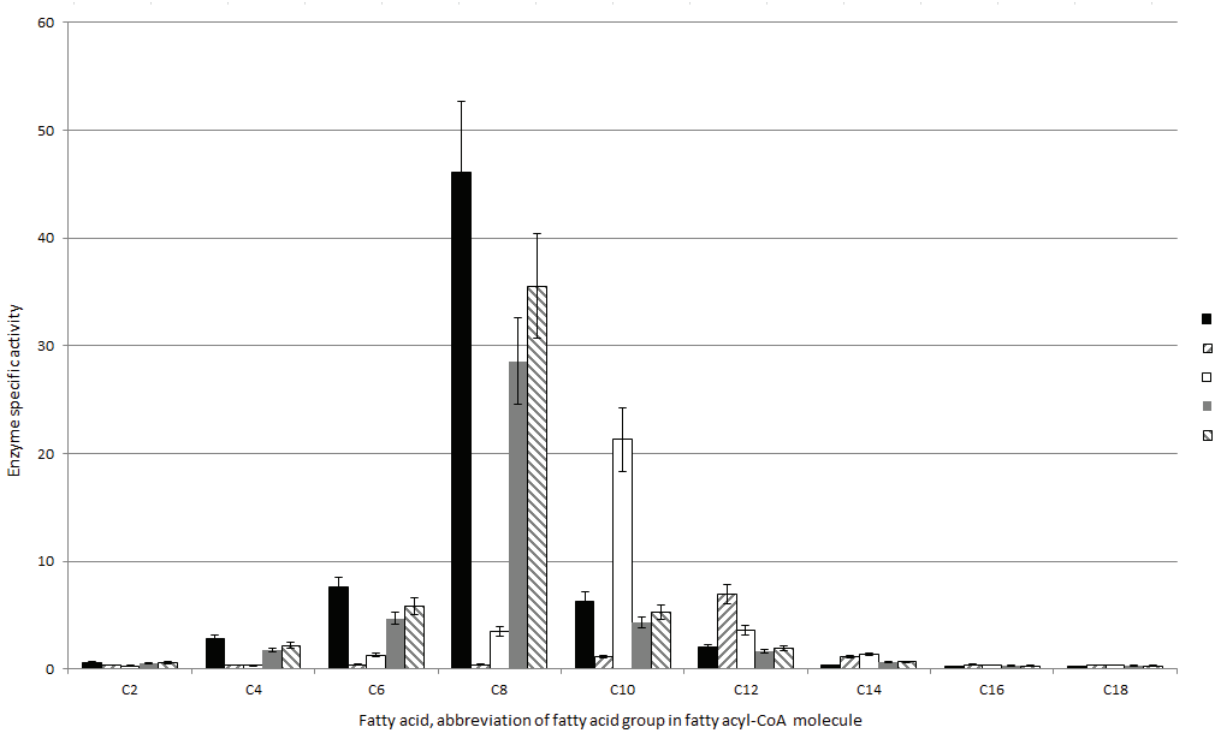


Figure 23. Selectivity of Eeb1p variants towards fatty acyl-CoA as substrate. Specific enzyme activities in nmol/(mg cell extract • min) with various fatty acids as substrate, abbreviations are C2: acetyl-CoA, C4: butyril-CoA, C6: n-hexanoyl-CoA, C8: n-octanoyl-CoA, C10: n-decanoyl-CoA, C12: lauroyl-CoA, C14: myristoyl-CoA, C16: palmitoyl-CoA and C18: oleyl-CoA.

4.3 Rational engineering of MhWS2

4.3.1 Rational mutagenesis by point mutations

In order to modify the alcohol selectivity of MhWS2, rational mutagenesis was performed by substituting a single residue that must be directly involved in alcohol selectivity. Site-directed mutagenesis of the involved nucleotide codon was achieved through performing a whole plasmid amplification by utilizing Phusion™ Site-Directed Mutagenesis Kit (Finnzymes, Finland). The employed plasmid was pSS-ST2-WS2 with a pair of PAGE purified (previously) 5´ phosphorylated oligonucleotide primers, for each point mutation, respectively (Table 16). PCRs were effectuated with one forward primer containing the changed codon for carrying on the point mutation for the desired residue substitution, paired to one constant reverse primer. Point mutated plasmids were transformed by the chemical method in competent *E. coli* DH5α cells. Subsequently, they were produced, recovered and sequenced (Eurofins MWG Operon, Ebersberg, Germany) for confirming. This substitution prediction was based on molecular modeling by similarity with a homolog enzyme (paPA5) aligned with a previously elucidated docking mechanism (pdb file 2H3P). By alignment of both protein sequences (Corpet, 1988) the predicted residue was Tyr-367 (Y367), which was substituted to other residues of the same chemical nature that would still confer hydrophobicity to this section of the enzyme but modify its properties in the sense of preference toward alcohols of different chain lengths. The tested residues were tryptophan (Y367W) with the expectation of conferring the enzyme with an alcohol preference towards shorter-chain alcohols, and as controls isoleucine (Y367I) and alanine (Y367A).

Table 16. List of oligonucleotide primers used in this study for plasmid construction.

Primer name	*Sense*	*Sequence 5′→3′*
WS-Y367W-fw	Forward	GCTGATGTCACCTTGGATCTTACAATTGA
WS-Y367I-fw	Forward	GCTGATGTCACCTATTATCTTACAATTGA
WS-Y367A-fw	Forward	GCTGATGTCACCTGCTATCTTACAATTGA
WS-Y367X-rv	Reverse	AGCATTGTGTATTGTGTTAGTGCTGAC

In wild-type MhWS2, this tyrosine residue Y367 is positioned towards influencing the active site (^{140}HHXXXDG^{146}), since the functional group of this

amino acid is oriented with its cyclic structure plane and its hydroxyl both directed toward the active site (see Figure 24), it should possess a direct influence in the interaction and fitting of the alcohol molecule introduced and bound in the enzyme for the active site to catalyze the transfer reaction occurring with a fatty acyl-CoA yielding a FAEE molecule (e.g. microbial biodiesel).

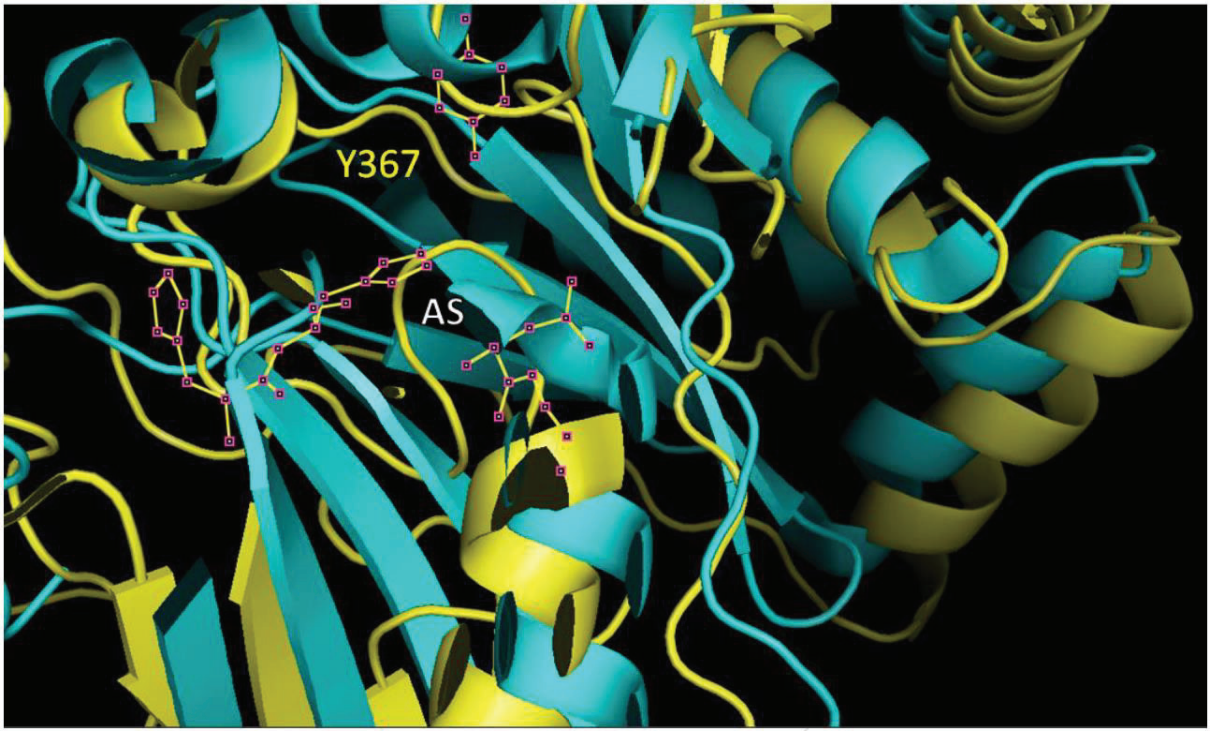


Figure 24. Image of the active site (AS) and the selected residue Y367 in MhWS2. Alignment of MhWS2 (yellow) with paPA5 (pdb file 1Q9J, cyan blue).

4.3.2 Alcohol selectivity

Specific enzyme activities were obtained with the purpose of investigating any effect of the single residue substitutions on the respective enzyme variant performance (see Figure 25). Specific enzyme activities with 1-dodecanol, 1-tetradecanol and 1-hexadecanol (C12, C14 and C16 alcohols) were always lower when compared to wild-type MhWS2. The residue modification to tryptophan (Y367W, tyrosine to tryptophan), had a positive effect on altering the substrate selectivity towards alcohols of shorter-chain length. Its activity by utilizing 1-butanol was twice the one of the wild-type. Apparently its activity with ethanol increases, although no statistical differences were found when compared to the wild-type. Opposite to these results, the specific activity of MhWS2-Y367W when exposed to long-chain alcohols was reduced. In the case of 1-hexadecanol utilization the activity decreased 12%, while by utilizing 1-octadecanol it was half (47%) of the wild-type activity. In the enzyme variants

utilized as controls, the contrary effect was observed. Both mutated versions, preferred long-chain alcohols while their specific activity by utilizing short-chain alcohols decreased and with medium-chain alcohols it was dramatically reduced. An apparent reduction in specific activity by consuming each alcohol by the alanine substitution variant (Y367A) as compared to the isoleucine substitution variant (Y367I) could be stated, although no statistical differences were found. The variant MhWS2-Y367A decreased its activity with 1-hexanol by 60% if compared to wild-type, while reductions were 43% and 32% when 1-butanol and 1-octanol were utilized, respectively. In contrary, an apparent increase in specific biocatalysis when employing 1-octadecanol could be mentioned, although no statistical difference was estimated.

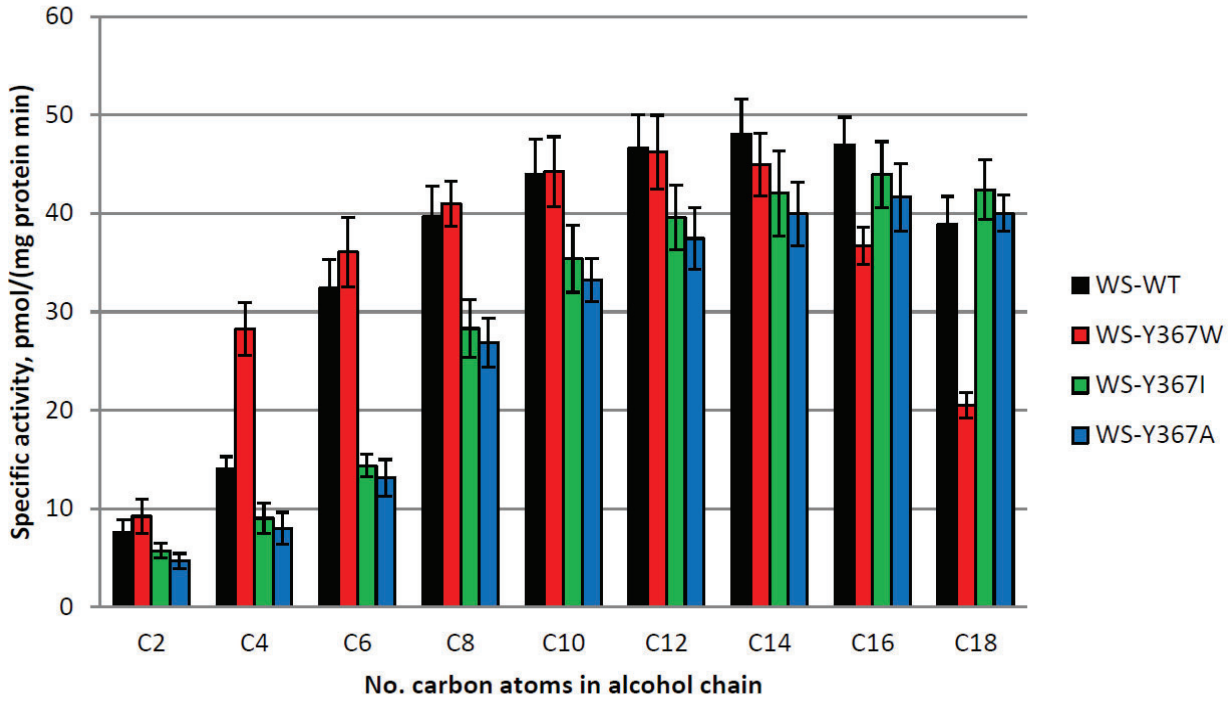


Figure 25. Alcohol selectivity of MhWS2 variants by single residue substitutions at residue position Y367. Specific enzyme activities with various alcohols as substrate, abbreviations: C2: ethanol, C4: 1-butanol, C6: n-hexanol, C8: n-octanol, C10: n-decanol, C12: n-dodecanol, C14: n-butadecanol, C16: n-hexadecanol, C18: n-octadecanol.

CHAPTER 5

Conclusions and Perspectives

CONCLUSIONS

By utilizing yeast *Saccharomyces cerevisiae*, it is possible to synthesize advanced microbial biodiesel (a.k.a. FAEEs) through heterologously expressing MhWS2 from *Marinobacter hydrocarbonoclasticus* (**Paper II**). Furthermore, by disrupting cells of the genes involved in the biochemical reactions of consuming free fatty acids, an increase in FAEEs can be achieved (**Paper III**). Moreover, through delta(δ) integration of the *ws2* gene to the yeast cell genome, an enhancement of *ws2* copy number can lead to a larger production (**Paper IV**).

Through engineering the enzymes MhWS2 and Eeb1p that catalyze the reaction of FAEEs formation, a more efficient synthesis can be obtained. By directed evolution of MhWS2, an increment in specific enzyme activity and alterations on selectivity to both substrates were obtained (**Paper V**). Whereas for Eeb1p, specific enzyme activity also increased while selectivity towards the fatty acyl molecule was modified, both linked to improvements in structure-function features mainly associated to a) the fore signal peptide, b) the lid-related initial section and c) the cap subdomain of this enzyme and its variants and d) residue substitutions T397P and L432F at high proximity to the active site (**Paper VI**). Rational engineering of MhWS2 can lead to an altered selectivity of the alcohol molecule being utilized as substrate, while FAEE production is conserved (**Paper VII**).

PERSPECTIVES

MhWS2 from *Marinobacter hydrocarbonoclasticus* is a wax ester synthase that has been naturally evolved to catalyze the biosynthesis of esters at high temperatures. Therefore, a scientific project regarding the investigation of this feature can lead to promising results. In this approach, *S. cerevisiae* cells harbouring the *ws2* gene, would be gradually adapted at higher temperatures in order to optimize the FAEE production.

The crystalization of MhWS2 and selected variants can be implemented for obtaining a highly reliable tertiary structure model. This would aim toward unraveling the requirements for increasing the biocatalytical performance toward FAEE formation.

Further studies must be done for deciphering the structure-function properties and features of the newly engineered selected enzyme variants, this will lead to a better comprehension regarding further explanations for these variants possessing a large number of residue substitutions at specific positions.

Crystalizing Eeb1p and variants would guide to a highly reliable tertiary structure model. This will improve precision for more firmly stating descriptions and features involved in the biocatalysis for ester production.

By acquiring more detailed structural descriptions of wild-types MhWS2, Eeb1p and selected variants, a possible future approach could be reached towards precisely discovering the biocatalytical process focusing on the mechanism of reaction for FAEE formation. Depending on the progress by these studies, possible clarifications could be obtained on the binding site mechanism of both substrates, where highly probable conformation changes would be found linked to the feasible generation of additional substrate binding sites, as suggested by the analysis of residue substitutions. Then, novel approaches for rational engineering of these biocatalyzers may be addressed.

The novel CRISPR-Cas9 gene editing biosynthetic tool can be employed for introducing any of the genes required for the FAEE production (including the wild-type enzymes and their selected variants). This could lead to a higher productivity and/or more efficient formation of microbial biodiesel.

A multidimensional engineering approach addressing several aspects regarding FAEE production, such as: enzyme improvement, metabolic pathway adaptation, and cultivation process optimization; all would represent an augment on the titer of FAEE formation, which is highly desired as for meeting

the future energy requirements demanding the next generation of transportation fuels.

8. REFERENCES

Alonso-Gutierrez, J., *et al.,* “Metabolic Engineering of *Escherichia coli* for limonene and perillyl alcohol production.” *Metabolic Engineering* **19**: 33-41 (2013).

Alvarez, H.M.*, et al.* "Accumulation and mobilization of storage lipids by *Rhodococcus opacus* PD630 and *Rhodococcus ruber* NCIMB 40126." *Applied Microbiology and Biotechnology* **54**(2): 218-223 (2000).

Athenstaedt, K., Daum, G. “YMR313c/*TGL3* encodes a novel triacylglycerol lipase located in lipid particles of *Saccharomyces cerevisiae*.” *The Journal of Biological Chemistry* **278**(26): 23317-23323 (2003).

Banerjee, A.*, et al.* "*Botryococcus braunii*: A Renewable Source of Hydrocarbons and Other Chemicals." *Critical Reviews in Biotechnology* **22**(3): 245-279 (2002).

Barney *et al*. “Identification of a residue affecting fatty alcohol selectivity in wax ester synthase” *Applied and Environmental Microbiology* **79**(1): 396-399 (2013).

Barney *et al*. “Altering small and medium alcohol selectivity in the wax ester synthase.” *Applied Microbiology and Biotechnology* **99**: 9675-9684 (2015).

Barth, S. *et al.* “Sequence and structure of epoxide hydrolases: a systematic analysis” *Proteins: Structure, Function and Bioinformatics* **55**(4): 846-855 (2004).

Basler, G, *et al*., “A *Pseudomonas putida* efflux pump acts on short-chain alcohols” *Biotechnology for Biofuels* **11**:136 (2018).

Bauer, T.L., *et al.* ”The modular structure of α/β hydrolases.” *The FEBS Journal* **287**: 1035-1053 (2019).

Becker, D. and L. Guarente. *“*High-efficiency transformation of yeast by electroporation.” *Methods in Enzymology*. C.G. GRF. New York, Academic Press. **194:** 182-187 (1991).

Beller, H.R., *et al*., ”Genes Involved in Long-Chain Alkene Biosynthesis in *Micrococcus luteus*.” *Applied and Environmental Microbiology* **76**(4): 1212-1223 (2010).

Beopoulos, A.*, et al.* "Control of lipid accumulation in the yeast *Yarrowia lipolytica*." *Applied and Environmental Microbiology* **74**(24): 7779-7789 (2008).

Bernard, A. *et al., “*Reconstitution of plant alkane biosynthesis in yeast demonstrates that Arabidopsis ECERIFERUM1 and ECERIFERUM3 are core

components of a very-long-chain alkane synthesis complex.” *Plant Cell* **24**: 3106–3118 (2012).
Branduardi, P., *et al*., “A novel pathway to produce butanol and isobutanol in *Saccharomyces cerevisiae*.” *Biotechnology for Biofuels* **6**:68 (2013).
Brennan, T., *et al.*, “Alleviating monoterpene toxicity using a two-phase extractive fermentation for the bioproduction of jet fuel mixtures in *Saccharomyces cerevisiae*.” *Biotechnology and Bioengineering* **109**(10): 2513-2522 (2012).
Brennan, T., *et al.*, “Physiological and Transcriptional Responses of *Saccharomyces cerevisiae* to d-Limonene Show Changes to the Cell Wall but Not to the Plasma Membrane.” *Applied and Environmental Microbiology* **79**(12): 3590-3601 (2013).
Brennan, T., *et al.*, “Evolutionary engineering improves tolerance for replacement jet fuels in *Saccharomyces cerevisiae*.” *Applied and Environmental Microbiology* **81**(10): 3316-3326 (2015).
Buijs, N.A., Siewers, V. & Nielsen, J. “Advanced biofuel production by the yeast *Saccharomyces cerevisiae*. *Current Opinion in Chemical Biology* **17**: 480–488 (2013).
Buijs, N.A., *et al.*, “Long-chain alkane production by the yeast *Saccharomyces cerevisiae.*” *Biotechnology and Bioengineering* **112**(6): 1275-1279 (2015).
Bzowka, M., *et al.* ”Evolution of tunnels in α/β-hydrolase fold proteins- What can we learn from studying epoxide hydrolases?” *PLOS Computational Biology* **18**(5) e1010119 (2022).
Cai, W., Zhang W., “Engineering modular polyketide synthases for production of biofuels and industrial chemicals” *Current opinion in biotechnology* **50**: 32-38 (2018).
Chen, X., *et al.* "Increased isobutanol production in *Saccharomyces cerevisiae* by overexpression of genes in valine metabolism.” *Biotechnology for Biofuels* **4**:21 (2011).
Chen, Y.*, et al.* "Enhancing the copy number of episomal plasmids in *Saccharomyces cerevisiae* for improved protein production." *FEMS Yeast Research* **12**: 598-607 (2012).
Cheng, J.B. and D.W. Russell. "Mammalian wax biosynthesis: II. Expression cloning of wax synthase cDNAs encoding a member of the acyltransferase enzyme family." *The Journal of Biological Chemistry* **279**(36): 37798–37807 (2004).
Corpet, F. "Multiple sequence alignment with hierarchical clustering." *Nucleic Acids Research* **16**: 10881-10890 (1988).

d'Espaux, L. *et al. "*Engineering high-level production of fatty alcohols by *Saccharomyces cerevisiae* from lignocellulosic feedstocks." *Metabolic Engineering* **42**: 115–125 (2017).
Denesyuk, A., *et al.* "The acid-base-nucleophile catalytic triad in ABH-fold enzymes is coordinated by a set of structural elements." *PLOS One* https://doi.org/10.1371/journal.pone.0229376 (2020).
Dmochowska, A.*, et al.* "Structure and transcriptional control of the *Saccharomyces cerevisiae* POX1 gene encoding acylcoenzyme A oxidase." *Gene* **88**(2): 247–252 (1990).
Duan, Y.*, et al.* "*De novo* Biosynthesis of Biodiesel by *Escherichia coli* in Optimized Fed-Batch Cultivation." *PLoS One* **6**(5): e20265 (2011).
Effmert, U., *et al*., "Volatile Mediated Interactions Between Bacteria and Fungi in the Soil." **38**: 665-703 (2012).
F.O. Licht. International Ethanol and Biofuels Report. **12:** 02 (2013).
Fretland, A.J., and Omiecinski, C.J. "Epoxide hydrolases: biochemistry and molecular biology." *Chemico-Biological Interactions* **129**(1-2): 41–59 (2000).
Guneser B.A., *et al.*, "Bioactives, Aromatics and Sensory Properties of Cold-Pressed and Hexane-Extracted Lemon (*Citrus Limon* L.) Seed Oils" *Journal of American Oil Chemistry Society* **94**: 723-731 (2017).
Harvey, B.G.*, et al.* "High-Density Renewable Fuels Based on the Selective Dimerization of Pinenes." *Energy & Fuels* **24**(1): 267-273 (2009).
Harvey, B.G.*, et al.* "High-Density Renewable Fuels Based on the Selective Dimerization of Pinenes." Patent US 2012/0205284 A1.
Heddergott, C., *et al.*, "The volatome of *Aspergillus fumigatus*" *Eukaryotic Cell* **13**(8): 1014-1026 (2014).
Heier, C., *et al.* "Identification of Yju3p as functional orthologue of mammalian monoglyceride lipase in the yeast *Saccharomyces cerevisiae*." *Biochimica et Biophysica Acta (BBA)* **1801**: 1063–1071 (2010).
Helalat, S.H., *et al.*, "Metabolic engineering of *Deinococcus radiodurans* for pinene production from glycerol" *Microbial Cell Factories* **20**(187): 1-14 (2021).
Herman, N.A. "Enzymes for fatty acid-based hydrocarbon biosynthesis." *Current Opinion in Chemical Biology* **35**: 22–28 (2016).
Hernández Lozada, N.J., *et al*. "Production of 1-octanol in *Escherichia coli* by a high flux thioesterase route" *Metabolic Engineering* **61**: 352-359 (2020).
Hernawan, T. and G. Fleet. "Chemical and cytological changes during the autolysis of yeasts." *Journal of Industrial Microbiology and Biotechnology* **14**: 440-450 (1995).
Holmquist, M. "Alpha beta-hydrolase fold enzymes structures, functions and mechanisms." *Current Protein and Peptide Science* **1**(2): 209-235 (2000).

Holtzapple, E. and C. Schmidt-Dannert. "Biosynthesis of isoprenoid wax ester in *Marinobacter hydrocarbonoclasticus* DSM 8798: Identification and characterization of isoprenoid coenzyme A synthetase and wax ester synthases." *Journal of Bacteriology* **189**: 3804-3812 (2007).
Hu, Y., Zhu, Z., Nielsen, J. & Siewers, V. “Heterologous transporter expression for improved fatty alcohol secretion in yeast.” *Metabolic Engineering* **45**: 51–58 (2018).
Hu, Z., *et al*., “Engineering *Saccharomyces cerevisiae* for production of the valuable monoterpene d-limonene during Chinese Baijiu fermentation.” *Journal of Industrial Microbiology and Biotechnology* **47**: 511-523 (2020).
Hung, R., *et al.*, “*Arabidopsis thaliana* as a model system for testing the effect of Trichoderma volatile organic compounds.” *Fungal Ecology* **6**(1): 19-26(2013).
Illman, A.M.*, et al.* "Increase in Chlorella strains calorific values when grown in low nitrogen medium." *Enzyme and Microbial Technology* **27**(8): 631-635 (2000).
Jetter, R. and L. Kunst. "Plant surface lipid biosynthetic pathways and their utility for metabolic engineering of waxes and hydrocarbon biofuels." *The Plant Journal* **54**: 670-683 (2008).
Jin, C.*, et al.* "Progress in the production and application of n-butanol as a biofuel." *Renewable and Sustainable Energy Reviews* **15**(8): 4080-4106 (2011).
Jongedijk, E., *et al.*, ”Capturing of the monoterpene olefin limonene produced in *Saccharomyces cerevisiae*“ *Yeast* **32**(1): 159-171 (2014).
Jongedijk, E., *et al.*, *“*Biotechnological production of limonene in microorganisms“ *Applied Microbiology and Biotechnology* **100**: 2927-2938 (2016).
Kalscheuer, R.*, et al.* "Synthesis of Novel Lipids in *Saccharomyces cerevisiae* by Heterologous Expression of an Unspecific Bacterial Acyltransferase." *Applied and Environmental Microbiology* **70**(12): 7119-7125 (2004).
Kalscheuer, R.*, et al.* "Microdiesel: *Escherichia coli* engineered for fuel production." *Microbiology* **152**(9): 2529-2536 (2006).
Kang, M.K., *et al.*, ”Functional screening of aldehyde decarbonylases for long-chain alkane production by *Saccharomyces cerevisiae*” *Microbial cell factories* **16**: 74 (2017).
Kelley, L.A. and M.J.E. Sternberg. "Protein structure prediction on the Web: a case study using the Phyre server." *Nature Protocols* **4**(3): 363-371 (2009).

Khozin-Goldberg, I.*, et al.* "Nitrogen starvation induces the accumulation of arachidonic acid in the freshwater green alga *Parietochloris incisa* (*trebuxiophyceae*)1." *Journal of Phycology* **38**(5): 991-994 (2002).
Kirby, J. and J.D. Keasling. “Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering.” *Annual Review of Plant Biology.* **60:** 335-355 (2009).
Knight, M.J., *et al.*, “The yeast enzyme Eht1 is an octanoyl-CoA: ethanol acyltransferase that also functions as a thioesterase.” *Yeast* **31**(12): 463-474 (2014).
Kondo, T., *et al.*, “Genetic engineering to enhance the Ehrlich pathway and alter carbon flux for increased isobutanol production from glucose by *Saccharomyces cerevisiae*.” *Journal of Biotechnology* **159**: 32-37 (2012).
Krivoruchko, A.*, et al.* "Improving biobutanol production in engineered *Saccharomyces cerevisiae* by manipulation of acetyl-CoA metabolism." *Journal of Industrial Microbiology & Biotechnology* **40**(9): 1051-1056 (2013).
Köffel, R., *et al.* ”The *Saccharomyces cerevisiae* YLL012/YEH1, YLR020/YEH2, and TGL1 genes encode a novel family of membrane-anchored lipases that are required for steryl ester hydrolysis.” *Molecular Cell Biology* **25**(5): 1655-1658 (2005).
Lardizabal, K.D.*, et al.* "Purification of a Jojoba Embryo Wax Synthase, Cloning of its cDNA, and Production of High Levels of Wax in Seeds of Transgenic Arabidopsis." *Plant Physiology* **122**: 645-655 (2000).
Leber, R.*, et al.* "Characterization of lipid particles of the yeast, *Saccharomyces cerevisiae*." *Yeast* **10**(11): 1421-1428 (1994).
Lee, F.W.F. and N.A.D. Silva. "Improved efficiency and stability of multiple cloned gene insertions at the δ sequences of *Saccharomyces cerevisiae*." *Applied Microbiology and Biotechnology* **48**(3): 339-345 (1997).
Lee, S.K.*, et al.* "Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels." *Current Opinion in Biotechnology* **19**(6): 556-563 (2008).
Lesage, P. and A.L. Todeschini. "Happy together: the life and times of Ty retrotransposons and their hosts." *Cytogenet. Genome Research* **110**(1-4): 70-90 (2005).
Li, Y.*, et al.* "High-density cultivation of oleaginous yeast *Rhodosporidium toruloides* Y4 in fed-batch culture." *Enzyme and Microbial Technology* **41**(3): 312-317 (2007).
Kaneko, T.*, et al.* "Loops govern SH2 domain specificity by controlling access to binding pockets." *Science Signaling* **3**(120): 1421-1428 (2010).

Liu, B. and Z. Zhao. "Biodiesel production by direct methanolysis of oleaginous microbial biomass." *Journal of Chemical Technology & Biotechnology* **82**(8): 775-780 (2007).
Liu, Q., *et al. "*Engineering an iterative polyketide pathway in Escherichia coli results in single-form alkene and alkane overproduction." *Metabolic Engineering* **28**: 82–90 (2015).
Lu, X.*, et al.* "Overproduction of free fatty acids in *E. coli*: Implications for biodiesel production." *Metabolic Engineering* **10**(6): 333-339 (2008).
MacDonald, B.*, et al.* "Determination of Sulfur in Biodiesel by X-Ray Fluorescence Spectroscopy." *National Institute of Standards and Technology* (2011).
Martin, V.J.J.*, et al.* "Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids." *Nature Biotechnology* **21**(7): 796-802 (2003).
Meadows, A.L., *et al.*, "Rewriting yeast central carbon metabolism for industrial isoprenoid production." **537**: 694-710 (2016).
Mendez-Perez, D., Begemann, M. B. & Pfleger, B. F. "Modular synthase-encoding gene involved in α-olefin biosynthesis in *Synechococcus* sp. strain PCC 7002." *Appl. Environ. Microbiol.* **77**(12): 4264–4267 (2011).
Meng, X.*, et al.* "Biodiesel production from oleaginous microorganisms." *Renewable Energy* **34**(1): 1-5 (2009).
Mitusinska, K., *et al.* "Structure-function relationship between soluble epoxide hydrolases structure and their tunnel network." *Computational and Structural Biotechnology Journal* **20**: 193-205 (2022).
Murphy, D.J. and J. Vance. "Mechanisms of lipid-body formation." *Trends Biochem. Sci.* **24**(3): 109-115 (1999).
Niu, F.X., *et al.* "Enhancing Production of Pinene in Escherichia coli by Using a Combination of Tolerance, Evolution, and Modular Co-culture Engineering" *Frontiers in Microbiology* **9**: 1623 (2018).
Niu, F.X., *et al.* "Genomic and transcriptional changes in response to pinene tolerance and overproduction in evolved *Escherichia coli*" *Synthetic and Systems Microbiology* **4**(3): 113-119 1623 (2019).
Noweck, K. and W. Grafahrend. Fatty Alcohols. Ullmann's Encyclopedia of Industrial Chemistry, Wiley-VCH Verlag GmbH & Co. KGaA (2000).
Ollis, D.L., *et al.* "The alpha/beta hydrolase fold." *Protein Engineering, Design and Selection* **5**(3): 197-211 (1992).
Peralta-Yahya, P.P.*, et al.* "Identification and microbial production of a terpene-based advanced biofuel." *Nature Communications* **2**: 483 (2011).

Petschnigg, J. *et al*. ”Good fat, essential cellular requirements for triacylglycerol synthesis to maintain membrane homeostasis in yeast.” *Journal of Biological Chemistry* **284**(45): 30981-30993 (2009).
Poust S., *et al*. “Divergent Mechanistic Routes for the Formation of gem-Dimethyl Groups in the Biosynthesis of Complex Polyketides” *Angewandte Chemie* **54**(8): 2370-2373 (2015).
Qiu, Y., *et al., “*An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis.” *Proc. Natl. Acad. Sci. U. S. A.* **109**: 14858–63 (2012).
Qureshi, N. and H.P. Blaschek. "Production of Acetone Butanol Ethanol (ABE) by a Hyper-Producing Mutant Strain of *Clostridium beijerinckii* BA101 and Recovery by Pervaporation." *Biotechnology Progress* **15**(4): 594-602 (1999).
Ratledge, C. and J. Wynn. "The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms." *Advances in Applied Microbiology* **51**: 1-44 (2002).
Rauwerdink, A., Kazlauskas, R.J., “How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: the Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes” *ACS Catalysis* **5**(10): 6153-6176 (2015).
Renewable Fuels Association. "2011 Ethanol Industry Outlook: Building Bridges to a More Sustainable Future**:** 2-3, 19-20 (2011).
Rottensteiner H., *et al*., “The ins and outs of peroxisomes: co-ordination of membrane transport and peroxisomal metabolism” *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1763**(12): 1527-1540 (2006).
Rui, Z. *et al. “*Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase.” *Proc. Natl. Acad. Sci.* **111**: 18237–18242 (2014).
Rui, Z., *et al*. “Discovery of a Family of Desaturase-Like Enzymes for 1-Alkene Biosynthesis.” *ACS Catal.* **5**: 7091–7094 (2015).
Röttig, A. and A. Steinbüchel. "Random mutagenesis of atfA and screening for *Acinetobacter baylyi* mutants with an altered lipid accumulation." *European Journal of Lipid Science and Technology* **115**(4): 394-404 (2013a).
Röttig, A. and A. Steinbüchel. ”Acyltransferases in bacteria.” *Microbiology and Molecular Biology Reviews* **77**(2): 277-321 (2013b).
Saerens, S.M.*, et al.* "The *Saccharomyces cerevisiae* EHT1 and EEB1 genes encode novel enzymes with medium-chain fatty acid ethyl ester synthesis and hydrolysis capacity." *J. Biol. Chem.* **281**(7): 4446-4456 (2006).
Sandager, L.*, et al.* "Storage lipid synthesis is non-essential in yeast." *J. Biol. Chem.* **277**(8): 6478 (2002).
Sarria, S., *et al.*, “Microbial synthesis of pinene” *ACS Synthetic Biololgy* **3**(7): 466-475 (2014).

Schirmer, A.*, et al.* "Microbial biosynthesis of alkanes." *Science* **329**(5991): 559-562 (2010).
Selvaraju, K.*, et al.* "*MGL2/YMR201w* encodes a monoacylglycerol lipase in *Saccharomyces cerevisiae*." *FEBS Letters* **590**: 1174-1186 (2016).
Shaojin, Y. and Z. Yioing. "Research and Application of Oleaginous Microorganism." *China Foreign Energy* **11**: 90-94 (2006).
Shi, S.*, et al.* "Prospects for microbial biodiesel production." *Biotechnology Journal* **6**: 277-285 (2011).
Shi, S.*, et al.* "Functional expression and characterization of five wax ester synthases in *Saccharomyces cerevisiae* and their utility for biodiesel production." *Biotechnology for Biofuels* **5**: 7-16 (2012).
Si, T., *et al*. "Utilizing and endogenous pathway for 1-butanol production in *Saccharomyces cerevisiae.*" *Metabolic Engineering* **22**: 60-68 (2014).
Siloto, R.M.P. *et al*., "Simple methods to detect triacylglycerol biosynthesis in a yeast-based recombinant system." *Lipids* **44**: 963 (2009).
Siu, Y., *et al.*, "Design and selection of a synthetic feedback loop for optimizing biofuel tolerance." *ACS Synthetic Biology* **7**(1): 16-23 (2018).
Slocombe, S.P.*, et al.* "Oil accumulation in leaves directed by modification of fatty acid breakdown and lipid synthesis pathways." *Plant Biotechnol. J.* **7**(7): 694-703 (2009).
Sorigué, D. *et al. "*An algal photoenzyme converts fatty acids to hydrocarbons." *Science* **357**: 903–907 (2017).
Srirangan, K., *et al.*, "Engineering *Escherichia coli* for microbial production of butanone." *Appl. Environ. Microbiol.* **82**(9): 2574-2585 (2016).
Steen, E.*, et al.* "Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol." *Microbial Cell Factories* **7**: 36 (2008).
Steen, E.*, et al.* "Microbial production of fatty-acid-derived fuels and chemicals from plant biomass." *Nature* **463**(7280): 559-562 (2010).
Stöveken, T.*, et al.* "The Wax Ester Synthase/Acyl Coenzyme A:Diacylglycerol Acyltransferase from Acinetobacter sp. Strain ADP1: Characterization of a Novel Type of Acyltransferase " *Journal of Bacteriology* **187**(4): 1369-1376 (2005).
Stöveken, T. and A. Steinbuchel. "Both histidine residues of the conserved HHXXXDG motif are essential for wax ester synthase/acyl-CoA:diacylglycerol acyltransferase catalysis." *European Journal Lipid Science Technology* **111**: 112-119 (2009).
Sukovich, D.J.*, et al.* "Widespread Head-to-Head Hydrocarbon Biosynthesis in Bacteria and Role of OleA." *Applied and Environmental Microbiology* **76**(12): 3850-3862 (2010).

Tan, X., *et al.* "Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria." *Metabolic Engineering* **13**(2): 169-176 (2011).
Tashiro, M., *et al.* "Bacterial production of pinene by a laboratory-evolved pinene-synthase" *ACS Synthetic Biology* **5**(9): 1011-1020 (2016).
Thomas, W., *et al.* "Screening for Lipid Yielding Microalgae: Activities for 1983." *SERI/STR-231-2207*: p. 31 (1984).
Tsukahara, K., *et al.* "Treatment of liquid fraction separated from liquidized food waste in an upflow anaerobic sludge blanket reactor." *Journal of Bioscience and Bioengineering* **87**(4): 554-556 (1999).
UNH Biodiesel Group (2010). from http://www.unh.edu/p2/biodiesel/article_alge.html.
Valle-Rodríguez, J.O., *et al.* "Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid ethyl esters, an advanced biofuel, by eliminating non-essential fatty acid utilization pathways." *Applied Energy* **115**(0): 226-232 (2014).
Valle-Rodríguez, J.O., *et al*. "Directed evolution of a wax ester synthase for production of fatty acid ethyl esters in *Saccharomyces cerevisiae." Applied Microbiology and Biotechnology*; 107:2921-2932 (2023).
Varthini, V.L., *et al.* "ROG1 encodes a monoacylglycerol lipase in *Saccharomyces cerevisiae*." *FEBS letters* **589**: 23-30 (2015).
Villa, J.A. *et al.* "Use of limited proteolysis and mutagenesis to identify folding domains and sequence motifs critical for wax ester synthase/acyl coenzyme A: diacylglycerol acyltransferase activity." *Applied and Environmental Microbiology* **80**(3): 1132-1142 (2014).
Walther, T.C. and R.V. Farese Jr. "The life of lipid droplets." *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1791**(6): 459-466 (2009).
Wang, X., *et al.* "G418 selection and stability of cloned genes integrated at chromosomal δ sequences of *Saccharomyces cerevisiae*." *Biotechnol. Bioeng.* **49**(1): 45-51 (1996).
Wei, L.J., *et al.,* "Increased Accumulation of Squalene in Engineered *Yarrowia lipolytica* through Deletion of *PEX10* and *URE2*" *Applied and Environmental Microbiology* **87**(17) (2021).
Westfall, P.J. and T.S. Gardner, "Industrial fermentation of renewable diesel fuels." *Current Opinion in Biotechnology* **22**(3): 344-350 (2011).
Westfall, P.J., *et al.* "Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent

artemisinin." *Proceedings of the National Academy of Sciences* **109**(3): E111-E118 (2012).
Wu, X., *et al*., “Biosynthesis of pinene in purple non-sulfur photosynthetic bacteria” *Microbial Cell Factories* **20**(101) (2021).
Vicente, G.*, et al.* "Direct Transformation of Fungal Biomass from Submerged Cultures into Biodiesel." *Energy & Fuels* **24**(5): 3173-3178 (2010).
Wältermann, M.*, et al.* "Key enzymes for biosynthesis of neutral lipid storage compounds in prokaryotes: Properties, function and occurrence of wax ester synthases/acyl-CoA:diacylglycerol acyltransferases." *Biochimie* **89**(2): 230-242 (2007).
Yan, Y. and J.C. Liao "Engineering metabolic systems for production of advanced fuels." *Journal of Industrial Microbiology & Biotechnology* **36**(4): 471-479 (2009).
Yang, J. *et al.,* “Metabolic engineering of *Escherichia coli* for the biosynthesis of alpha-pinene” *Biotechnology for biofuels* **6**(60): 1-10 (2013).
Yang, H., Wong, M.W., “Oxyanion Hole Stabilization by C–H···O Interaction in a Transition State—A Three-Point Interaction Model for Cinchona Alkaloid-Catalyzed Asymmetric Methanolysis of meso-Cyclic Anhydrides” Journal of the American Chemical Society **135**(15) (2013).
Youngquist, J.T., *et al*., “Production of medium chain length fatty alcohols from glucose in *Escherichia coli*” *Metabolic Engineering* **20**: 177-186 (2013).
Yuzawa, S.*, et al.* "Heterologous production of polyketides by modular type I polyketide synthases in *Escherichia coli*." *Current Opinion in Biotechnology* **23**(5): 727-735 (2012).
Yuzawa, S.*, et al.* “Insights into polyketide biosynthesis gained from repurposing antibiotic-producing polyketide synthases to produce fuels and chemicals” *The Journal of antibiotics* **69**: 494-499 (2016).
Yuzawa, S.*, et al.* “Short-chain ketone production by engineered polyketide synthases in *Streptomyces albus*” *Nature Communications* **9**: 4569 (2018).
Zahedi, R.P.*, et al.* “Proteomic analysis of the yeast mitochondrial outer membrane reveals accumulation of a subclass of preproteins” *Molecular Biology of the Cell* **17**: 1436-1450 (2006).
Zargar A, *et al*., “Engineering polyketide synthases for production of designer biofuels” *Chemical Society* **1155** (2017).
Zeldin, D.C., *et al.* " Metabolism of epoxyeicosatrienoic acids by cytosolic epoxide hydrolase: substrate structural determinants of asymmetric catalysis." *Archives of Biochemistry and Biophysics* **316**(1):443–451 (1995).
Zhao, Y., *et al.* “High-efficiency production of bisabolene from waste cooking oil by metabolically engineered *Yarrowia lipolytica*.“ **14**(6): 2497-2513 (2021).

Zhou, Y. J., *et al.*"Harnessing Yeast Peroxisomes for Biosynthesis of Fatty-Acid-Derived Biofuels and Chemicals with Relieved Side-Pathway Competition." *J. Am. Chem. Soc.* **138**: 15368–15377 (2016).
Zhu, Z. *et al.* "Enabling the synthesis of medium chain alkanes and 1-alkenes in yeast." *Metabolic Engineering* **44**: 81–88 (2017).